A puzzle concerning local symmetries and their empirical significance

Sebastián Murgueitio Ramírez

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 argue that there are strong reasons to resist Friederich's proof because one of its assumptions is, at the very least, highly controversial. Second, I argue that, despite criticism by Brading and Brown (2004) and Friederich (2015), 't Hooft's Beam-Splitter experiment is indeed a concrete example 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

Contents

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, the conventional wisdom has come under attack. In particular, Greaves and Wallace (2014) developed a general framework within which it is transparent that there are cases where local symmetries produce empirically distinct physical states. By using the

Constrained Hamiltonian formalism, Teh (2015) further clarifies the conditions under which this result holds. In response, by generalizing the framework of Greaves and Wallace (GW henceforth) and by appealing to some very general principles regarding the individuation and composition of physical systems, Friederich (2015, 2016) has attempted to prove the result that gauge transformations[1] 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').[2] This leads to a puzzle: how can it be that if one adopts GW's framework and the Constrained Hamiltonian formalism, then some local 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 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, p. 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 frameworks, but a consequence of the fact that Friederich's proof involves

[1]In this paper, I follow physicists in using 'local transformations' and 'gauge transformations' as interchangeable. As I use it here, a local (or a gauge) transformation is a function from space-time to a gauge group.

[2]It 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)).

seemingly false assumptions. In particular, as I will explain in detail in section 3, Friederich introduces an assumption that entails—contrary to both GW and Teh's detailed analyses on this matter—that neither the asymptotic behaviour of gauge transformations, nor the relations between subsystem transformations and environment transformations, are of any relevance to the empirical significance of subsystem transformations.

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 argue that it is not sound. In section 4, I argue 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 (for instance, 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 phase of a wave-function 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, whereas a global transformation does not depend on space-time. For example, transformations of the four-vector potential in electromagnetism are local transformations, whereas boosts are global transformations. Now, it is helpful to note that the terms 'gauge transformation' and 'local transformation' are sometimes taken to mean transformations that relates different descriptions of the same physical state. 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.[3]

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, I will say that a symmetry lacks empirical significance if and only if any two symmetry-related states of the subsystem (or the universe) are empirically indistinguishable (or, equivalent, I will say that in such a case, the symmetry 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, I will say that a symmetry is empirically significant if there exists at least two symmetry-related states which are em-

[3]Teh (2015, p. 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.

pirically distinguishable.[4] 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 (for example, 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 the essential concepts of that framework.

At the heart of GW's 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 tuple 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 every combination of states of the environment and states of the subsystem will yield 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,

[4]For 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)).

2014, p. 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 $s \in \mathcal{S}$, $e \in \mathcal{E}$, $\sigma(u) = \sigma(s * e) = \sigma_S(s) * \sigma_E(e)$ where 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, $\sigma_s(s) * e$ 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, only non-interior symmetries could be empirically significant.

According to GW, there are two main cases 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 a Faraday cage scenario in electrostatics is an example of this kind of case—henceforth 'Type II' case—, where the subsystem is the interior region of the cage and the environment is the surface of the conductor. As we know, shifts in the scalar potential in the interior (these shifts are both an internal and a global symmetry) are associated to changes in the surface charge. And changes in the surface charge 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.[5]

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 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 accommodates 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 does their framework allows for the possibility of local transformations that are empirically significant (a possibility that Brading and Brown (2004, p. 657), Healey (2009) and Friederich (2015) want to deny)[6], but GW also say that 't Hooft's Beam Splitter experiment is precisely a concrete realization of a local Type I case. As we will study to greater detail in section 4, 't Hooft's Beam Splitter corresponds to a modified version of a standard double-slit experiment, where one adds a phase-shifter behind one of the slits. For GW, the change in interference pattern coming from the phase shift on the corresponding beam suggests that a local transformation can be empirically significant.

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

[5]For a careful recent analysis of Type II cases in the context of gauge theories, see Murgueitio Ramírez and Teh (2019).

