

A puzzle concerning local symmetries and their empirical significance

Abstract

In the last five years, the controversy about whether or not gauge transformations can be empirically significant has intensified. On the one hand, Greaves and Wallace (2014) developed a framework according to which, under some circumstances, gauge transformations can be empirically significant, and Teh (2015) further supported this result by using the Constrained Hamiltonian formalism. On the other hand, Friederich (2015, 2016) claims to have proved that gauge transformation can never be empirically significant. In this paper, I accomplish two tasks: first, I show that Friederich's proof is not valid, and that once it is corrected, it entails a result that is compatible with the treatments by Greaves, Wallace, and Teh. Second, I show that, despite criticism by Brading and Brown (2004) and Friederich (2015), t'Hooft's Beam-Splitter experiment is indeed a concrete realization of a case where a local gauge symmetry has empirical significance. By shedding light on these two points, this paper shows that recent arguments that claim gauge transformations cannot be empirically significant are not satisfactory.

Keywords: Symmetry, gauge transformations, empirical significance

Word Count: 7900

1 Introduction

According to conventional wisdom, local transformations, of either the whole universe or of subsystems within it, never relate universe states that are physically distinguishable from one another. Recently, however, in a series of papers, many of which have been published in this journal, this point has given rise to new debates. On the one hand, Greaves and Wallace (2014) developed a general framework from which it is transparent that there are cases where local symmetries produce empirically distinct physical states, and by using the *Constrained Hamiltonian* formalism, Teh (2015) further clarifies the conditions under which this result holds. On the other hand, by generalizing the framework of Greaves and Wallace (GW henceforth) and by appealing to some very general and plausible principles regarding the individuation and composition of physical systems, Friederich (2015, 2016) has attempted to prove the result that gauge transformations are *never* empirically significant (that is, they never give rise to situations that are empirically distinct, but see section 2 for a detailed explanation of ‘empirical significance’).¹ This leads to a puzzle: how can it be that if one adopts GW’s framework and the Constrained Hamiltonian formalism, then some local gauge transformations can be shown to be empirically significant and, on the other hand, if one further develops GW’s framework and adopts some general principles about the individuation of subsystems, then suddenly gauge transformations can be proven to lack empirical significance? One of the two main goals of this paper is to offer a solution to this question.

Friederich, who is very much aware of the puzzle just introduced, proposes some sort

¹It is worth mentioning that Brading and Brown (2004) and Healey (2009) also offer arguments with the purpose of defending the conventional wisdom, but for the sake of space, I will restrict this particular discussion to the debates between Greaves, Wallace and Teh, on the one hand, and Friederich, on the other (for a recent argument against the conventional wisdom from the perspective of the Standard Model, see Dougherty (2019)).

of pluralism according to which different answers regarding the empirical status of gauge symmetries are all acceptable because they come from different (possibly incompatible) frameworks that disagree, among other things, about how to individuate physical subsystems (Friederich, 2016, 9). In this paper, I will offer a different (non-pluralistic) solution to this dilemma. In particular, I will show that the tension between the results of GW and Teh on the one hand, and the results of Friederich on the other, is not a matter of incompatible assumptions or frameworks, but a consequence of the fact that Friederich's result is not valid. As I will explain in detail in section 3, the problem with his argument is that it attempts to derive a result about any kind of gauge transformation by just focusing on a very particular kind (namely, transformations that go to the identity at some point of the environment). Once we identify this problem, it becomes clear that the alleged proof only establishes the weaker result that a certain kind of gauge transformation (i.e., one that 'asymptotes' to the identity in a smooth way) is not empirically significant. Importantly, both GW and Teh's treatments also agree that this very same kind of gauge transformation is not empirically significant (but their treatments also show that other kind of gauge transformations not affected by Friederich's proof can indeed be empirically significant). Contrary to appearances, then, the different frameworks converge to the very same result, and the puzzle is solved.

The second goal of this paper is to show that, despite criticism by Brading and Brown (2004) and Friederich (2015), t'Hooft's Beam-Splitter experiment is indeed a concrete realization of a case where a local gauge symmetry has empirical significance. As I will explain in section 4, the main problem with these criticisms is that they mischaracterize the feature that provides empirical significance to the symmetry at hand. Thus, by clarifying how t'Hooft's Beam-Splitter is indeed an experiment where a gauge transformation is empirically significant, and by showing that Friederich's proof does not work, I will reinforce GW and

Teh’s claim that local gauge transformations can be empirically significant.

The structure of the paper goes as follows. In section 2, I explain why Greaves and Wallace and Teh think that, simply as a matter of very general considerations, local symmetries can be empirically significant. In section 3, I discuss Friederich’s alleged proof that local symmetries can never be empirically significant, and I show why it is not logically valid. Furthermore, I argue that the correct conclusion of the proof actually agrees with the results of Greaves and Wallace (2014) and Teh (2015). In section 4, I show that the main arguments against the view that t’Hooft’s Beam-Splitter experiment constitutes an empirical realization of a local symmetry with empirical significance do not succeed.

2 Greaves and Wallace’s formalism

There are three main distinctions we should be aware of for the upcoming discussions. One is the distinction between external and internal symmetries. An external symmetry is a transformation of the space-time parameters of a theory (i.e., x, y, z, t) and an internal symmetry is a transformation of any other parameter that is not a space-time one. For example, boosts and rotations are external symmetries, whereas transformations of the spin of a particle are internal. The second distinction is that between global symmetries and local symmetries. Roughly, a local symmetry is a transformation from space-time to a gauge group (such as the transformations of the four-vector potential in electromagnetism), whereas a global transformation does not depend on space-time (the term “gauge transformation” is sometimes taken to mean *a local transformation that relates different descriptions of the same physical state*, but clearly this is not what it is meant in this paper for otherwise there would not be any substantive question about the empirical significance of these

transformations).²

The third important distinction is that between symmetries of a subsystem and symmetries of the whole universe. As the name indicates, a symmetry of a subsystem (a system that is not the universe itself) relates states of the subsystem, whereas a symmetry of the whole universe relates states of the universe. I will follow Greaves and Wallace (2014) in taking symmetry-related states of the universe to always represent empirically indistinguishable states of affairs. Furthermore, if two states of a subsystems (or the universe) that are related by a symmetry are empirically indistinguishable, then I will say that the symmetry lacks empirical significance (or, equivalent, I will say that it is not empirically significant). Obviously, then, from GW’s set-up it follows that symmetries of the universe lack empirical significance. On the other hand, if two states that are related by a symmetry are empirically distinguishable, then I will say that the symmetry is empirically significant.³ For example, a world where a certain ship is at rest with respect to the shore and a world where the same ship is moving with respect to that same shore are empirically distinct worlds, and so the symmetry relating the states of the ship (i.e., a given boost) is an empirically significant symmetry. But a boost is an external symmetry, and in this paper we will exclusively focus on internal local transformations of subsystems. According to conventional wisdom, these always lack empirical significance.

