Do First-Class Constraints Generate Gauge Transformations? A Geometric Resolution*

Clara Bradley

Department of Logic and Philosophy of Science,

University of California, Irvine

Abstract

The standard definition of a gauge transformation in the constrained Hamiltonian formalism traces back to Dirac (1964): a gauge transformation is a transformation generated by an arbitrary combination of first-class constraints. On the basis of this definition, Dirac argued that one should extend the form of the Hamiltonian in order to include all of the gauge freedom. However, there have been some recent dissenters of Dirac's view. Notably, Pitts (2014) argues that a first-class constraint can generate "a bad physical change" and therefore that extending the Hamiltonian in the way suggested by Dirac is unmotivated. In this paper, I use a geometric formulation of the constrained Hamiltonian formalism to argue that there is a flaw in the reasoning used by both sides of the debate, but that correct reasoning supports the standard definition and the extension to the Hamiltonian. In doing so, I clarify two conceptually different ways of understanding gauge transformations, and I pinpoint what it would take to deny that the standard definition is correct.

1 Introduction

Gauge transformations represent local symmetries in physics that are often taken to indicate arbitrariness in the mathematical formalism of a theory. How to interpret this arbitrariness

*Draft of August 16, 2024.

is widely disputed and is connected to a wider literature on "surplus structure" in physics.[1] However, there is a different kind of dispute about gauge transformations that will be the focus here: if gauge transformations are conceptualized as transformations that indicate arbitrariness, what is the correct *formal* definition of a gauge transformation?

There is a longstanding tradition of using a formalism known as the "constrained Hamiltonian formalism" to establish the gauge transformations of a theory. The standard definition arising from the formalism is attributed to Dirac (1964): a gauge transformation is a transformation generated by an arbitrary combination of the first-class constraints, where constraints are relationships on the Hamiltonian variables that restrict the dynamically-allowed states and first-class constraints are those that "commute" with each other.[2] This definition is taken to have important consequences for the formulation of a Hamiltonian theory. In particular, Dirac argued on the basis of this definition that the Hamiltonian that generates the dynamics should be understood as an equivalence class of Hamiltonians, called the "Extended Hamiltonian", in order to include all of the gauge freedom.

However, there have been several recent dissenters of Dirac's view of gauge transformations. For example, Pitts (2014*b*) argues, using the example of Electromagnetism, that a first-class constraint does not generate a gauge transformation but rather "a bad physical change". Similarly, Pons (2005) argues that Dirac's analysis of gauge transformations is "incomplete" since it does not provide an accurate account of the symmetries between solutions to the equations of motion. Both authors conclude that extending the Hamiltonian in the way suggested by Dirac is unmotivated. If correct, these arguments could have implications for other issues in the foundations of the constrained Hamiltonian formalism. Notably, there is a puzzle called the "Problem of Time" that arises in the constrained Hamiltonian formalism for theories that are time-reparameterization invariant when one adopts the standard definition of a gauge transformation. If gauge transformations are not given by the standard definition, then this could be an avenue to avoiding the Problem of Time.[3]

More recently, Pooley & Wallace (2022) argue, contra Pitts (2014*b*), that the Extended

[1] For more on the notion of surplus structure and its connection to symmetries of a theory, see, for example, Ismael & Van Fraassen (2003), Earman (2004), Baker (2010).

[2] This will be made more precise in Section 2.

[3] See Pitts (2014*a*) for a response of this kind. For an introduction to the Problem of Time and its philosophical implications, see Thébault (2021).

Hamiltonian formalism is empirically equivalent to the non-extended formalism for Electromagnetism, and therefore that Dirac's orthodoxy is vindicated for this theory. In this paper, I extend the observations of Pooley & Wallace (2022) by arguing that the Extended Hamiltonian formalism is motivated on theoretical grounds. In more detail, I use a standard geometric formulation of the constrained Hamiltonian formalism to show that the extension to the Hamiltonian can be motivated independently from consideration of the gauge transformations, and, under the dynamics generated by this extended Hamiltonian, the standard account of gauge transformations as being generated by arbitrary first-class constraints is correct. In doing so, I argue that there is a common assumption made about the relationship between gauge transformations and the form of the Hamiltonian that is part of the source of the debate, but that this assumption is unnatural in the geometric framework. This leads to a revised account of the definition of gauge transformations in the constrained Hamiltonian formalism that sheds light on a particular source of contention: what the relationship is between gauge transformations on *states* and gauge transformations on *solutions*.

The paper will go as follows. In Section 2, I present Dirac's version of the constrained Hamiltonian formalism and his argument that arbitrary combinations of first-class constraint generate gauge transformations. In Section 3, I spell out the example that Pitts (2014*b*) gives as a counterexample to Dirac's view. In Section 4, I discuss where the disagreement lies between Dirac and Pitts' views, and I highlight a crucial assumption made on both sides of the debate. In Section 5, I consider the response to Pitts (2014*b*) given by Pooley & Wallace (2022) and pinpoint the way in which it fails to provide a complete response. In Section 6, I present the geometric formulation of the constrained Hamiltonian formalism, and I use this formulation in Section 7 to argue that the issue in the debate lies in the way that gauge transformations are both understood and motivated. In Section 8, I consider two possible counterarguments, before concluding.

2 Dirac's Theory

Dirac's version of the constrained Hamiltonian formalism is constructed by starting with the Lagrangian formalism. In the Lagrangian framework, one has a finite N number of degrees of freedom $q_n, n = 1, ..., N$, with corresponding velocities $\frac{dq_n}{dt} = \dot{q}_n$, where we assume an independent time variable t.[4] The dynamics are given by specifying a Lagrangian $L = L(q_n, \dot{q}_n)$ with corresponding action $I = \int L(q_n, \dot{q}_n)dt$, from which one derives the equations of motion called the Euler-Lagrange equations:

$$\frac{d}{dt}\frac{\partial L(q_n, \dot{q}_n)}{\partial \dot{q}_n} = \frac{\partial L(q_n, \dot{q}_n)}{\partial q_n}$$

To move to the Hamiltonian framework, one introduces "canonical momenta" variables $p_n = \frac{\partial L}{\partial \dot{q}_n}$. When these momenta are not independent of each other, there are constraints of the form $\phi_m(q_n, p_n) \approx 0$ for $m = 1, ..., M$ where M is the number of constraints and the meaning of $\approx$ is that the relationship holds *weakly*: one can substitute the left hand side for the right only on the subspace where the equation holds. This means in practical terms that one must evaluate any expression involving $\phi_m(q_n, p_n)$ and its derivatives before setting $\phi_m(q_n, p_n) = 0$. Constraints of this kind are called the *primary constraints*.