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

Galileo ship-type scenarios. By using the Constrained Hamiltonian formalism, Teh (2015) sheds light onto those conditions by showing that the empirical significance of a gauge transformation is sensitive to the asymptotic behaviour of the transformation.[7] 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 deformed 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 $\mathcal{G}_0^\infty$, the group of gauge transformations that, asymptotically, is smoothly connected to the identity. Another one is $\mathcal{G}^\infty$, which corresponds to the group of transformations that go asymptotically to the identity (but might not be smoothly deformed 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, p. 115):

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

$\mathcal{G}_0^\infty$ 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 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 I case) that are not in $\mathcal{G}_0^\infty$ (for instance, 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

[7] In this paper, I do not examine 'non-asymptotic ways' by which local transformations can be empirically significant (but see Gomes (2019a) and Gomes (2019b) for very recent studies of those cases).

boundary).[8] Finally, let us point out that Teh also agrees with GW that 't Hooft's Beam Splitter experiment offers an example of a local transformation that has empirical significance (Teh, 2015, p. 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 could endorse the framework (or part of it) and show that, despite the fact that it seems to allow for the possibility of Type I cases of local symmetries, there are reasons to believe that Type I cases of 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 elegantly classifies uncontroversial cases of empirical significance. It makes precise how external transformations (such as boosts and rotations) of isolated subsystems can be empirically significant; how shifts in the scalar potential in the interior of a conductor are empirically significant in the case of electrostatics; and, finally, it explains the manner in which 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

[8] Teh 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.

to be on the right track.

Second, the framework has the virtue of being so general that it can easily be 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, p. 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, p. 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, p. 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 are important problems with the proof by Friederich.

The following fact about gauge symmetries plays a crucial role in the upcoming argument:

> 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, p. 548).[9]

In other words, FACT* says that we can always extend a subsystem gauge transformation σ_s in such a way that its extension, σ_{se}, can be smoothly deformed to the identity at some point in the environment (here the subscript '*se*' stands for 'extension of the subsystem symmetry'). 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 asymptotically can be smoothly deformed to the identity (notice

[9] Although 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, p. 9) paper.

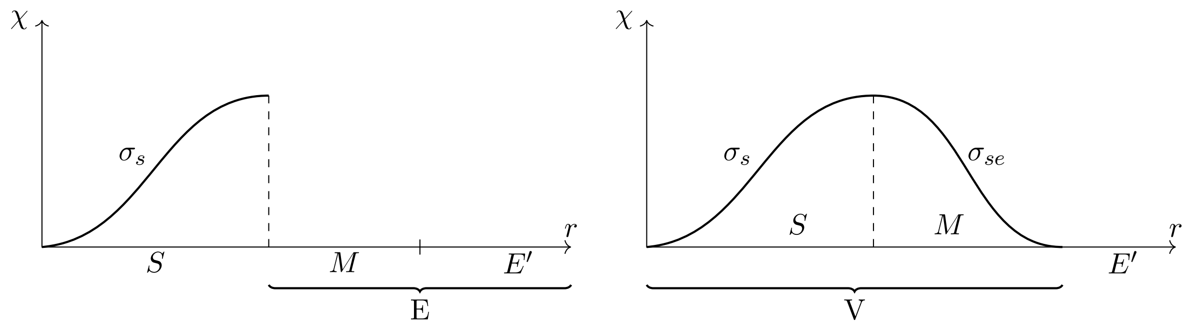

Figure 1: An illustration of FACT*. On the left, we have a gauge transformation σ_s acting on the subsystem S. On the right, we extend this symmetry by means of σ_{se}, a transformation that asymptotically can be smoothly deformed to the identity. The resulting symmetry $\sigma_v = \sigma_s * \sigma_{se}$ is interior on the bigger subsystem V (V is the subsystem consisting of S and the region labelled 'M' in the figure). Notice that M together with E' constitute the environment of S (that we call 'E').

that the composition of the original symmetry with a trivial extension constitutes an interior symmetry). With this terminology, we can rewrite FACT* as follows:

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

Besides FACT, Friederich's proof requires three other assumptions that he calls 'DES', 'SUL' and 'MAH'. He also uses the symbol '$\sim$' to express the relation of representing or denoting the same physical state (2015, p. 546). In particular, '$s \sim s'$' means that s and s' represent (or denote) the same subsystem state, '$u \sim u'$' means that u and u' represent the same universe state, and so on. Notice that if $u \sim u'$, then the physical state represented by 'u' and the physical state represented by 'u'' are, trivially, empirically equivalent because these are one and the very same physical state. And for Friederich, if the physical state represented by 'u' and the physical state represented by 'u'' truly are empirically equivalent, then these physical states must be one and the same. In particular, he says that 's and s' designate the same physical state if they are empirically equivalent both from within the subsystem and from the perspectives of arbitrary external observers' (2016, p. 4). Hence, 'physical equivalence' ('$\sim$') and 'empirical equivalence' are themselves equivalent terms.