In order to tackle the question of what precise conditions have to be satisfied for a symmetry to be empirically significant, GW develop a general framework that allows us to answer that question for different kinds of theories and systems. Here I will briefly introduce

²Teh (2015, 97) calls ‘Redundant’ local transformations defined in this way (that is, defined as transformations that relate redundant descriptions), and ‘Formal’ those local transformation from space-time to a gauge group.

³For this reason, symmetries of the universe are sometimes called “theoretical symmetries” and some symmetries of subsystems are called “empirical symmetries” (Healey, 2009). However, I will not be using that terminology here. Others use the term ‘Direct Empirical Significance’ to refer to what I am here calling ‘empirical significance’ (for example, see Brading and Brown (2004), Friederich (2015) or Teh (2015)).

the essential concepts of that framework.

At the heart of GW proposal is the remark that we should be able to split states of the whole universe into states of subsystems of the universe. In particular, one of GW’s main assumptions is that we should be able to describe the state of the universe, u , in terms of a Cartesian product of the state of a subsystem s and the state of the environment, e : $u = \langle s, e \rangle$ (here u is an element of the set \mathcal{U} that contains all physically possible universe states, s is an element of the set \mathcal{S} that contains all physically possible states of the subsystem, and e is an element of \mathcal{E} , the set containing all physically possible states of the environment). Importantly, GW introduce an operation denoted by ‘ $*$ ’ that should be interpreted as follows: “ $s_1 * s_2$ ” means “the composition of state s_1 and state s_2 .” So, in particular, $u = \langle s, e \rangle$ can be written as $u = s * e$. Importantly, not any combination between states of the environment and states of the subsystem will yield to a well-defined state of the universe, and so $s * e$ will not always yield a state u in \mathcal{U} (so \mathcal{U} is a subset of the Cartesian product $\mathcal{S} \times \mathcal{E}$, see (Greaves and Wallace, 2014, 68)).

Next, let’s consider how GW represent symmetries in this formalism. A given symmetry σ of the universe acting on a state of the universe u gives us a new state u' in \mathcal{U} : $u' = \sigma(u)$ (unless the symmetry in question is the identity, in which case $u' = \sigma(u) = u$). And just as universe states can be decomposed into states of the subsystem and the environment, symmetries of the universe can be uniquely decomposed into symmetries of the subsystem and the environment. In particular, for all $s \in \mathcal{S}$, $e \in \mathcal{E}$, $\sigma(u) = \sigma(s * e) = \sigma_S(s) * \sigma_E(e)$ for some maps σ_S and σ_E (here σ_S is the restriction of σ to the subsystem and σ_E is the restriction of σ to the environment).

We have now the main ingredients needed to answer the question of when, according to GW, subsystem symmetries are empirically significant (recall that a universe symmetry always relates empirically equivalent states of the universe, and so only subsystem symmetries

can be empirically significant). According to GW, the distinction between symmetries that are empirically significant and those that are not does not hinge, as conventional wisdom has presupposed, on a distinction between global symmetries and local symmetries. Rather, it hinges on a distinction between what GW call “interior” and “non-interior” symmetries (we should not confuse interior symmetries with internal symmetries).

Interior symmetries: A subsystem symmetry σ_s is interior iff for all $s \in S$ and all $e \in E$ for which $s * e$ is defined, $\sigma_s * e$ is well-defined and it is empirically equivalent to $s * e$.

In other words, if σ_s is interior, then if we start with an arbitrary universe state $s * e$ and apply σ_s to s , we recover an empirically equivalent state of the universe. Thus, obviously, only non-interior symmetries could be empirically significant.

According to GW, there are two main ways in which a subsystem symmetry can be empirically significant: first, a symmetry can be empirically significant if, for some states s and e , (a) $s * e$ and $\sigma_s(s) * e$ are both well-defined and (b) they represent empirically distinct states of the universe. Indeed, this is precisely the case of a ‘Galileo-ship scenario’; if we represent the ship at rest with respect to the shore by $s * e$, then we will represent the ship moving with respect to the shore by $\sigma_s(s) * e$ (where σ_s is a boost, and the state of the environment stays the same), and clearly, $s * e$ and $\sigma_s(s) * e$ are empirically distinct states of the world. I will call this kind of case, where the symmetry is empirically significant in the way a Galileo-ship case is, “Type I”.

Importantly, notice that measurements confined to the subsystem alone (to the ship cabin in the last example) or measurements confined to the environment alone (to the shore) would not be able to distinguish between $s * e$ (the ship at rest) and $\sigma_s(s) * e$ (the ship moving). Thus, the empirical content of such a symmetry is associated with measurements

of relational properties between the environment and the subsystem. For example, the boost of the ship obviously induces changes in the relative speed between the ship and the shore.

Second, for cases where the subsystem is not appropriately isolated from the environment, it can happen that the symmetry in question disrupts the relevant boundary conditions so that it maps a well-defined state $s * e$ of the universe to an ill-defined state $\sigma_s(s) * e$. In that case, the action of the symmetry requires us to alter the environment state, $e \mapsto e'$, in such a way that $\sigma_s(s) * e'$ is then well-defined (here e' and e are physically distinct states of the environment). Notice that this is a case analogous to a Faraday cage scenario in electrostatics, where the subsystem is the interior region of the cage and the environment is the surface of the conductor; shifts in the scalar potential in the interior (an internal and global symmetry) are associated to changes in the surface charge (and of course, these are detectable changes of the state of the environment). However, the main debate around the empirical status of gauge transformations has centered around Type I cases, and in this paper I will follow the literature in that respect.