The *Hamiltonian* is defined as $H(q_n, p_n) = p^n q_n - L$ where we implicitly have a sum over n. However, it is not uniquely defined when the system is constrained, since one can add a linear combination of primary constraints and it will weakly be the same Hamiltonian. We call the addition of this linear combination of primary constraints the *Total Hamiltonian*, $H_T = H + u^m \phi_m$ where u^m are arbitrary functions of the canonical variables and again we implicitly have a sum over m. Thus, the Total Hamiltonian should be thought of as an equivalence class of Hamiltonians, differing over the choices of u^m. From the variation in H_T, one can derive Hamilton's equations of motions with constraints:

$$\dot{q}_n = \frac{\partial H}{\partial p_n} + u^m \frac{\partial \phi_m}{\partial p_n}$$

$$\dot{p}_n = -\frac{\partial H}{\partial q_n} - u^m \frac{\partial \phi_m}{\partial q_n}$$

More generally, for any dynamical variable g, $\dot{g} \approx \{g, H\} + u^m \{g, \phi_m\} = \{g, H_T\}$ where $\{\}$ is the Poisson bracket, satisfying the following properties:

[4] In order to consider the Problem of Time, it is useful to drop this assumption and treat the time variable as an additional dynamical variable, but we keep this assumption for the purposes here.

1. Leibniz rule: for any functions f, g, h, $\{fg, h\} = f\{g, h\} + g\{f, h\}$.

2. Jacobi identity: for any functions f, g, h, $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$.

3. Constant function: If k is a constant, then for any function f, $\{k, f\} = 0$.

In order for the solutions to the equations of motion to be consistent with the primary constraints, in the sense that the primary constraints hold at all times along a solution to the equations of motion, it ought to be the case that $\dot{\phi}_m \approx 0$. In other words, it ought to be the case that $\{\phi_m, H\} + u^{m'}\{\phi_m, \phi_{m'}\} \approx 0$. For each m, this equation either is identically satisfied with the primary constraints, reduces to an equation independent of the u's of the form $\chi_k(q_n, p_n) \approx 0$, or it imposes conditions on the u's.

In the second case, we say that $\chi_k(q_n, p_n) \approx 0$ are *secondary constraints*, since they arise from applying the equations of motion to the primary constraints. If we have a secondary constraint, then we get new consistency conditions by requiring $\dot{\chi}_k \approx 0$, which is again one of the three kinds above. One can continue this process until one has found all of the secondary constraints and one is left with the consistency conditions of the third kind. We can combine the primary and secondary constraints, writing them as $\phi_j \approx 0$ for $j = 1, ..., M + K$ where K is the number of secondary constraints.

For the remaining consistency conditions that do not reduce, we can find solutions $u^m = U^m(q_n, p_n)$ up to $V^m(q_n, p_n)$ where $V^m\{\phi_j, \phi_m\} \approx 0$, giving the general solution $u^m = U^m + v^a V_a^m$ where v^a is arbitrary.

Substituting into the Total Hamiltonian, we get

$$H_T = H' + v^a \phi_a$$

where $H' = H + U^m \phi_m$ and $\phi_a = V_a^m \phi_m$. Notice that we have satisfied all the consistency conditions but still have coefficients v^a that are arbitrary functions of the canonical variables.

A dynamical variable $R(q_n, p_n)$ is said to be *first-class* if $\{R, \phi_j\} \approx 0$. In other words, a dynamical variable is first-class if the Poisson bracket with any constraint equals a linear function of the constraints. If it is not first-class, it is called *second-class*. Importantly, H' and ϕ_a are first-class. This means that H_T is an equivalence class of Hamiltonians given by a sum of a first-class Hamiltonian and a linear combination of primary, first-class constraints.

Given some initial state $(q_n(t_0), p_n(t_0))$, the q's and p's at later times are underdetermined because of the arbitrariness in the coefficients v^a. One might take this to be a mark of *indeterminism* in the theory: there are multiple possible evolutions from an initial state. However, we might also think that this indeterminism is an artifact of our mathematical description, in that it indicates that our theory contains "redundancy". It is this direction of thought that led Dirac to propose the following definition of a gauge transformation:

> **State Gauge Transformation**: A gauge transformation relates any two states that are possible evolutions from an initial state under the dynamics generated by the Total Hamiltonian at some fixed (infinitesimal) interval δt.

In other words, Dirac proposes that physically equivalent states as precisely those that result from the arbitrariness in v^a in evolving a system's state. We can determine these transformations in the following way:

For a given dynamical variable g with initial value g_0, its value after some infinitesimal δt under a specified choice of coefficients v^a is:

$$g(\delta t) = g_0 + \dot{g}\delta t = g_0 + \{g, H_T\}\delta t = g_0 + \delta t[\{g, H'\} + v^a\{g, \phi_a\}] \tag{1}$$

However, one could have made different choices for v^a. Call another set of choices v'^a. The difference between the two values for g at δt under these two choices of coefficients is given by:

$$\Delta g(\delta t) = \delta t(v^a - v'^a)\{g, \phi_a\} = \varepsilon^a\{g, \phi_a\} \tag{2}$$

where ε^a is an arbitrary small number. This change will describe the same physical state, since it corresponds to a change from one state to another that arises merely from a different choice of arbitrary coefficient in the evolution from some initial state. Since ϕ_a are just the primary first-class constraints, Dirac concludes:

> All primary first-class constraints generate gauge transformations.

However, this isn't the end of the story. If we apply a second transformation generated by $\kappa^{a'}\phi_{a'}$ for arbitrary coefficients $\kappa^{a'}$, we get $\Delta g = \varepsilon^a\kappa^{a'}\{g, \{\phi_a, \phi_{a'}\}\}$, which also won't change the physical state since it is made up of transformations that don't change the physical state. Therefore, $\{\phi_a, \phi_{a'}\}$ is another generating function of a gauge transformation. The ϕ_a's are first-class constraints, and the Poisson bracket of two first-class quantities is first-class, so this generating function is a first-class constraint. However, it need not be a primary first-class constraint; it could be a *secondary* first-class constraint. Observing this, Dirac presents the following conjecture:

> **Dirac Conjecture**: All secondary first-class constraints generate gauge transformations.

We therefore have the following criteria for a State Gauge Transformation:

> Arbitrary combinations of first-class constraints generate a gauge transformation.

However, we are now in a situation where the dynamics is given by the Total Hamiltonian, which includes the arbitrariness associated with the primary first-class constraints, but we also have arbitrariness associated with the secondary first-class constraints. This mismatch between the dynamics and the arbitrariness led Dirac to suggest that one should also add the first-class secondary constraints to the Total Hamiltonian, giving rise to the *Extended Hamiltonian*, $H_E = H_T + w^b\chi_b$ where χ_b are the first-class secondary constraints and w^b are arbitrary functions of the canonical variables. The equations of motion then read: $\dot{g} = \{g, H_E\}$.

Finally, we can define an *observable* as a function f that has the property that $\{f, \gamma_j\} \approx 0$ for all first-class constraints γ_j. Observables are functions that are gauge-invariant, in the sense that they take the same value under the transformations generated by the first-class constraints. On the other hand, the *gauge variables* are the functions that are not observables.