The assumption denoted by 'DES' says the following:

> **(DES)**: A subsystem symmetry σ_s has direct empirical significance iff $\sigma_s(s) \nsim s$ for some s.

Now, in order to explain how the notion of *direct empirical significance* connects to universe states, Friederich introduces the following assumption:

> **(SUL)**: For all $s, s' \in \mathcal{S}$,
> $s \sim s'$ iff $s * e \sim s' * e$ for all $e \in \mathcal{E}$ for which $s * e$ and $s' * e$ are defined.

If we take DES and SUL together, the following result follows immediately:

> DES-SUL: if $s * e$ and $\sigma_s(s) * e$ are empirically distinct states of the universe for some e for which they are well-defined (if $s * e \nsim \sigma_s(s) * e$), then σ_s has (direct) empirical significance.

Notice that DES-SUL essentially states what a Galileo-ship case is, where the environment is not altered and only the subsystem state is transformed.

Finally, Friederich introduces an assumption related to how the states of composite systems relate to the states of the subsystems. To facilitate the presentation of both the assumption and the proof, it will be useful to distinguish between three regions of the universe, namely, the region corresponding to subsystem S (with states $s \in \mathcal{S}$), the region corresponding to subsystem M (with states $m \in \mathcal{M}$), and the region corresponding to E' (with states $e' \in \mathcal{E}'$). As figure 1 illustrates, region M together with region E' constitute the environment E (with states $e \in \mathcal{E}$) of S. Also, S together with M correspond to a subsystem

V (with states $v \in \mathcal{V}$) whose environment is E'. Given this, the very same universe state can be written in different ways depending on what subsystems we are focusing on. For example, a given universe state u can be written as $u = s * m * e'$ if we focus on S, M and E', or it could be written as $u = s * e$ if we focus on S and E (with $e = m * e'$), or it could be written as $v * e'$ if we focus on V and E' (with $v = s * m$).[10] We then have all the tools to properly understand the next assumption.

> **(MAH)**: For all $s, s' \in \mathcal{S}$ and $m, m' \in \mathcal{M}$,
> if $s * m * e \sim s' * m' * e$ for all $e \in \mathcal{E}$ for which $s * m * e$ and $s' * m' * e$ are defined,
> then $s \sim s'$ and $m \sim m'$.

Roughly, MAH says that if two universe states are empirically equivalent, then the apparently distinct subsystem states of those universe states are actually one and the same physical state (recall that '$s \sim s'$' means that 's' and 's'' denote the same state). And so, in particular, it follows from MAH that if $s * m * e \sim \sigma_s(s) * \sigma_m(m) * e$ for all e for which they are defined, then $s \sim \sigma_s(s)$. Given DES, this means that σ_s is not empirically significant.

Given these ingredients, we can run Friederich's proof:

(1) Consider an arbitrary gauge transformation σ_s of a subsystem S, and an arbitrary state $s \in \mathcal{S}$.

(2) From FACT, it follows that there is a Trivial Extension σ_{se} of σ_s.

[10] Notice that when Friederich uses '$*$', he is not using it exclusively as an operation between two states that together form a universe state (even though this was the way Greaves and Wallace (2014, p. 68) introduced it). For example, when he writes things such as '$s * m * e'$', we see that $s * m$ and $m * e'$ are not universe states but rather subsystem states. So Friederich uses '$*$' as a generic operator that represents the composition of subsystem states independently of whether or not these states form a putative universe state. For simplicity, instead of introducing a new symbol to represent the operation in question, I will just follow Friederich's conventions. For convenience, I will also use '$*$' to denote the composition of symmetries, such as when we write $\sigma_s * \sigma_m$ (it should be obvious from context how the symbol in question is being used).