The crucial take-away message of GW's framework, for the purposes of the present paper, is this: GW's description of Type I cases is general enough in that it does not specify anything about the kind of subsystem symmetry or the kind of physical theory we are dealing with. That is, on the face of it, GW's framework is so general that it not only accomodates Type I cases involving global symmetries (such as a classic Galileo ship scenario) but, more interestingly, it seems to allow for Type I cases of local symmetries. In particular, a Type I case for a local symmetry σ_s would be one for which (a) for some states of the subsystem, $s * e$ and $\sigma_s(s) * e$ are both well-defined, and (b) $s * e$ and $\sigma_s(s) * e$ represent empirically distinct states of the universe. Not only their framework allows for the possibility of local transformations that are empirically significant (a possibility that

Brading and Brown (2004, 657), Healey (2009) and Friederich (2015) want to deny)⁴, but GW also say that t’Hooft’s Beam Splitter experiment is precisely a concrete realization of a local Type I case (we will study this case in section 4).

So far the discussion has been very general, and so at this point it would be helpful to say a bit more about the conditions under which local gauge transformations can lead to Galileo-type scenarios. By using the Constrained Hamiltonian formalism, Teh (2015) sheds light into those conditions. In particular, he explains that we should formally distinguish between the group of the so-called ‘small’ gauge transformations (transformations that, asymptotically, can be smoothly transformed to the identity), and the group of ‘large’ gauge transformations (transformations that cannot be smoothly deformed to the identity)(Teh, 2015, Sec. 4). Indeed, we should distinguish between three subgroups of the group of gauge transformations \mathcal{G} . One is \mathcal{G}_0^∞ , the group of gauge transformations that is smoothly connected to the identity. Another one is \mathcal{G}^∞ , which corresponds to the group of transformations that go asymptotically (not necessarily in a smooth way) to the identity. Finally, there is \mathcal{G}_I , the group of transformations that leave invariant the boundary conditions of the fields (so that the subsystem remains invariant). Then, we get the following hierarchy (2015, 115):

$$\mathcal{G}_0^\infty \subset \mathcal{G}^\infty \subset \mathcal{G}_I \subset \mathcal{G}. \tag{1}$$

\mathcal{G}_0^∞ contains all the small gauge transformations and these cannot exhibit empirical significance (they are ‘interior’ in GW’s sense). Therefore, the only candidates of gauge

⁴For example, Brading and Brown (2004, 657) say that “in conclusion, there can be no analogue of the Galilean ship experiment for local gauge transformations.” And Friederich (2015), as we will see in section 3, intends to prove this.

transformations that can be empirically significant are those in \mathcal{G}_I (for they need to preserve the boundary conditions in order to count as a Type one case) that are not in \mathcal{G}_0^∞ (i.e. those in $\mathcal{G}_I/\mathcal{G}_0^\infty$). These are then the ‘large’ transformations, and the way they are empirically significant is by inducing relational changes with respect to an environment (where the environment here is taken to correspond to a fixed frame at the asymptotic boundary).⁵ Finally, let us point out that Teh also agrees with GW in that t’Hooft’s Beam Splitter experiment offers an example of a local transformation that has empirical significance (Teh, 2015, 109).

3 Friederich’s proof

At this point, someone defending the orthodoxy—in this particular context, the view that Type I cases of local symmetries are impossible—can proceed in at least two ways. She could either try to resist GW’s general framework altogether, or she can endorse the framework (or part of it) and show that, despite the fact that it seems to allow for their possibility, there are reasons to believe that Type I cases for local symmetries are impossible (of course, that person would also have to explain why Teh’s analysis of the constrained Hamiltonian formalism is problematic, but let me focus here on GW’s general framework). Although in this paper I will be concerned with the second route, let me briefly offer three reasons for why the first route—the total rejection of GW’s framework—does not seem to be a very promising way of defending the orthodoxy.

First, GW’s framework gets right all of the uncontroversial cases of empirical significance. It entails that external transformations (such as boosts and rotations) of isolated subsystems

⁵Teh points out that for a long time, physicists have taken seriously the possibility of local symmetries exhibiting Type I empirical significance. And indeed, ‘large’ transformations are crucial to the construction of charges.

are empirical significant; that shifts in the scalar potential in the interior of a conductor are empirical significant in the case of electrostatics; that universe states related by universe symmetries are physically indistinguishable; and, finally, that gauge transformations that asymptote to the identity (and hence are trivial there) are to be understood as mere redundancies of our descriptions of the subsystem at hand. Thus, at least with respect to these uncontroversial cases, the framework seems to be on the right track.

Second, the framework has the virtue that it is so general that it can be easily applied to many different physical theories. Indeed, it is hard to imagine that the framework could not be applied to a given physical theory, for as long as the theory in question deals with the state space of a system, transformation rules between these states, and composition rules for states of different subsystems (all of which are very general features of physical theories), then GW's framework could be applied to that theory. Thus, if our project is that of answering when symmetries are empirically significant, it seems that rejecting a framework that is capable of modelling very different physical theories is not the best way to go (unless a replacement is put forward). And third, as far as I know, nobody has offered reasons to reject the framework itself (of course, rejecting the framework merely as a result of it entailing that there could be local symmetries with empirical significance would be question-begging in the context of the present dispute).

Let's consider now the only argument in the literature that, while explicitly accepting GW's framework, attempts to prove that Type I local symmetries are impossible. In particular, the argument intends to show that "on a natural development of the Greaves-Wallace framework, a version of the standard view can be vindicated, which says that only global symmetries can have direct empirical significance" (Friederich, 2015, 540). Or as he puts in his 2016 paper (my emphasis):

as I will show, one obtains a result according to which all (subsystem-restricted) gauge transformations in local gauge theories are without any direct empirical significance, *whether or not they reduce to the identity transformation on the subsystem boundary and whether or not they connect topologically inequivalent configurations* (Friederich, 2016, 5).

At this point, the puzzle mentioned in the introduction appears. How can it be that on the one hand, according to both GW’s framework and the Constrained Hamiltonian formalism, local gauge transformations that are not trivial at the boundary with the environment can be empirically significant, and yet, on the other hand, gauge transformations can be proven to always lack empirical significance? Friederich, who is aware of this puzzle, suggests some sort of pluralism according to which different answers regarding the empirical status of gauge symmetries are all acceptable because they come from different (possibly incompatible) frameworks (Friederich, 2016, 9). As I will explain now, however, there is no need to adopt this kind of pluralism around the empirical significance of gauge symmetries, for there is an important problem with the alleged proof by Friederich.