The final picture of Dirac's theory is:

1. The symmetries of the theory are "State Gauge Transformations" that are generated by arbitrary combinations of first-class constraints.

2. The dynamics is generated by an equivalence class of Hamiltonians represented by the Extended Hamiltonian.

Whether this picture is correct will be the subject of the rest of the paper.

3 An Argument Against Dirac

Although Dirac's formalism and the notion of a gauge transformation that arises from it have been widely accepted as the standard framework, there have been several recent arguments in the literature that Dirac's picture is flawed.[5] Here, I focus on a supposed counterexample to Dirac's picture argued for by Pitts (2014*b*) of classical Electromagnetism.

The Lagrangian for classical Electromagnetism can be written in observer-dependent form as

$$\mathcal{L}(\vec{A}, V; \dot{\vec{A}}, \dot{V}) = \int \frac{1}{2}(\dot{\vec{A}} - \nabla V)^2 - \frac{1}{2}(\nabla \times \vec{A})^2 - (V\rho + \vec{A} \cdot \vec{J})$$

where $\vec{A}$ and V are time-dependent functions on $\mathbb{R}^3$ and the integral is over $\mathbb{R}^3$. The conjugate momenta are $p_{\vec{A}} = \frac{\delta L}{\delta \dot{\vec{A}}} = \dot{\vec{A}} - \nabla V$ and $p_V = \frac{\delta L}{\delta \dot{V}} = 0$. This means that there is one primary constraint, $\phi_0 = p_V$. The Total Hamiltonian is:

$$H_T = \int \frac{1}{2}(p_{\vec{A}}^2 + \vec{B}^2) + \lambda p_V + p_{\vec{A}} \cdot \nabla V + (V\rho + \vec{A} \cdot \vec{J}) \tag{3}$$

where the integral is over $\mathbb{R}^3$ and λ is an arbitrary function of the canonical coordinates. Integrating by parts with appropriate boundary conditions, we can rewrite the Total Hamiltonian as:

$$H_T = \int \frac{1}{2}(p_{\vec{A}}^2 + \vec{B}^2) + \vec{A} \cdot \vec{J} + \lambda p_V - V(\nabla \cdot p_{\vec{A}} - \rho) \tag{4}$$

We can then find the evolution of the primary constraint:

$$\{p_V, H_T\} = \frac{\delta H}{\delta V} = \nabla \cdot p_{\vec{A}} - \rho. \tag{5}$$

So there is a secondary constraint given by $\phi_1 = \nabla \cdot p_{\vec{A}} - \rho$. The evolution of the secondary constraint is zero, so there are two constraints in total, and both constraints are first-class.

The equations of motion for $\vec{A}$ and V are given by:[6]

[5] See in particular Pitts (2014*a*,*b*) and Pons (2005) but also Pons et al. (1997) and Barbour & Foster (2008).

[6] We leave out the equations of motion for $p_{\vec{A}}$ and p_V for convenience, since they aren't important for the argument.

$$\begin{aligned}\frac{\partial \vec{A}}{\partial t} &= \{\vec{A}, H_T\} = \frac{\partial H_T}{\partial p_{\vec{A}}} = p_{\vec{A}} + \nabla V \\ \frac{\partial V}{\partial t} &= \{V, H_T\} = \frac{\partial H_T}{\partial p_V} = \lambda\end{aligned} \tag{6}$$

The question that Pitts (2014*b*) asks is whether the arbitrary combinations of the primary and secondary constraint generate gauge transformations for these equations. In other words, we want to know whether, if $(\vec{A}(t), V(t); p_{\vec{A}}(t), p_V(t))$ satisfies these equations of motion, then transforming this solution by an arbitrary combination of the first-class constraints, $\int \alpha\phi_0 + \beta\phi_1$, also satisfies the equations of motion, where α and β are arbitrary functions of the canonical coordinates and time.

We have that:

$$\begin{aligned}\{\vec{A}, \int \alpha\phi_0 + \beta\phi_1\} &= \{\vec{A}, \int \alpha p_V + \beta(\nabla \cdot p_{\vec{A}} - \rho)\} \\ &= \{\vec{A}, \int \alpha p_V\} + \{\vec{A}, \int \beta(\nabla \cdot p_{\vec{A}} - \rho)\}\end{aligned} \tag{7}$$

The first term vanishes. Since $\int \beta\nabla \cdot p_{\vec{A}} = -\int p_{\vec{A}} \cdot \nabla\beta$ by integration by parts (with appropriate boundary conditions), the second term is equal to $\{\vec{A}, -\int p_{\vec{A}} \cdot \nabla\beta + \beta\rho)\} = \nabla\beta$. Therefore, the transformed quantity is given by $A' = A + \nabla\beta$.

Similarly:

$$\{V, \int \alpha\phi_0 + \beta\phi_1\} = \{V, \int \alpha p_V\} + \{V, \int \beta(\nabla \cdot p_{\vec{A}} - \rho)\} \tag{8}$$

The second term here vanishes, and the first term is equal to α. Thus, the transformed potential is given by $V' = V + \alpha$.

We also have that $\{p_{\vec{A}}, \int \alpha p_V + \beta(\nabla \cdot p_{\vec{A}} - \rho)\} = \{p_V, \int \alpha p_V + \beta(\nabla \cdot p_{\vec{A}} - \rho)\} = 0$ and so the conjugate momenta do not change under the transformation generated by an arbitrary combination of the constraints. We can therefore write the transformed equations of motion for $\vec{A}$ and V as:

$$
\begin{aligned}
\frac{\partial \vec{A}'}{\partial t} &= \frac{\partial \vec{A}}{\partial t} + \frac{\partial \nabla \beta}{\partial t} = p_{\vec{A}} + \nabla(V + \alpha) \\
\frac{\partial V'}{\partial t} &= \frac{\partial V}{\partial t} + \frac{\partial \alpha}{\partial t} = \lambda
\end{aligned}
\tag{9}
$$

Since we assumed that $\frac{\partial \vec{A}}{\partial t} = p_{\vec{A}} + \nabla V$, the first equation is satisfied only when $\frac{\partial \nabla \beta}{\partial t} - \nabla \alpha = 0$. In particular, in the case where either α or β is zero (where one considers the transformation generated by only one of the primary or secondary constraints), the first equation is not satisfied.

On the basis of this argument, Pitts (2014*b*) concludes that arbitrary combinations of first-class constraints do not generate gauge transformations. Rather, only a particular combination of first-class constraints generates a gauge transformation. So, the argument goes, Dirac was wrong about what the gauge transformations are.