(3) The transformation consisting of σ_s and σ_{se}, $\sigma_v = \sigma_s * \sigma_{se}$ is *interior*. That is, $\sigma_s(s) * \sigma_{se}(m) * e' \sim s * m * e'$ (for all e' for which these are defined).

(4) From MAH, it follows that $\sigma_s(s) \sim s$.

(C) From DES, it follows that σ_s is not empirically significant.

Although the proof is logically valid, there are good reason to doubt its soundness. For notice how the proof works: it uses the fact that local transformations that are trivially extended do not lead to empirically distinct states of the universe (for instance, it uses the fact that $\sigma_s(s) * \sigma_{se}(m) * e' \sim s * m * e'$) in order to conclude that no local transformation, including the ones that are not trivially extended, is empirically significant. The problem with this kind of argument is that transformations that are trivially extended are a very particular type of transformation because they exhibit features (to be explained below) that transformations that could be empirically significant do not. Thus, we should not expect that trivially extended transformations offer reliable guidance regarding the empirical significance of precisely those transformations that are putative candidates for being empirically significant. Having said this, the proof is logically valid, and so to resist it we need to reject at least one assumption.

3.1 Problems with the proof

It is not hard to realize that the most controversial assumption is MAH. For notice that MAH allows us to assert a very general claim about the (lack of) empirical significance of σ_s just from the fact that universe state $\sigma_s(s) * \sigma_{se}(m) * e'$, where σ_s is trivially extended, is empirically equivalent to universe state $s * m * e'$ (see steps 3 and 4 of the proof). In particular, MAH links the empirical equivalence between $\sigma_s(s) * \sigma_{se}(m) * e'$ and $s * m * e'$ to both $\sigma_s(s) \sim s$ and to $\sigma_{se}(m) \sim m$. But if 's' and '$\sigma_s(s)$' denote the same subsystem state (for instance, if $\sigma_s(s) \sim s$), and if 'm' and '$\sigma_{se}(m)$' also denote the same subsystem state, then, obviously, '$\sigma_s(s)*\sigma_{se}(m)*e'$' and '$\sigma_s(s)*m*e'$' denote the same universe state whenever these states are well-defined. But that this follows from MAH is precisely the reason this premise is controversial. For the states $\sigma_s(s) * \sigma_{se}(m) * e'$ and $\sigma_s(s) * m * e'$ are different in at least three relevant respects. First, the environment state of S in $\sigma_s(s) * \sigma_{se}(m) * e'$ is $\sigma_{se}(m) * e'$, which is different from $m * e'$—the environment state of S in $\sigma_s(s) * m * e'$. Second, and related to the first point, in universe state $\sigma_s(s) * \sigma_{se}(m) * e'$, σ_s is trivial on the boundary between S and M. However, in universe state $\sigma_s(s) * m * e'$, σ_s is not trivial on the boundary between S and M. Third, the asymptotic behaviour of σ_s is very different in both cases. In one case it goes to the identity, in the other one it does not. I will now explain why these three differences are crucial to the question of the empirical significance of $\sigma_s(s)$, which will in turn explain why MAH is problematic.

3.1.1 Problem 1: from changed to unchanged environment states

Let's start by recalling from section 2 what a standard 'Galileo-ship case' of empirical significance is (I called it a 'Type I' case). This is a case where the subsystem is transformed and the environment is not. In particular, it is a case where $s * e$ and $\sigma_s(s) * e$ are both well-defined and represent empirically distinct states of the universe. In the present context, where $m*e'$ is the state of the environment of S, this means that $\sigma_s(s)*m*e'$ and $s*m*e'$ are both well-defined and represent empirically distinct states of the universe. Having said this, notice that in Friederich's proof we do not explicitly consider whether or not $\sigma_s(s) * m * e'$ and $s*m*e'$ are empirically equivalent—which is the question we are ultimately interested in answering when asking about σ_s's empirical significance. Rather, in the proof, we are asked to compare $\sigma_s(s) * \sigma_{se}(m) * e'$ and $s * m * e'$ and to notice they are empirically equivalent. And due to MAH, we can then infer that $\sigma_s(s) * m * e'$ and $\sigma_s(s) * \sigma_{se}(m) * e'$ are empirically equivalent simply because $\sigma_s(s) * \sigma_{se}(m) * e'$ and $s * m * e'$ are (assuming, again, that all these are well-defined universe states). But this kind of inference should be particularly controversial because the environment state of S is not the same in $\sigma_s(s) * \sigma_{se}(m) * e'$ and in $\sigma_s(s) * m * e'$—in the first case, it is $\sigma_{se}(m) * e'$ and in the second is it is $m * e'$.