The following fact about gauge symmetries plays a crucial role in the upcoming argument (I find the other assumptions or facts used by Friederich to be very natural extensions of GW’s formalism, and so, for reasons of space, I will not discuss them):

FACT*: Any local symmetry σ_s defined on a subsystem S can be extended to an interior symmetry defined on a larger subsystem V of which S is a part (Friederich, 2015, 548).⁶

In other words, FACT* says that we can always extend a subsystem gauge transforma-

⁶Although in his (2015) paper this fact was taken to be an assumption (there called “Ext” for ‘extendability’), it is stated as a derivable fact of gauge theories in his (2016, 9) paper.

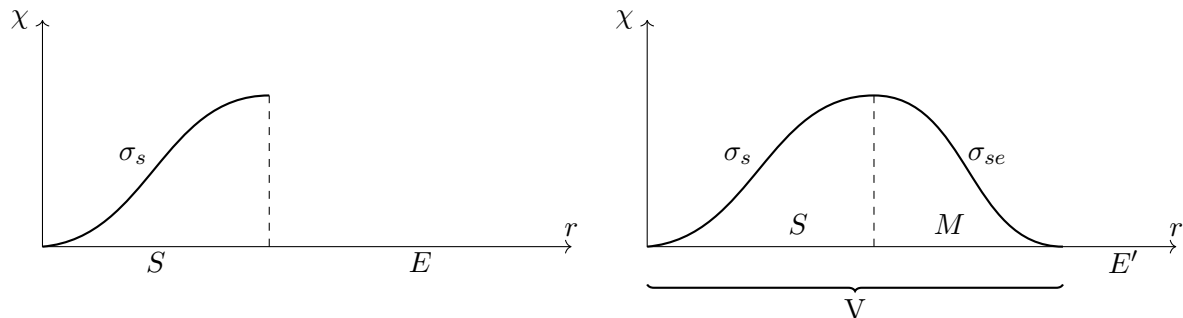


Figure 1: An illustration of FACT*. In the left, we have a gauge transformation σ_s acting on the subsystem S . In the right, we extend this symmetry by means of σ_{se} , a transformation that asymptotically goes to the identity (in a smooth manner) at some point in the environment. The resulting symmetry $\sigma_v = \sigma_s * \sigma_{se}$ is interior on the bigger subsystem V (V is the subsystem consisting of S and the region labelled ‘ M ’ in the figure).

tion σ_s in such a way that its extension, σ_{se} , goes smoothly to the identity at some point in the environment (here the subscript “ se ” stands for ‘extension of the subsystem symmetry’).⁷ The total transformation thus obtained (the one given by the initial symmetry and its extension) is interior on a bigger subsystem (see figure 1 for an illustration).

For the sake of simplicity, let me call ‘Trivial Extension’ any extension of a subsystem symmetry σ_s such that it asymptotes (in a smooth way) to the identity at some point in the environment. With this terminology, we can rewrite FACT* as follows:

FACT: for any gauge symmetry σ_s , there is always a Trivial Extension σ_{se} .

Here is then a simplified version of Friederich’s proof according to which gauge symmetries cannot be empirically significant:

- (1) Consider an arbitrary gauge transformation σ_s of a subsystem S .

⁷Friederich is not very clear about whether he means that any transformation can be extended to the identity in a smooth way or in a general (non-necessarily smooth) way. From the context of the discussion, it seems he means “smooth”, and so I will assume that in what follows. But my criticism still holds if he does not require smoothness.

- (2) From FACT, it follows that there is a Trivial Extension σ_{se} of σ_s .
 - (3) The transformation consisting of σ_s and σ_{se} , $\sigma_v = \sigma_s * \sigma_{se}$, is interior.
 - (4) Since it is interior, applying σ_v to the universe is equivalent to applying the universe symmetry $\sigma = \sigma_v * I_r$, where I_r refers to the identity transformation on the region of the environment not including V (that is, the transformation on the environment of V that is denoted by E' in figure 1).
 - (5) Thus, to apply σ_s together with σ_{se} is equivalent to applying the universe symmetry $\sigma = \sigma_v * I_r = \sigma_s * \sigma_{se} * I_r$.
 - (6) When a universe symmetry is applied to the universe, all of its restrictions to subsystems lack empirical significance.⁸
 - (7) By construction, σ_s is a restriction of the universe symmetry σ to subsystem S (see (5) above).
- C) Thus, σ_s does not have empirical significance.

(Given that σ_s is arbitrary, the above result holds for any gauge subsystem symmetry whatsoever.)

The problem with the previous argument is that the claim that σ_s does not have empirical significance does not follow from the premises. The reason is simply that in the proof it is assumed that σ_s is applied to S *together* with a Trivial Extension. Therefore, Friederich's argument only establishes the following weaker claim:

C*: When σ_s and a Trivial Extension σ_{se} are jointly applied, σ_s lacks empirical significance.

⁸Friederich derives (a particular instance of) this result from his assumptions SUL and DES (see the first equation in page 9 of his (2016)).

Importantly, this weaker claim is not a threat to the coherence of local Type I cases because C^* does not entail that σ_s lacks empirical significance in those cases where a Trivial Extension *is not* applied, and these are precisely the Type I cases. In particular, recall that Type I cases are characterized by transformations of the form $\sigma_s * I_e$ that leave the environment of s completely unchanged, whereas in the previous proof we are asked to consider a transformation of the form $\sigma_s * \sigma_{se}$, and obviously σ_{se} alters the state of a region of the environment of s (in figure 1, the region affected by σ_{se} is M).

To reinforce my point, it is helpful to attend to what Friederich actually says at the very beginning of his argument (my emphasis):

Let σ_s be a local symmetry defined on a subsystem state space \mathcal{S} . By Ext [FACT], *we can* extend it to an interior symmetry σ_v on \mathcal{V} , the state space associated with a larger subsystem $S \in V$ (Friederich, 2015, 549).

Notice that in the second sentence of the previous passage, Friederich correctly says that we *can* extend the subsystem symmetry (this is a consequence of FACT). But of course, from the fact that we can do so it does not follow that we must do so (for example, we can also extend the subsystem symmetry in a rigid manner that does not go to the identity!). From this point of the argument onward, Friederich focuses completely on that case where σ_s and σ_{se} are jointly applied, and he says nothing about cases where σ_s is applied in the absence of a Trivial Extension.⁹ In short, then, Friederich's proof does not establish that σ_s always lacks empirical significance but only that it lacks it when applied together with a Trivial Extension.