Remember also that the motivation for Dirac to move to the "Extended Hamiltonian" was precisely that secondary first-class constraints generate gauge transformations in addition to primary first-class constraints. But the above argument shows that this is not strictly true: in fact, there are only as many arbitrary functions of time as there are primary first-class constraints. To see this, notice that since $\nabla \alpha = \frac{\partial \nabla \beta}{\partial t}$, we can write the gauge transformations as being generated by $\int \dot{\epsilon}\phi_0 + \epsilon\phi_1$. In other words, we only need one arbitrary function (and its time derivative) to specify the gauge transformations. Therefore, one might also take this argument to show that the Extended Hamiltonian is not motivated, or more strongly, that the Extended Hamiltonian is the wrong equivalence class of Hamiltonians, since it suggests that there is "more" arbitrariness in the dynamics than there in fact is.

4 Where The Disagreement Lies

There is an immediate sense in which the above argument fails on its own to show that Dirac was wrong. In Section 2, we interpreted Dirac as giving an account of what I called "State Gauge Transformation": transformations relating two states that are possible evolutions from some initial state. However, the argument I just ran, following Pitts (2014*b*), doesn't consider whether two *states* are equivalent; it considers whether two *solutions* are equivalent. That is, it considers whether arbitrary combinations of first-class constraints generate a transformation that takes one from a solution to the equations of motion to another solution. We might alternatively call this notion of a gauge transformation "Solution Gauge Transformation":

> **Solution Gauge Transformation**: A gauge transformation relates any two states that are possible evolutions from an initial state under the dynamics generated by the Total Hamiltonian *at any time* t.

What Pitts' argument demonstrates is that the Solution Gauge Transformations are not generated by arbitrary combinations of first-class constraints in the context of classical Electromagnetism. Indeed, arbitrary combinations of first-class constraints *do* generate State Gauge Transformations in classical Electromagnetism. To see this, recall that we can write the Solution Gauge Transformations as $\int \dot{\epsilon}\phi_0 + \epsilon\phi_1$. At a particular fixed time, ϵ and $\dot{\epsilon}$ become independent of each other. And so, we can write the State Gauge Transformations as $\int \alpha\phi_0 + \beta\phi_1$, as would be the case if arbitrary combinations of first-class constraints generate gauge transformations. So what Pitts (2014*b*) shows is that Solution Gauge Transformations do not always match the State Gauge Transformations.

At this point one might want to say: what this shows is that we really have two distinct notions of a gauge transformation, 'State Gauge Transformation' and 'Solution Gauge Transformation', and it turns out that these notions do not coincide. This would suggest that there is not really a debate here at all; different parties in the debate are just focusing on different notions, and we can accept that both are right.

Although *formally* this thought seems correct, there is a *conceptual* issue with accepting both notions of a gauge transformation, since it would mean accepting that individual states along two curves can be gauge-equivalent without it being the case that if one curve is a solution, then the other also is. The reason is that the transformations that generate Solution Gauge Transformations are more restrictive than (are a subset of) those that generate State Gauge Transformations. But if gauge equivalence is supposed to mean *physical* equivalence, then this would be to say that two curves can be such that each individual state along one curve is physically equivalent to a state along the other curve but the curves *as a whole* are not physically equivalent to one another. Conceptually, this is not coherent: solutions just consist of a series of states, and so if all of these states are physically equivalent to some other series of states, then the solutions ought to also be physically equivalent.

Therefore, it seems that if one wants to accept that "Solution Gauge Transformation" is the right definition of gauge transformations on solutions and that gauge equivalence is a notion of physical equivalence, one has to accept that there is no independent notion of a gauge transformation on states. That is, any notion of a state gauge transformation must be derivative to that of the solution gauge transformations: a state gauge transformation must be the special case of the solution gauge transformations where the solutions are considered to be infinitesimally short in terms of time.

This helps to set up the rest of the paper: I will argue that one can maintain separate notions of state and solution gauge transformations as notions of physical equivalence, but it means that one has to deny that "State Gauge Transformation" and "Solution Gauge Transformation" as I defined them above are the right characterizations of gauge transformations on states and solutions respectively. In particular, one common part of the definition "State Gauge Transformation" and "Solution Gauge Transformation" is the commitment to gauge transformations being determined by considering curves that are generated by the *Total Hamiltonian.* I will argue that 1. gauge transformations on states do not require a commitment to a particular form of the Hamiltonian and 2. gauge transformations on solutions ought to be determined by considering curves that are generated by the Extended Hamiltonian rather than the Total Hamiltonian. This second argument bears close resemblance to a recent response to Pitts (2014*b*) by Pooley & Wallace (2022), so it will be helpful to spell out their argument first and pinpoint the way in which it falls short of providing a complete resolution to the debate before detailing the two arguments.

5 Pooley and Wallace's Response to Pitts

Pooley & Wallace (2022) show that in the example of classical Electromagnetism, if one starts with the Extended Hamiltonian, arbitrary combinations of first class constraints generate gauge transformations of solutions. Their argument can be summarised as follows. Consider the Extended Hamiltonian for classical Electromagnetism, where we add to the Total Hamiltonian the secondary constraint multiplied by an arbitrary function μ:

$$H_E = \int \frac{1}{2}(p_{\vec{A}}^2 + \vec{B}^2) + \vec{A}\cdot\vec{J} + \lambda p_V - (V+\mu)(\nabla\cdot p_{\vec{A}} - \rho) \tag{10}$$

With this Hamiltonian, the equations of motion become:

$$\begin{aligned} \frac{\partial \vec{A}}{\partial t} &= \frac{\partial H_E}{\partial p_{\vec{A}}} = p_{\vec{A}} + \nabla(V+\mu) \\ \frac{\partial V}{\partial t} &= \frac{\partial H_E}{\partial p_V} = \lambda \end{aligned} \tag{11}$$

When we now consider the transformation generated by an arbitrary combination of primary and secondary constraints, $\int \alpha\phi_0 + \beta\phi_1$, we find:

$$\begin{aligned} \frac{\partial \vec{A}'}{\partial t} &= \frac{\partial \vec{A}}{\partial t} + \frac{\partial \nabla\beta}{\partial t} = p_{\vec{A}} + \nabla(V+\mu+\alpha) \\ \frac{\partial V'}{\partial t} &= \frac{\partial V}{\partial t} + \frac{\partial \alpha}{\partial t} = \lambda \end{aligned} \tag{12}$$

We can rewrite the first equation as $\frac{\partial \vec{A}'}{\partial t} = \frac{\partial \vec{A}}{\partial t} = p_{\vec{A}} + \nabla(V+\mu+\alpha-\dot{\beta})$. Notice that μ, α and $\dot{\beta}$ are all arbitrary functions, so we can write this equation as

$$\frac{\partial \vec{A}'}{\partial t} = \frac{\partial \vec{A}}{\partial t} = p_{\vec{A}} + \nabla(V+\mu')$$

where μ' is arbitrary. This is just the untransformed equation of motion, with μ' in place of μ. In other words, if $(\vec{A}(t), V(t); p_{\vec{A}}(t), p_V(t))$ is a solution to $\frac{\partial \vec{A}}{\partial t} = p_{\vec{A}} + \nabla(V+\mu)$, then $(\vec{A}(t) + \nabla\beta, V(t)+\alpha; p_{\vec{A}}(t), p_V(t))$ is also a solution. Therefore, arbitrary combinations of first-class constraints generate gauge transformations on solutions, for the dynamics generated by the Extended Hamiltonian.