To reinforce this last point, let me note that, *prima facie*, the empirical significance of a subsystem transformation should depend on whether or not the environment is transformed. For example, in a universe with two objects, a boost of one object will be empirically significant if the other object (the environment) is not boosted, but it is not empirically significant if that second object is boosted in the same way. In general, whether a transformation is empirically significant (in the Galileo-ship sense) should depend on whether or not the environment is transformed. Yet, if MAH is true, then the transformations on the environment (given by $m * e' \mapsto \sigma_{se}(m) * e'$ in our case) are not relevant to the empirical significance of local transformations σ_s in the subsystem.

3.1.2 Problem 2: from triviality on the boundary to non-triviality

Second, in the context of field theories, GW are clear that Type I cases will only occur when σ_s is not trivial on the boundary with the environment (2014, p. 79). This is due to the fact that the empirical significance of Type I cases is due to relational differences between the subsystem and the environment, and such relational differences require that the transformation in question is not trivial on the respective boundary with the environment. In the present context, for it to produce relational changes, σ_s must not be trivial on the boundary between S and M (recall that M is contiguous to S). But how do we know that σ_s is not trivial on the boundary with M? The answer depends on what transformations we consider on M. If we do not change M's states, then any transformation σ_s that does not go to the identity on a neighbourhood of the boundary with M will not be trivial (as in the left image of figure 1). However, if we change the states in M, then to determine if σ_s is trivial or not requires finding out if the shift induced by σ_s on states in a neighbourhood of the boundary is of the same magnitude as the shifts induced by the transformation σ_m acting on states $m \in \mathcal{M}$ on the boundary. If the shifts are of the same magnitude (for instance, $\sigma_s|_{\partial S} - \sigma_m|_{\partial S} = 0$), so that $\sigma_s * \sigma_m$ is continuous on the boundary, then σ_s is indeed trivial on the boundary and will not induce the relational changes required for Type I cases (as in the right image of figure 1). If, on the other hand, the shifts on the boundary are not the same and a discontinuity between σ_s and σ_m arises, then σ_s will not be trivial on the boundary and so might be empirically significant. Having said this, it is clear that σ_s cannot produce any relational differences with respect to M in universe state $\sigma_s(s) * \sigma_{se}(m) * e'$, for in that case it is trivial on the boundary (σ_{se} and σ_s perfectly match in the boundary). So it is clear that for GW, $\sigma_s(s) * \sigma_{se}(m) * e'$ is not a universe state that is relevant for the investigation of σ_s's empirical significance. And yet, according to MAH, cases where σ_s is trivial on the boundary (for instance, $\sigma_s(s) * \sigma_{se}(m) * e'$) suffice to prove that σ_s is not empirically significant even when it is not trivial in the boundary. We should be doubtful of a principle that helps us prove a very general result just from the consideration of a very particular kind of case (i.e, the case where σ_s is trivial on the boundary).

3.1.3 Problem 3: from small transformations to large transformations

Third, Teh's analysis directly contradicts MAH (in footnote 28, Teh explicitly says that MAH is problematic (2015, p. 110)). For recall that one of the main points of Teh's analysis is that the asymptotic behaviour of σ_s (its 'extension') is central to its empirical significance. In particular, as we discussed in section 2, Teh shows that we need to distinguish between 'small' transformations (those that asymptotically can be smoothly deformed to the identity) and 'large' transformations (those that cannot be smoothly deformed to the identity but still satisfy the subsystem's boundary conditions). Teh clearly explains that only large transformations can be empirically significant. However, according to MAH, it is sufficient to use small transformations in order to derive the result that large transformations are not empirically significant. To be as explicit as possible, let me prove this in three main steps.