Interestingly, the fact that σ_s lacks empirical significance in the presence of a Trivial Extension is precisely the result we expected from both GW's framework and from the

⁹In his terminology, σ_{se} corresponds to $\sigma_{s'}$, and σ_v is expressed as $\sigma_v = \sigma_s * \sigma_{s'}$.

Constrained Hamiltonian approach as presented by Teh (2015). After all, when σ_s is applied together with σ_{se} , we are effectively applying a symmetry transformation σ_v that smoothly asymptotes to the identity and so that it is interior. And as we explained in section 2, ‘small’ transformations (those that go to the identity in a smooth way) always correspond to redundant descriptions of the physical content of the theory (2015, Sect 4). Thus, if σ_v is applied, σ_s and σ_{se} both must lack empirical significance precisely because σ_v is small. In other words, because of the essential role that FACT plays in the proof, Friederich’s result only concerns the transformations in \mathcal{G}_0^∞ , but not the ‘large’ ones (not the ones in $\mathcal{G}_I/\mathcal{G}_0^\infty$).

Although it does not succeed in proving that subsystem gauge transformations can never be empirically significant, Friederich’s proof brings to the forefront the fact that there is an ambiguity in the question ‘does applying σ_s to a subsystem bring about a physically distinct state of the universe?’ Naturally, to answer the question we consider a universe state that results from the application of σ_s to the subsystem and then see if that state is physically distinct from the initial one. Consider, for instance, this state: $\sigma_s(s) * \sigma_{se}(e) * I_r(r)$. From the fact that $\sigma_s * \sigma_{se}$ is interior, it follows that this state is empirically equivalent to $s * e * r$ (this is explained in the proof above). Then, the answer to our question is ‘no’, and so we infer that the subsystem symmetry is not empirically significant. But consider now the following application of σ_s to the subsystem: $\sigma_s(s) * I_e(e) * I_r(r)$. Since $\sigma_s * I_e$ is not interior, $\sigma_s(s) * I_e(e) * I_r(r)$ and $s * e * r$ do not need to represent the same states of affairs. Thus, in this case the answer to the question is ‘yes’ (or ‘maybe’, depending on the details of the theory and of the subsystem), and so we infer that the very same subsystem symmetry is (or can be) empirically significant.¹⁰ As explained before, Friederich’s argument only concerns

¹⁰This fact might seem trivial from the perspective of external global transformations, after all, boosting a ship is empirically significant if the shore is not boosted but it is not if the shore and the ship are boosted together. But here we are talking of local transformations, and there was not a *prima facie* reason to believe that this feature of external transformations was to be found in local transformations. Friederich’s argument shows that it is found.

the $\sigma_s(s) * \sigma_{se}(e) * I_r(r)$ case, and says nothing about the $\sigma_s(s) * I_e(e) * I_r(r)$ case.

So, to sum up, Friederich's argument does not establish that local subsystem symmetries always lack empirical significance but it establishes the weaker (but not for that uninteresting) claim that they do when a Trivial Extension accompanies the subsystem transformation. And incidentally, Friederich's argument is helpful at illustrating the fact that the empirical significance of a subsystem local transformation is not just a matter of what the transformation does to the state of the subsystem, but also a matter of what additional transformations we consider on the environment. In this way, the apparent tension between the results of Greaves and Wallace (2014) and Teh (2015) on the one hand, and Friederich (2015; 2016), simply disappear.

4 The divisive beam splitter

Having explored the more general arguments regarding the issue of whether or not gauge transformations can be empirically significant, let me focus now on a more concrete question. Is t'Hooft's Beam Splitter, presented by Hooft (1980), a concrete realization of a gauge transformation having empirical significance? Kosso (2000, 95), Greaves and Wallace (2014) and Teh (2015) think that it is, while Hooft (1980, 98), Brading and Brown (2004), and Friederich (2015) think it is not. But before considering the main arguments on the matter, let me briefly present the experiment.

4.1 The set-up

Consider a set-up similar to the double slit experiment, where a matter wave is sent towards a screen with two slits. Imagine that we manage to separate the outgoing two beams, and we place a phase-shifter that affects only the upper beam. At the end, we let the beams hit

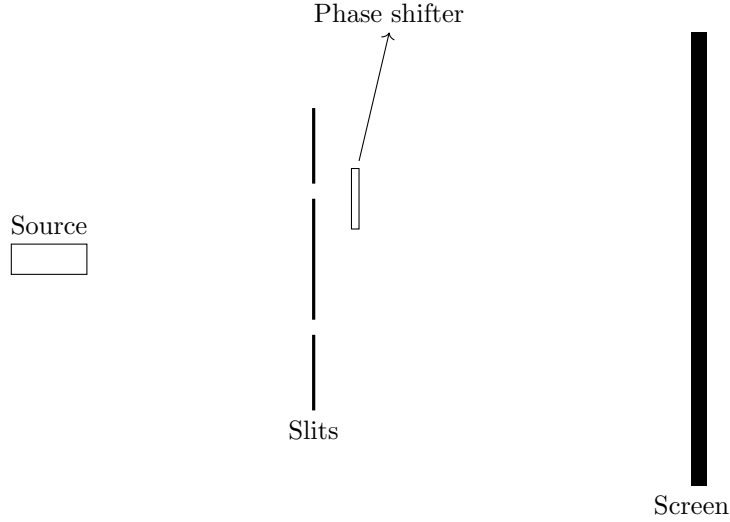


Figure 2

a screen, where we can see an interference pattern (see figure 2).

This situation can be modeled by what GW call “Klein-Gordon-Maxwell’ electrodynamics, where a matter wave ψ couples to a background electromagnetic field. The Lagrangian of the theory is

$$\mathcal{L} = (\partial_\mu \psi - iqA_\mu \psi)^* (\partial^\mu \psi - iqA^\mu) - m^2 \psi^* \psi + \mathcal{L}_{EM}, \quad (2)$$

where \mathcal{L}_{EM} is the Lagrangian of Maxwell Electrodynamics without matter. The Lagrangian of the theory is invariant under the following conjoint gauge transformation of the potential and the matter wave:

$$\begin{aligned} \psi(x) &\rightarrow e^{-iq\chi(x)} \psi(x) \\ A_\mu(x) &\rightarrow A_\mu(x) + \partial_\mu \chi(x), \end{aligned} \quad (3)$$

where $\chi(x)$ is a real-valued smooth-function on space-time that parametrizes the gauge transformation.