Although this argument shows that when we start with the Extended Hamiltonian, the gauge transformations are generated by arbitrary combinations of first-class constraints, it leaves open the question of what the justification is for starting with the Extended Hamiltonian. Indeed, it seems that the proponents of "Solution Gauge Transformation" will deny that this is the right starting point; they would say that it is the Total Hamiltonian that one should use to determine the gauge transformations.

Pooley & Wallace (2022) do provide one kind of response: the dynamics generated by the Extended Hamiltonian is *empirically equivalent* to the dynamics generated by the Total Hamiltonian: the predictions regarding invariant quantities are the same. In particular, what they notice is that the difference between the solutions of the Total and Extended Hamiltonian lies in what quantity plays the role of the electric field: when the Total Hamiltonian is used to generate the dynamics, it is $\dot{\vec{A}} - \nabla V$ that plays the role of the electric field, but when the Extended Hamiltonian is used, it is $p_{\vec{A}}$. And so, given that our access to these quantities is through the role they play in the equations of motion, there is no empirical difference between these choices of Hamiltonian.

Although I take this response to be both convincing and informative, I will argue that we can go further: the Extended Hamiltonian can be motivated purely on mathematical grounds, and therefore there are *theoretical* reasons for using the Extended Hamiltonian to determine the gauge transformations. When combined with Pooley and Wallace's argument, I think this provides a strong case in favour of the claim that arbitrary combinations of first-class constraints generate gauge transformations for solutions.

To make this argument, I will use a standard geometric way of expressing the constrained Hamiltonian formalism since it provides a neutral framework for illuminating the issues of concern. In particular, the geometric framework allows us to see clearly what the role of the first-class constraints is within the structure of the formalism. This will help to make clear the sense in which there are theoretical motivations for definitions of state and solution gauge transformations.

6 Geometric Formulation

The constrained Hamiltonian formalism can be expressed naturally in a geometric way using the theory of symplectic manifolds.[7] A symplectic manifold consists of a pair (M, ω) where M is a smooth manifold and ω is a *symplectic form*: it is a two-form (a smooth, anti-symmetric

[7] This formalism is widely used to express the constrained Hamiltonian formalism. For further details of this formalism, see Henneaux & Teitelboim (1994), Butterfield (2006).

tensor field of rank (0,2)), that satisfies the following conditions:

1. ω is *non-degenerate*, i.e. if $\omega(X_i, X_j) = 0$ for all $X_j \in TM$ and some $X_i \in TM$, then $X_i = 0$.

2. ω is *closed*, i.e., $\mathbf{d}\omega = \mathbf{0}$, where $\mathbf{d}$ is the exterior derivative operator, which is such that $\mathbf{d}f = df$, the differential of a function f, $\mathbf{d}(\mathbf{d}\alpha) = 0$ where α is a k-form, and $\mathbf{d}(f\alpha) = df \wedge \alpha + f\mathbf{d}\alpha$.

There is a sense in which every symplectic manifold comes equipped with "Poisson structure": Let (M, ω) be a symplectic manifold and $C^\infty(M)$ the space of smooth maps on M. In addition, let ω' be the inverse of ω (a smooth, anti-symmetric tensor field of rank $(2,0)$).[8] Then the map $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ defined by $f, g \mapsto \{f, g\} = \omega'(df)(dg)$ is a Poisson bracket on M.

A constrained Hamiltonian theory can be defined as a symplectic manifold in the following way. The manifold is *phase space*, consisting of the points $\{(q_i, p_i), i = 1, ..., N\}$, which can be understood as the cotangent bundle of configuration space, T^*Q, where Q consists of the points q_i. T^*Q comes equipped with a one-form, the *Poincaré one-form*, given by $\theta = p_i dq^i$. The corresponding two-form is given by $\omega = \mathbf{d}\theta = dp_i \wedge dq^i$, which is symplectic.

Given a function f, one can uniquely define a smooth tangent vector field X_f through:

$$\omega(X_f, \cdot) = \mathbf{d}f \tag{13}$$

where $\{\cdot\}$ represents any vector field tangent to T^*Q. In particular, one can uniquely define a vector field corresponding to the Hamiltonian $H = p^i q_i - L$ through $\omega(X_H, \cdot) = dH$. This provides an alternative way to write Hamilton's equations. In particular, $\{f, H\} = \omega(X_f, X_H) = df(X_H) = \mathcal{L}_{X_H}(f)$. If we define the flow parameter of X_H to be time, then this says that $\{f, H\} = \frac{df}{dt}$, which is Hamilton's equation.

We can understand the constraints $\varphi_i(q, p) = 0$ for $j = 1, .., M$ where M is the total number of constraints as giving rise to a smooth, embedded sub-manifold of phase space of dimension $N - M$, which we call the *constraint surface*, given by $\Sigma = \{(q, p) \in \Gamma | \forall_i : \varphi_i(q, p) = 0\}$. The

[8]This is well-defined because ω is non-degenerate.

first-class constraints are those constraints whose associated vector field is tangent to Σ, while the second-class constraints are those constraints whose associated vector field is not tangent to Σ.[9] For our purposes, we only consider the case where all the constraints are first-class, since these are the gauge-generating constraints.

We can define an induced two-form on the constraint surface $\tilde{\omega}$ as the pullback along the embedding $i : \Sigma \to \Gamma$ of ω. This induced two-form is in general *degenerate* i.e. it is not invertible. In particular, it possesses M linearly independent null vector fields that form the null space of $\tilde{\omega}$. These are the vector fields that satisfy $\tilde{\omega}(X_j, \cdot) = 0$ where $\{\cdot\}$ is any vector field tangent to Σ. But these are precisely the vector fields that off the constraint surface satisfy $\omega(X_j, \cdot) = d\gamma_j$ where γ_j are the first-class constraints, since $d\gamma_j|_\Sigma = 0$. Thus, we will write X_{γ_j} for these null vector fields to indicate that they are the tangent vector fields associated with the first-class constraints.

This means that one cannot associate a unique vector field with any smooth function on the constraint surface through the equation $\tilde{\omega}(X_f, \cdot) = \mathbf{d}f$, since if X_f satisfies this equation, so does $X_f + X_{\gamma_j}$. We call the geometry of such a surface *presymplectic*. The integral curves of the null vector fields are called the *gauge orbits*. Equivalently, the gauge orbits consist of the set of points that can be joined by a curve with null tangent vectors. They are M-dimensional surfaces on the constraint surface spanned by the null vectors.