STEP 1: consider a universe state with three regions, S, M and E, and assume that E is asymptotically far away from S. Now consider a transformation given by $\sigma_s * \sigma_{se}$ (where σ_{se} is a trivial extension of σ_s) that asymptotically goes to the identity (in particular, the transformation vanishes at E). Thus, $\sigma_s * \sigma_{se}$ is a small transformation and so lacks empirical significance (in GW's terminology, this transformation is interior). Hence, we have that $\sigma_s(s) * \sigma_{se}(m) * e \sim s * m * e$. From MAH, we then get that

Result 1: $\sigma_{se}(m) \sim m$.

STEP 2: consider another very simple transformation on universe state $s * m * e$ that consists of a rigid shift by the same amount everywhere. To be concrete, consider the transformation $\sigma_c(s) * \sigma_c(m) * \sigma_c(e)$ that shifts s, m and e by amount c. For simplicity, let's choose c to match σ_s (from the first step) on the boundary between S and M. Obviously, $\sigma_c(s) * \sigma_c(m) * \sigma_c(e) \sim s * m * e$ (a constant shift must produce an empirically equivalent state of the universe). From MAH, we then get

Result 2: $\sigma_c(m) \sim m$.

STEP 3: from Result 1 and Result 2 it follows that $\sigma_{se}(m) \sim \sigma_c(m)$. And if this is true, then

Result 3: $\sigma_s(s) * \sigma_{se}(m) * e \sim \sigma_s(s) * \sigma_c(m) * e$

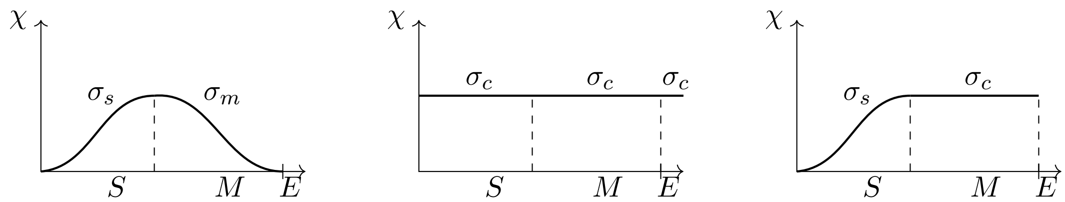

Figure 2: We assume that E is asymptotically far away from S. In that case, the left figure represents a small transformation given by $\sigma_s * \sigma_m$, where σ_m is a trivial extension of σ_s. The middle figure represents a constant shift everywhere (including E). The right figure represents a large transformation given by $\sigma_s * \sigma_c$. Notice that it rigidly extends just upon E. By using MAH alone, one can show that the universe states that result from these three transformations are all empirically equivalent when defined.

when these states are well-defined. But the transformation given by $\sigma_s * \sigma_c$ is a *large* transformation for it extends rigidly up until E (see figure 2). Hence, just by using MAH and the meaning of '$\sim$', we have proved that an arbitrary small transformation $\sigma_s * \sigma_{se}$ and an arbitrary large transformation $\sigma_s * \sigma_c$ produce empirically equivalent universe states. Since small transformations are not empirically significant, then it must follow that large transformations are not either.

Using the notation of section 2, what I just showed is that MAH allows us to infer that none of the transformations in $\mathcal{G}_I$ are empirically significant simply from the fact that the transformations in $\mathcal{G}_0^\infty$ (the smaller subgroup of the hierarchy in question) is not empirically significant. That is, MAH entails, contrary to Teh's analysis, that the asymptotic behaviour of gauge transformations is not relevant to their empirical significance.

3.1.4 Problem 4: unmotivated

Not only is MAH problematic for the reasons just explained, but I want to end this section by pointing out that MAH is not well motivated by Friederich himself. In order to motivate MAH, Friederich invites us to think of cases whereby changing the state of one subsystem, one can 'compensate' for these changes (for example, 'cancel their effect') by considering changes to another subsystem (Friederich, 2015, p. 547). In particular, he asks us to consider this kind of case: imagine a universe that has three objects, A (a ship), B (a shore) and C (another island). If we leave the state of C fixed, it is clear that any (non-trivial) boost on the first object will inevitably lead to an empirically distinct state of the universe, and this is true no matter how we boost B. In other words, if A is boosted, there is no transformation of B that will have the result that the boost on A is no longer empirically significant (assuming C stays as it is). If we use Friederich's terminology, what we have just shown is that, for the case of boosts, if $a * b * c \nsim \sigma_a(a) * b * c$, then there is no σ_b such that $a * b * c \sim \sigma_a(a) * \sigma_b(b) * c$. For Friederich, this example motivates MAH for the case of boosts (Friederich, 2016, p. 7).