After the initial (matter) beam goes through the double slit, it splits into two beams, ψ_s and ψ_e , which can be separated enough so that they are effectively isolated from another (at least for some time before reaching the screen). Imagine, for example, that ψ_s is only non-zero within region R_s , whereas ψ_e is only non-zero within region R_e , where R_s and R_e do not overlap. Given this set-up, GW identify the subsystem with R_s , and its states are given by $s = (\psi_s, A_{\mu_s})$ (that is, by the matter and electromagnetic fields ‘living’ in region R_s).

In the case where there is not a phase-shifter (Situation *A*), we take (ψ_s, A_{μ_s}) to represent the state of the upper subsystem and (ψ_e, A_{μ_e}) to represent the state of the lower subsystem after the initial beam goes through the slits. In the case where we place a phase-shifter for the upper beam (Situation *B*), $(e^{-iq}\psi_s, A_{\mu_s})$ represents the state of the upper subsystem and (ψ_e, A_{μ_e}) represents the state of the lower subsystem just after the beam goes through the slits (notice that the phase shifter shifts the phase of the upper subsystem and does nothing to its potential). Furthermore, we know that the interference pattern at the screen will be different in Situation *A* and Situation *B* due to the fact that the relative phase between the subsystems is different. We can infer, then, that

INFERENCE: The interference pattern is sensitive to whether the state of the upper subsystem is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$. In particular, given the interference pattern, we can determine if the state of the upper beam was (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$.

So far we have not said anything about symmetries, but notice that $\psi_s \mapsto e^{-iq}\psi_s$ together

with $A_{\mu_s} \mapsto A_{\mu_s} + 0$ is precisely the action of the gauge transformation 3, *restricted* to the subsystem, for the specific case where χ is constant at the boundary between the subsystems (if $\sigma = e^{-iq\chi(x)}$ is the gauge transformation on the universe matter field, $\sigma_s = e^{-iq\chi_s(x)}$ is the restriction of that transformation to the subsystem matter field). Suppose that $\chi_s(x)$ approaches a non-zero constant value at the boundary between the subsystems (take that value to be 1 for simplicity). As it is constant there, $\partial_\mu \chi_s(x)$ will be zero and so the vector-potential of the subsystem will not be altered at the boundary. Similarly, as $\psi_s(x)$ is assumed to be zero at the boundary, $e^{-iq\chi_s(x)}\psi_s(x) = \psi_s(x) = 0$. Thus, $\sigma_s = e^{-iq\chi_s(x)}$ will preserve the boundary conditions for the matter and electromagnetic fields in question, and so $\sigma_s(s) * e$ will be well-defined.

Given this set-up, in GW's formalism Situation *A* will be represented as $s * e$ and Situation *B* as $\sigma_s(s) * e$ (the local symmetry in question, recall, does not affect the state of the lower subsystem, and so we can use e for both situations). Furthermore, given that $s * e$ and $\sigma_s(s) * e$ are empirically distinguishable (as INFERENCE states), it follows that

RESULT: the experiment in question corresponds to a Type I situation where a local symmetry, $\sigma_s = e^{-iq\chi_s(x)}$, has empirical significance (at the boundary, this symmetry simply is $\sigma_s = e^{-iq}$).¹¹

4.2 Friederich's criticism

Friederich offers two main criticisms of GW's analysis of the experiment just described. The first one is this:

(C1): Situation A and Situation B are not adequately represented by $s * e$ and

¹¹To be more precise, the gauge symmetry consists of both $e^{-iq\chi_s(x)}$ and $\partial_\mu \chi_s(x)$, but since for constant $\chi_s(s)$ the transformation on the gauge fields is the identity transformation, I will be omitting it in the following discussion.

$\sigma_s(s) * e$ respectively, but by $s * e$ and $\sigma_s(s) * e'$ (where $e \neq e'$). Thus, contrary to what RESULT states, this is not a Type I scenario.

The justification for (C1) is that if ψ_s is the subsystem, then not only ψ_e but everything else, including the screen where the interference is manifested, is part of the environment. And clearly, the state of the screen is different in Situation A and Situation B (the interference pattern is different), and so the state of the environment is different as well. In his words:

If the upper half-beam plays the role of the subsystem \mathcal{S} and the rest of the set-up plays the role of the environment E , then, if two situations with different interference pattern are compared, they must evidently be represented by physically distinct environment states $e \neq e$ (Friederich, 2015, 553).

Let me notice that what Friederich is expressing here is tightly connected to what Brading and Brown (2004) say when arguing against the suggestion that the transformation on ψ_s is a local transformation that has empirical significance. In particular, they say that “an interference pattern occurs only where Ψ_I and Ψ_{II} overlap” (2004, 653), strongly suggesting that in the absence of an overlap region (such as the screen), we should not attribute empirical significance to a local transformation on the isolated beams. Thus, by emphasizing the importance of taking into consideration the states of the screen (i.e., of an overlap region), Friederich seems to be echoing this remarks by Brading and Brown.

Now, I think that (C1) involves a misscharacterization of Type I scenarios. For in these scenarios, the term ‘environment’ is not supposed to refer to things like observers or measurement devices or other objects capable of detecting the (putative) effects of subsystem symmetries. That is, when representing the state of the environment with ‘ e ’, GW (or Teh (2015)) are deliberately not intending to also represent the state of measurement devices or

observers. So, in the case at hand, the environment is not supposed to include the screen, which simply serves as a measurement device of the putative physical effects of the action of the subsystem symmetry in question.¹²

That this is the intended use of ‘environment’—one not including observers or measurement devices—is clear from the analysis of a classic Galileo case scenario involving a ship passing by a shore. It is clear, not only from GW’s and Teh’s presentations of the case but also from Galileo’s own description (Galilei et al., 2001, 216), for whom the state of the shore is not supposed to include measurement devices (and the same is true of the cabin). Of course, a speedometer (or an observer) will be in a different state in the case where the ship is at rest and in the case where it is boosted, so in order to guarantee that the state of the shore remains invariant, we have to exclude speedometers and observers. And just as speedometers and observers are not taken to be part of the environment in a standard Galileo-ship case, screens or observers should not be taken to be part of the environment in the beam splitter experiment here discussed.