The gauge orbits coincide with the notion of a gauge transformation in the Dirac formalism in the following sense: it is the null vector fields that generate the gauge orbits on the constraint surface, and these coincide with the vector fields X_{γ_j} corresponding to the first-class constraints. And so, arbitrary combinations of first-class constraints effectively generate a transformation that takes one 'along' a gauge orbit.

We can also understand the *observables* in the geometric formulation as the functions that are constant along the gauge orbit. In other words, the observables are the functions for which $\omega(X_f, X_{\gamma_j}) = 0$ on Σ, since $\omega(X_f, X_{\gamma_j}) = \mathcal{L}_{X_{\gamma_j}}(f)$ i.e. it is the flow of f along the gauge orbit.

[9]This coincides with the definition of first-class and second-class given in Section 2.

7 Geometric Resolution

To setup the argument in this section, recall that at issue is the question of how to reconcile the notion of gauge transformations of states and gauge transformations of solutions. Both Dirac (1964) and Pitts (2014*b*) take gauge transformations to be determined through the dynamics generated by the Total Hamiltonian, but this leads to different definitions in the case of states and of solutions, and consequently different opinions about whether one should extend the Hamiltonian or not. We can summarize the reasoning common to both sides of the debate as follows:

1. First, one determines the gauge transformations using the Total Hamiltonian.
2. Then, one uses the gauge transformations to say whether one should extend the Hamiltonian or not.

I will argue that this reasoning is flawed in three parts. First, I argue that Extended Hamiltonian is motivated independently from consideration of the gauge transformations, and so (2) is wrong: the gauge transformations do not determine the correct form of the Hamiltonian. Second, I argue that the gauge transformations on states arise naturally from the structure of the constraint surface, without considering the solutions to the equations of motion, and so (1) is wrong: the gauge transformations on states are not simply a special case of the gauge transformations on solutions. Finally, I use these two arguments to show that the gauge transformations on solutions (properly understood) are generated by arbitrary combinations of first-class constraints.

7.1 Motivating the Extended Hamiltonian

First, let's start with why the Extended Hamiltonian is motivated. It is clear that on the constraint surface, Hamiltonians related by an arbitrary combination of first-class constraints are equivocated. However, I think it is more important to recognize that on the constraint surface, the vector fields corresponding to solutions to the equations of motion for some Hamiltonian are defined only up to arbitrary combinations of vector fields associated with the first-class constraints. Take a (first-class) Hamiltonian vector field X_H and transform it to $X_H + a^j X_{\gamma_j}$

where X_{γ_j} are the null vector fields associated with the first-class constraints γ_j and a^j are arbitrary functions. We have that

$$\tilde{\omega}(X_H + a^j X_{\gamma_j}, \cdot) = \tilde{\omega}(X_H, \cdot) = dH|_\Sigma$$

since X_{γ_j} are null vectors. But this means that transforming X_H by an arbitrary linear combination of the vector fields associated with the first-class constraints preserves the dynamical equations on the constraint surface. In other words, the structure of the constraint surface is such that the evolution generated by X_H and that generated by $X_H + a^j X_{\gamma_j}$ is not distinguished: if f satisfies $\tilde{\omega}(X_f, X_H) = \frac{df}{dt}|_\Sigma$, then it satisfies $\tilde{\omega}(X_f, X_H + a^j X_{\gamma_j}) = \frac{df}{dt}|_\Sigma$. Therefore, we can think of the vector fields $X_H + a^j X_{\gamma_j}$ on the constraint surface as characterizing the equivalence class of vector fields that generate solutions to the equations of motion. Let us call this equivalence class of vector fields the "Extended Hamiltonian vector fields".

Notice that in such reasoning, we have not made any assumptions about the X_{γ_j} being associated with primary or secondary first-class constraints, nor about what the gauge-transformations are; each first-class constraint constitutes a null direction of the constraint surface, and it is this property that is important in determining which transformations of the Hamiltonian vector field are dynamically equivalent. In particular, notice that the sense of dynamical equivalence here is just that these Hamiltonian vector fields form an equivalence class, relative to the structure of the constraints surface. Inasmuch as this structure is how one makes predictions in the theory, these Hamiltonian vector fields generate the same predictions.

This provides one way in which restricting to the Total Hamiltonian is unnatural in the geometric framework: it distinguishes a class of null vectors (those that correspond to primary first-class constraints) that cannot be distinguished from other null vectors in terms of the structure of the constraint surface.[10]

[10] One can distinguish the secondary constraints through the fact that they correspond to time derivatives of the primary constraints, but this is not the relevant kind of difference in determining the equivalence class of Hamiltonians.

7.2 State Gauge Transformations

Second, let's consider the notion of gauge transformations on states. Again, restrict ourselves to the constraint surface. On the constraint surface, the induced two-form $\tilde{\omega}$ only acts on vector fields that are tangent to the constraint surface. In other words, the induced two-form only applies to functions g such that $X_g d\gamma_j = 0$ on the constraint surface. But notice that these are precisely the functions for which $\omega(X_g, X_{\gamma_j}) = 0$ when restricted to Σ i.e. the observables. So the functions f for which $\omega(X_f, X_{\gamma_j}) \neq 0$ on Σ must be such that X_f is not tangent to Σ, and thus one cannot define the action of $\tilde{\omega}$ on them. This means that for functions which vary along the gauge orbits (the gauge variables), the induced two-form effectively cannot 'see' this change, since it only acts on those vector fields tangent to the constraint surface. And so, points along a gauge orbit are equivalent in the sense that one cannot distinguish the value of a function at different points along a gauge orbit using the structure of the constraint surface.

Notice that this reasoning does not make reference to the dynamics, in particular, it doesn't make reference to the Total Hamiltonian; it relies only on the structure of the constraint surface. This suggests a revision to the definition of the state gauge transformations:

> **State* Gauge Transformation:** A (state) gauge transformation is a transformation that relates any two states on the constraint surface that cannot be distinguished by the induced two-form.

This emphasizes that what makes states along a gauge orbit equivalent has to do with their role in the structure of the constraint surface. Notice that on this definition, arbitrary combinations of first-class constraints generate gauge transformations precisely because they give rise to the gauge orbits. We therefore have a definition of the gauge transformations on states that is motivated independently from the gauge transformations on solutions, but which agrees with both sides of the debate about the generators of gauge transformations on states.

We can also use this argument to oppose a claim made by Henneaux & Teitelboim (1994). They say:

> "The identification of the gauge orbits with the null surfaces of the induced two-form relies strongly on the postulate made throughout the book that all first-class constraints generate gauge transformations." (p. 54)

In other words, they suggest that one must independently maintain that first-class constraints generate gauge transformations in order to interpret the null surfaces as the gauge-equivalent points. But the argument above shows that this interpretation is motivated from within the geometric formulation.