The problem is that Friederich presents no similar example motivating MAH in the field case, which is, after all, the relevant case for our discussion. And there is a good reason Friederich cannot offer such an example, for there is an important difference between the boost case and the case of local transformations. To see the difference, let's again consider three regions of the universe: S, M and E. The relevant question now is this: if we were to leave the state E as it is and if there was a transformation on S that had empirical significance, would that transformation stop having empirical significance if M was transformed in some specific way? If the answer is *no* (as it was in the case of boosts), then we would have motivated MAH (assuming Friederich's own example motivates MAH). But in the field case, the answer to this question is clearly *yes.* If we transform M by using σ_{se} (a trivial extension of σ_s), then the total transformation is interior, and so $\sigma_s(s) * \sigma_m(m) * e' \sim s * m * e'$. That would mean that the transformation σ_S no longer has empirical significance. Thus, if there was a field case analogous to the boost case, it not clear why would we even hold MAH. In a field setting, which is the setting relevant to Friederich's proof, MAH looks rather implausible.

To summarize, in this section I have shown that, in the light of GW and Teh's discussions, MAH is problematic for at least three reasons: it entails that the empirical significance of a subsystem transformation is not sensitive to (i) transformations in the environment, (ii) its behaviour on the subsystem boundary, and (iii) its asymptotic behaviour. Furthermore, MAH is not motivated for field cases. But even if these are good reasons to reject MAH, perhaps the most direct argument against the soundness of Friederich's proof is that there are physical models for which $s * e$ (or $s * m * e'$) and $\sigma_s(s) * e$ (or $\sigma_s(s) * m * e'$) are both well-defined and represent empirically distinct states of the universe. Given DES-SUL, it follows that σ_s has to be empirically significant even by Friederich's own lights, and so it follows that the proof cannot be sound. One such model is 't Hooft-Polyakov Monopole discussed by Teh (2015) in the appendix of his paper. The other one is 't Hooft's Beam Splitter that we will discuss in greater detail now.

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, p. 95), Greaves and Wallace (2014) and Teh (2015) think that it is, while Hooft (1980, p. 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.

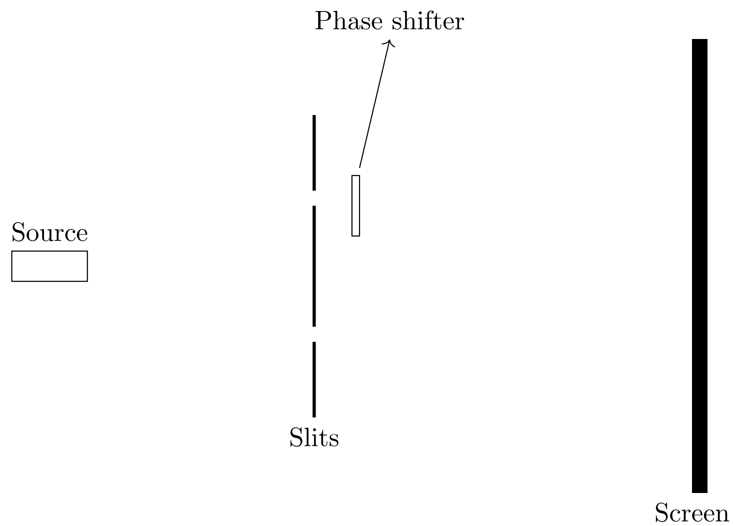

Figure 3

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 a screen, where we can see an interference pattern (see figure 3).

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

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

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:

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

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 on 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.

Given this set-up, in GW's formalism Situation A will be represented as $s*e$ (here e is the state of the lower beam) 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]

[11] To 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.

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 $\sigma_s(s) * e$ 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 $e \neq e'$. The physical situation of the screen is different, and this must be accounted for by a physical difference between states e and e' (Friederich, 2015, p. 553).

Note 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, p. 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 (for example, of an overlap region), Friederich seems to be echoing these remarks by Brading and Brown.