The previous objection to (C1)—namely, that (C1) mischaracterizes Type I scenarios—can be further strengthened by noticing that GW’s analysis is meant to be objective in the sense that if a symmetry is empirically significant (or not) according to their analysis, then this is so independently of whether something *actually* measures the effects of such a symmetry. That is, for a symmetry to be empirically significant there need not be any *actual* experiments showing its effects, but rather, the following counterfactual condition suffices; *if there were* observers or measurement devices, then they *would* detect the effects of the symmetry. Hence, even if an actual situation does involve devices and observers, abstracting away these should not affect judgments regarding the empirical significance of

¹²Friederich seems to anticipate this response in the second paragraph of (2015, 553). I will discuss his own objection to this response below, when I talk of (C2).

subsystem symmetries, and so GW are justified in taking the state of the environment to remain invariant in the case at hand. For example, abstracting away the presence of speedometers or observers in the shore should not undermine or decrease our confidence in the claim that boosts of ships are empirically significant. Similarly, abstracting away the presence of the screen in the beam splitter example should not affect the judgment that changes in the relative phase between ψ_s and ψ_e are empirically significant.

Let me consider now a different but related criticism by Friederich. He seems to think that, perhaps against our intuitions, the fact that the state of the screen is different in Situation A and in Situation B actually undermines the case for attributing empirical significance to the subsystem symmetry in question. More precisely, Friederich argues that:

(C2) It is not true that given the interference pattern, we can determine if the state of the upper beam is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$ (this goes against INFERENCE). Hence, we should not attribute empirical significance to the action of the subsystem symmetry represented by e^{-iq} (recall that $\chi = 1$ at the boundary).

To understand why Friederich thinks this, it is helpful to go back to GW's description of the case. Recall that in Situation A, (ψ_s, A_{μ_s}) represents the state of the upper subsystem and (ψ_e, A_{μ_e}) represents the state of the lower subsystem. In Situation B, $(e^{-iq}\psi_s, A_{\mu_s})$ represents the state of the upper subsystem after going through the phase-shifter, and (ψ_e, A_{μ_e}) represents the state of the lower subsystem. Given the fact that the interference patterns are different in these situations, and the fact that (ψ_e, A_{μ_e}) is the state of the lower beam in both situations, it seems that *the interference pattern is sensitive to whether the state of the upper subsystem is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$* (this is what INFERENCE says). And because of this, it seems natural to infer that the subsystem symmetry has empirical

significance (this corresponds to RESULT).

Now, in order to understand how Friederich intends to resist the previous reasoning, we have to start by distinguishing ψ_s and ψ_e when isolated, from ψ_s and ψ_e when overlapping at the screen. Friederich shows that we can recover the interference pattern of Situation B and Situation A by swapping the states that GW use. That is, he shows that we can recover the interference pattern of Situation B by using (ψ_s, A_{μ_s}) (instead of $(e^{-iq}\psi_s, A_{\mu_s})$) to represent the state of the upper beam when isolated, and we can recover the interference pattern of Situation A by taking $(e^{iq}\psi_s, A_{\mu_s})$ (instead of (ψ_s, A_{μ_s})) to represent the state of that beam when isolated (of course, this is only possible by modifying the way we represent the states of the beams in the overlapping region (Friederich, 2015, 554)).¹³ Hence, it seems that from the interference pattern alone we would not be able to tell if the subsystem state of the upper beam is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$ (by a suitable choice of the states in the overlapping region, both of these states can lead to the same interference pattern). And so, the argument goes, we should not attribute empirical significance to the local transformation (given by e^{-iq}) of the upper subsystem.

Let me explain now why Friederich's criticism fails. The problem lies in Friederich overlooking an underlying assumption of GW's set-up, namely, that when GW say that the interference pattern is sensitive to whether the state of the upper subsystem is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$, they are implicitly assuming that (ψ_s, A_{μ_s}) represents the state of the upper beam in the absence of a phase-shifter (that is, since GW express the operation of the gauge transformation acting on the wave-function as $\psi_s \mapsto e^{-iq\chi(x)}\psi_s$, then it follows that they take ψ_s to be the state prior to any transformation). Given that assumption, the interference pattern does allow us to determine if the state of the upper beam is (ψ_s, A_{μ_s}) or $(e^{-iq}\psi_s, A_{\mu_s})$ (it is (ψ_s, A_{μ_s}) in the absence of a phase-shifter, and $(e^{-iq}\psi_s, A_{\mu_s})$ in the

¹³Notice that the swapping is not perfect, we need to use a plus sign and not a negative sign.

presence of one).

Of course, Friederich is right in that we could use (ψ_s, A_{μ_s}) to represent the state of that same beam in Situation B, but in that case we would have to assume that in the absence of a phase-shifter, the state of the beam is $(e^{+iq}\psi_s, A_{\mu_s})$ so that the transformation by e^{-iq} yields (ψ_s, A_{μ_s}) . But even then we could still use the interference pattern to discriminate between the relevant states of the beam (as represented in this alternative form). That is, under this new convention for representing states, the interference pattern will still allow us to determine if we are in Situation A or Situation B (Situation A will be characterized by $(e^{+iq}\psi_s, A_{\mu_s})$ and Situation B by (ψ_s, A_{μ_s})).

So, to sum up, the main point of GW's argument is not, contrary to what Friederich seems to be suggesting in (C2), that the interference pattern will by itself allow us to determine the state of the upper and lower beams (this is impossible because of the phase-freedom of quantum states). Rather, the point is that given a convention of what state s we take to represent the state of the upper beam in the absence of a phase-shifter, the interference pattern in the screen will allow us to determine if the state of the upper beam remains as it is (i.e., s) or if it changes (i.e., $e^{-iq}(s)$). In other words, all GW require in order to attribute empirical significance to the action of the local symmetry in question is that, no matter what representation of the relevant state we use, the symmetry in question will induce changes in the relative phase between the two subsystems, and these changes will be measured. Thus, the interference pattern would be able to track whether the subsystem symmetry in question "acted" or not on that beam, from which we infer that the symmetry is empirically significant.