7.3 Solution Gauge Transformations

Finally, let us consider the gauge transformations on solutions. From the previous two arguments, we have determined that adding an arbitrary combination of null directions to the Hamiltonian vector field generates a curve whose derivative along the Hamiltonian vector field is the same (on the constraint surface). This curve differs only with regards to where on the gauge orbit it lies at each point in time. Moreover, each state along a gauge orbit forms an equivalence class of states. Thus, transforming a solution by an arbitrary amount along the gauge orbit at each point gives rise to another solution generated by a Hamiltonian vector field with a different combination of null vectors. We have a natural reason to think these are physically equivalent, since the Hamiltonians of this kind form an equivalence class, and the states along a gauge orbit form an equivalence class.

To see this more precisely: take a curve $s(t)$ defined on the constraint surface whose tangent vector is a solution to the equations of motion $\tilde{\omega}(X_s, X_H + a^j X_{\gamma_j}) = \frac{ds}{dt}$. Now take another curve $s'(t) = h(t) \cdot s(t)$ where $h(t)$ is a smooth function that "moves" $s(t)$ by some amount along the gauge orbit at each point. Then $X_{s'}$ will also be a solution to $\tilde{\omega}(X_{s'}, X_H + a^j X_{\gamma_j}) = \frac{ds'}{dt}$, since this equation of motion determines the tangent vector to the dynamical trajectory only up to the addition of an arbitrary (time-dependent) combination of null vectors. Therefore, an arbitrary combination of first-class constraints generates a transformation that takes solutions of the Extended Hamiltonian to other solutions.

This motivates the following characterization of the solution gauge transformations:

> **Solution* Gauge Transformation:** A (solution) gauge transformation relates any two curves that are possible evolutions from an initial state under the dynamics generated by the Extended Hamiltonian vector fields.

Notice that this definition is supported on two fronts. First, we have independently motivated the Extended Hamiltonian vector fields as the correct equivalence class (which we argued earlier was lacking in Pooley & Wallace (2022)). Second, we have independently motivated that the equivalence class of states is given by the gauge orbits. This provides another way in which restricting to the Total Hamiltonian is unnatural geometrically: it would be to say that the dynamics can distinguish states along a gauge orbit, even though the structure of the constraint surface is such that it cannot distinguish these states. So we shouldn't think that gauge transformations on states are a special case of those on solutions; rather, there are two independent notions that are coherent with each other.

In summary, we can diagnose the debate about the right characterization of a gauge transformation as follows: the debate takes for granted that gauge transformations are determined by the evolution generated by the Total Hamiltonian. This leads to a disagreement about the generators of gauge transformations, and consequently the right equivalence class of Hamiltonians. What I have argued here is that this reasoning is flawed: the Extended Hamiltonian can be motivated as the right equivalence class of Hamiltonians prior to determining the gauge transformations, and the gauge transformations on states can be determined without directly considering the evolution generated by the equivalence class of Hamiltonians. This allows one to maintain a clear conceptual difference between gauge transformations on states and gauge transformations on solutions, and it allows one to maintain that both of these notions capture a notion of physical equivalence without conceptual tension.

8 Possible Counterarguments

Before concluding, let's consider how one might respond to the argument given in the previous section; in particular, how one might defend "Solution Gauge Transformation" over "Solution* Gauge Transformation", since it is these notions that lead to different characterizations of the transformations that generate gauge transformations. One notable aspect of the argument is its commitment to the geometry of the constraint surface as a guide to the symmetries of the theory. So let us consider two objections one might have to this. First, that we shouldn't restrict to the constraint surface. Second, that we shouldn't think that the geometrical formulation of the constrained Hamiltonian formalism is adequate.

Starting with the first objection: Inasmuch as constraints are understood to provide the "physically allowed states", it seems natural to think that the points off of the constraint surface are unnecessary for describing the dynamics of the theory. However, one might want to maintain that these points still have importance as "kinematically possible" states. That is, one might want to maintain that we ought to consider states off the constraint surface as important for describing the physical theory as a whole, even if the dynamics is restricted to the constraint surface. In particular, the secondary constraints are fixed by thinking about the consistency of the primary constraints *with the dynamics*. And so it might seem that at least when it comes to secondary constraints, there is no logical inconsistency with specifying a theory in terms of points where the secondary constraints do not hold. And the vector fields associated with the secondary constraints are not null vectors of the two form on the full phase space (nor on the primary constraint surface); the full phase space is symplectic, and so it is non-degenerate by definition. So, the counterargument goes, we cannot use the fact that these vector fields are null to argue that they generate gauge transformations.

One natural response is that the points off of the constraint surface are 'excess structure': although there is nothing inconsistent about including them, the content of the theory is given by the constraint surface. Another response is to point out that the idea that we start out with the primary constraints and then generate the secondary constraints through the dynamics is somewhat an accident of the way that the Hamiltonian formalism is usually set up. As I presented Dirac's version of the theory, one starts with a Lagrangian function, from which one derives the primary constraints. Only once we have the primary constraints and the Hamiltonian in hand do we determine the secondary constraints. But we could have set up the Hamiltonian formalism in a different way: we could say that our theory is given by specifying a Hamiltonian function, a symplectic two-form, and a collection of constraints. In this way of setting up the formalism, although there is a functional relationship between the primary and secondary constraints, there is no clear difference in the role that they play. In particular, the only relevant difference seems to be which constraints are first-class; these are the ones that generate transformations that keep one along the constraint surface, and which are null vectors of the induced two-form on the constraint surface.

In order to push back on this response, one would have to argue that there is something wrong with setting up the Hamiltonian formalism on the constraint surface. This leads to the second objection, namely that the geometric formulation of the constrained Hamiltonian formalism is not adequate. The first thing to note here is that this geometric formulation is a natural extension of a widely accepted formulation of Hamiltonian mechanics without constraints using symplectic manifolds, and so in this sense is well motivated. But one might want to argue that it is inadequate in a different way. In particular, one might want to argue that the Hamiltonian formalism is derivative from the Lagrangian formalism; the Lagrangian formalism is the "fundamental" one, and the Hamiltonian formalism is just an alternative way of expressing the Lagrangian one. Indeed, on the standard way of presenting Dirac's formalism, one begins with a Lagrangian, and uses it to define the Hamiltonian and constraints. On this view, there is a difference between the primary and secondary constraints that comes from the Lagrangian viewpoint and that isn't captured purely through consideration of the geometry of the constrained Hamiltonian formalism. The difference is that the primary constraints are necessary to ensure that the Hamiltonian formalism is equivalent to the Lagrangian formalism, while the secondary constraints are 'extra' constraints on the Hamiltonian side that are not motivated from the Lagrangian perspective. In particular, it is only the primary constraints that have to be imposed in order for the map from the Lagrangian to Hamiltonian state spaces (the Legendre transformation) to be invertible.