At this point, going back to the original Galileo ship scenario can be helpful. Imagine that Situation A consists of the ship staying at rest relative to the shore, and Situation B consists of a boost of the ship. Then, it is natural for us to say that if the state of the ship

in Situation A is represented by $s = 0 \text{ m/s}$, then its state in Situation B is represented by $\sigma_s(s) = v$ (where σ_s represents a boost of non-zero velocity v). But of course, we could also say that $s = 0 \text{ m/s}$ represents the state of the ship in Situation B in case where we use a frame in which $\sigma'_s(s) = -v$ represents the state in Situation A (in the latter case, we take the ship and the shore to be moving with the same negative velocity of magnitude v). In short, because of its symmetries, classical mechanics entails that there is a lot of freedom in the way we represent the states of the ship and the shore. But it would be a mistake to think that the fact that we can easily move from one representation to the other (say, from the state of the ship at rest being represented by 0 m/s to it being represented by $-v$) undermines in any way the claim that boosts of the ship are empirically significant. Analogously, from the fact that we can represent the state of the upper beam in different ways (i.e. as $(e^{-iq}\psi_s, A_{\mu_s})$ or (ψ_s, A_{μ_s})) we should not infer that the local transformation acting on the upper subsystem is not empirically significant.

4.3 An ambiguity in GW's description

Let me end with a criticism expressed by Brading and Brown (BB henceforth) in section 3.1.2 of their (2004) paper. Although GW respond to the main part of their criticism (Greaves and Wallace, 2014, 83), there is an important bit that remains unanswered. In particular, BB say that (my emphasis)

either the transformation of the electromagnetic potential results in the potential being discontinuous at the boundary between the 'two subsystems,' *in which case the relative phase relations of the two components are undefined* (it is meaningless to ask what the relative phase relations are), or the electromagnetic potential remains continuous, in which case we have a special case of a local

gauge transformation on the entire system (Brading and Brown, 2004, 656).

I agree with BB that if the transformation is continuous between the two subsystems, then this will not constitute a case of a gauge transformation with empirical significance. However, it is not clear that if the gauge transformation in question is discontinuous at the boundary, then the relative phase relations of the two components (i.e., the two beams) are undefined. Indeed, given what was said in section 4.1, it seems that these phase-relations are well-defined: if the relative phase between the beams is well-defined after the placement of a phase shifter for the upper beam, then the relative phase between the beams has to be well-defined after a gauge transformation that shifts the phase of the upper beam without affecting the phase of the lower beam. The reason is that in the set-up of the experiment, the gauge transformation in question perfectly replicates the change of relative phase coming from the placement of a phase-shifter on the upper beam. Thus, as long as it replicates the (well-defined) change in relative phase that is coming from the phase shifter, it should not matter that the transformation in question is discontinuous (again, if it was continuous, it would not actually induce the change in relative phase that we need).

Now, there is a sense in which the criticism by BB actually goes deeper than what I just said above. For by protesting about the discontinuity of the transformation at the boundary between the subsystems, they bring attention to the following confusing feature of the description of the experiment: the two beams are treated as if they were both two subsystems that interact at a given finite boundary (at a screen, say). Leaving aside the subtle issue about how to individuate subsystems in quantum mechanics, this description is problematic for two reasons. First, because, as it was explained in section 2, local transformations of a subsystem can be empirically significant (in a Galileo-type way) only if (1) we take the subsystem to be asymptotically isolated from the environment and (2) take the

environment to play the role of an asymptotic fixed reference frame (with respect to which the gauge transformation in question induces some relational changes) and whose dynamics is deliberately ignored. Thus, strictly speaking, the lower beam should be taken as something like a fixed frame at infinity whose dynamics is omitted in the analysis. And second and more importantly, if we were to take seriously that these two beams are two subsystems that interact at a finite boundary, then we would have to deal with very subtle issues (that GW do not address) regarding the composition of gauge subsystems in the presence of boundaries. For example, it is an open question whether or not gauge transformations leave the subsystem invariant in the presence of boundaries (see Gomes (2019) for a good discussion of this problem, and for a novel solution).

5 Conclusion

In this paper, I showed that the tension between the results of Greaves and Wallace (2014) and Teh (2015) and those of Friederich (2015, 2016) is only superficial. Once we realize that Friederich’s argument is only about transformations that asymptote in a smooth way to the identity at some point of the environment, it becomes clear that the proof entails that ‘interior’ transformations (in GW’s sense) or ‘small’ transformations (in Teh’s sense), are not empirically significant. Since the proof is not about transformations that are not trivial at the boundary, it is silent about precisely those cases that both GW and Teh take to exhibit some empirical significance. There is, then, no actual tension.

I also showed that the main arguments against the interpretation of t’Hooft’s Beam Splitter as a Galileo-ship scenario for gauge symmetries do not succeed. In particular, I explained that (i) the state of the screen should not be considered as part of the environment, that (ii) the freedom in the representation of the states of the upper and lower beams (when

isolated) does not preclude in any way the fact that the symmetry leads to empirically distinct scenarios, and that (iii) the way of understanding the empirical significance of this set-up requires us to treat the lower subsystem as a mere reference frame at the asymptotic boundary.

References

- Brading, K. and Brown, H. R. (2004). Are gauge symmetry transformations observable? *The British Journal for the Philosophy of Science*, 55(4):645–665.
- Dougherty, J. (2019). Large gauge transformations and the strong cp problem. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.
- Friederich, S. (2015). Symmetry, empirical equivalence, and identity. *The British Journal for the Philosophy of Science*, 66(3):537–559.
- Friederich, S. (2016). Symmetries and the identity of physical states. *EPSA15 Selected Papers*, 5:153–166. A follow-up to <http://philsci-archive.pitt.edu/9816/>.
- Galilei, G., Drake, S., Einstein, A., and Heilbron, J. (2001). *Dialogue Concerning the Two Chief World Systems, Ptolemaic and Copernican*. Modern Library Classics. Modern Library.
- Gomes, H. (2019). Gauging the boundary in field-space. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67:89 – 110.

- Greaves, H. and Wallace, D. (2014). Empirical consequences of symmetries. *The British Journal for the Philosophy of Science*, 65(1):59–89.
- Healey, R. (2009). Perfect symmetries. *The British Journal for the Philosophy of Science*, 60(4):697–720.
- Hooft, G. (1980). Gauge theories of the forces between elementary particles. *Scientific American*, 242(6):104.
- Kosso, P. (2000). The empirical status of symmetries in physics. *The British Journal for the Philosophy of Science*, 51(1):81–98.
- Teh, N. J. (2015). Galileos gauge: Understanding the empirical significance of gauge symmetry. *Philosophy of Science*, 83(1):93–118.