Therefore, this argument goes, restricting to the secondary constraint surface – and consequently having the view that arbitrary combinations of first-class constraints generate gauge transformations – leads to a theory that is inequivalent to the Lagrangian theory, and so is not the right theory to consider. Indeed, one can show that the Total Hamiltonian formalism, understood as relying on the primary constraint surface, gives rise to solutions that are equivalent to the solutions to the Euler-Lagrange equations (Batlle et al. (1986)). Therefore, it seems that restricting to the constraint surface (including the secondary constraints) gives rise to a theory that although is *empirically* equivalent to the Lagrangian formalism, is not strictly the same theory. And so, if one takes the view that the Lagrangian formalism is more fundamental, then this might motivate one to say that our definition of a gauge transformation should be inherited from this formalism, and thus not the definition motivated by the geometry of the constraint surface.[11]

There are several deep and subtle questions here about what makes one theory more "fundamental" than another and how to characterize the equivalence of theories that I do not have space to answer here.[12] However, what the arguments here suggest is that if one wants to advocate for "Solution Gauge Transformation" over "Solution* Gauge Transformation", these are questions that one is forced to face. That is, in order for one to maintain that "Solution Gauge Transformation" is the correct definition, one has to say what it is that provides the distinction between primary and secondary constraints in characterizing the equivalence class of Hamiltonians, and it seems that this requires one to adopt views regarding the relationship between the Hamiltonian and Lagrangian formalisms.

9 Conclusion

To summarize, I have argued that the debate about the correct characterization of the gauge transformations in the constrained Hamiltonian formalism rests on assumptions about the relationship between gauge transformations and the form of the Hamiltonian that are unnatural from the perspective of the geometric formulation of the constrained Hamiltonian formalism. Using the geometric formulation, I showed that we can distinguish between gauge transformations on states and gauge transformations on solutions in a conceptually clear way and that both are generated by arbitrary combinations of first-class constraints, thereby supporting the orthodox view. However, this allowed us to pinpoint more clearly where disagreement can be found. In particular, I suggested that there are crucial questions about the relationship between Lagrangian and Hamiltonian theories in the presence of gauge symmetry, where different answers to these questions can lead to different views regarding the correct form of the Hamiltonian, and thus to what the correct characterization of the gauge transformations is.

One important topic that I have not discussed in this paper is the "Problem of Time". Recall: for theories that are time-reparameterization invariant, the standard account of gauge

[11] On the other hand, Gryb & Thébault (2023, ch.8) argue that on a proper understand of Noether's Second Theorem, the secondary Hamiltonian constraints can be derived from the Lagrangian perspective. This suggests that the Extended Hamiltonian is not in conflict with the Lagrangian account.

[12] In Bradley (2024), I address the question of the equivalence between Lagrangian and Hamiltonian gauge theories.

transformations implies that time evolution is itself a gauge transformation since the Hamiltonian is a first-class constraint. In supporting the standard account of gauge transformations as being generated by arbitrary combinations of first-class constraints, it might appear that we are also left with the issues surrounding the Problem of Time. That is, we haven't seemed to do anything to deny that a Hamiltonian first-class constraint generates a gauge transformation. However, I think that the distinctions drawn out here highlight what is interesting about the case of a Hamiltonian first-class constraint. In particular, the claim that we can conceptually distinguish the gauge transformations on states and the gauge transformations on solutions does not seem to be true in the case where the Hamiltonian is a first-class constraint: the gauge orbits are just the solutions to the equations of motion, and so the states along a gauge orbit cannot be understood independently from the dynamics. Thus, it is less clear whether one can distinguish two notions of physical equivalence as well. This suggests that the puzzle surrounding the Problem of Time at least partly comes down to the fact the transformation generated by a Hamiltonian first-class constraint doesn't fall neatly into the categories defined here. But more work needs to be done to say what exactly is distinct about this case, inasmuch as when the Hamiltonian is a first-class constraint, it seems to play the same role geometrically as any other first-class constraint since it is a null vector field of the induced two-form on the constraint surface. To answer this would require a more careful consideration of the role of the Hamiltonian and whether there is a more fine-grained distinction between different kinds of constraints.[13] I hope that the work here has at least provided support for the claim that the Problem of Time is not the result of an incorrect definition of the gauge transformations in the constrained Hamiltonian formalism; rather, it must be treated on its own terms.

References

Baker, D. J. (2010), 'Symmetry and the metaphysics of physics', *Philosophy Compass* **5**(12), 1157–1166.

Barbour, J. & Foster, B. Z. (2008), 'Constraints and gauge transformations: Dirac's theorem is not always valid', *arXiv preprint arXiv:0808.1223* .

[13]For an analysis of this kind, see Gryb & Thébault (2023, ch.13).

Batlle, C., Gomis, J., Pons, J. M. & Roman-Roy, N. (1986), 'Equivalence between the Lagrangian and Hamiltonian formalism for constrained systems', *Journal of Mathematical Physics* **27**(12), 2953–2962.

Bradley, C. (2024), 'The relationship between lagrangian and hamiltonian mechanics: The irregular case'. Unpublished Manuscript.

Butterfield, J. (2006), On symplectic reduction in classical mechanics, *in* J. Butterfield & J. Earman, eds, 'The Handbook of Philosophy of Physics', North Holland, pp. 1–131.

Dirac, P. A. M. (1964), *Lectures on Quantum Mechanics*, Dover.

Earman, J. (2004), 'Laws, symmetry, and symmetry breaking: Invariance, conservation principles, and objectivity', *Philosophy of Science* **71**(5), 1227–1241.

Gryb, S. & Thébault, K. (2023), *Time Regained: Volume 1: Symmetry and Evolution in Classical Mechanics*, Oxford University Press.

Henneaux, M. & Teitelboim, C. (1994), *Quantization of Gauge Systems*, Princeton University Press.

Ismael, J. & Van Fraassen, B. C. (2003), Symmetry as a guide to superfluous theoretical structure, *in* 'Symmetries in physics: Philosophical reflections', Cambridge University Press Cambridge, pp. 371–392.

Pitts, J. B. (2014*a*), 'Change in hamiltonian general relativity from the lack of a time-like killing vector field', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **47**, 68–89.

Pitts, J. B. (2014*b*), 'A first class constraint generates not a gauge transformation, but a bad physical change: The case of electromagnetism', *Annals of Physics* **351**, 382–406.

Pons, J. M. (2005), 'On dirac's incomplete analysis of gauge transformations', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **36**(3), 491–518.

Pons, J. M., Salisbury, D. C. & Shepley, L. C. (1997), 'Gauge transformations in the lagrangian and hamiltonian formalisms of generally covariant theories', *Phys. Rev. D* **55**, 658–668.

Pooley, O. & Wallace, D. (2022), 'First-class constraints generate gauge transformations in electromagnetism (reply to pitts)', *arXiv preprint arXiv:2210.09063* .

Thébault, K. P. Y. (2021), The problem of time, *in* E. Knox & A. Wilson, eds, 'The Routledge Companion to Philosophy of Physics', Routledge.