Now, I think that (C1) involves a misscharacterization of Type I scenarios. And in particular, I think that it is not correct to say, as Friederich does, that 'if two situations with different interference pattern are compared, they must evidently be represented by physically distinct environment states'. For in Type I cases, 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 represent the state of measurement devices (such as the screen) or observers. So, in the case at hand, the environment is not supposed to include the screen, which simply serves as a way of revealing the putative physical effects of the action of the subsystem symmetry in question (just as a radar would simply reveal the effects of boosts on a ship).[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, p. 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, the states of screens or observers should not be taken to be part of the

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

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 these away should not affect judgments regarding the empirical significance of 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.[13]

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

[13]In Teh's formalism, this point—that measurement devices are supposed to be ignored in the study of Type I symmetries—is even clearer, for he proposes to treat the environment as a (fixed) asymptotic reference frame with respect to which subsystem symmetries can have empirical significance. In that case, the environment is not modeled as a subsystem with its own dynamics, and so the dynamics of measurement devices living in the environment is deliberately ignored.

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 after going through the beam splitter, 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 same beam (of course, this is only possible by modifying the way we represent the states of the beams in the overlapping region (Friederich, 2015,

p. 554)).[14] 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 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 (for example, the upper 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

[14] Notice that the swapping is not perfect, in one case we need to use a plus sign and not a negative sign.

seems to be suggesting in (C2), that the interference pattern will by itself allows us to determine the state of the upper beam (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 (for example, s) or if it changes (for example, $e^{-iq}(s)$). 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 can be measured. Thus, the interference pattern will 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.

It is helpful to briefly explain why focusing on the overlapping region, as Friederich suggests, does not really help in the current discussion. The reason is simple, namely, a phase-shifter changes the phase of the system it acts on even if there is no overlapping region. Imagine, for example, that the beams are not allowed to interact after the upper beam goes through the phase-shifter (so there is no overlapping region). Even if we cannot measure an interference patter, it is still true that the relative phase between the two isolated subsystems changes as a result of the phase-shifter acting on the upper subsystem. And as I explained earlier, changes in relative phase are empirically significant even if we do not measure the interference corresponding to these changes (just as boosts of a ship with respect to a shore are empirically significant, even if there is no agent or device that detects these changes).

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 $s = 0\,\mathrm{m/s}$, 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 $s = 0\,\mathrm{m/s}$ represents the state of the ship in Situation B in a 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 $0\,\mathrm{m/s}$ 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 (for example, 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, p. 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 electromag-

> netic potential remains continuous, in which case we have a special case of a local gauge transformation on the entire system (Brading and Brown, 2004, p. 656).

I agree with BB that if the transformation is continous 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 (for example, 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 the following reason. 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.[15] For example, it is an open question whether or not gauge transformations actually leave the subsystem invariant in the presence of boundaries (see Gomes (2019a), Gomes (2019b), and Murgueitio Ramírez and Teh (2019) for a careful discussion of this problem, and for some novel solutions).

5 Conclusion

In this paper, I showed that the tension between the arguments of Greaves and Wallace (2014) and Teh (2015) and those of Friederich (2015, 2016) are due to the fact that Friederich's proof appeals to MAH, which is an assumption that is inconsistent with the results presented by both GW and Teh. In particular, I showed that MAH entails—contrary to what GW and Teh show—that the empirical significance of a subsystem transformation is not sensitive to (i) transformations in the environment, (ii) its behaviour on the subsystem boundary, and (iii) its asymptotic behaviour. In the last part of my paper I argued that 't Hooft's Beam Splitter is precisely a case of a local symmetry with empirical significance. In particular, I explained that in that experiment, contrary to Friederich's criticism, (i) the state of the screen should not be considered to be part of the environment, and that (ii) the freedom regarding the representation of the states of the upper and lower beams (when isolated) does not preclude in any way the fact that the symmetry in question leads to empirically distinct scenarios.

[15]Notice that we can try to accommodate this case to Teh's analysis, as presented in section 2. To do so, we would need to treat the upper beam as the subsystem and the lower beam as something like a fixed frame at infinity whose dynamic is omitted in the analysis (clearly, that would not really be a way of treating this case as involving two subsystems interacting at a finite boundary, but would still allow us to conceive of this case as one where a local transformation exhibits empirical significance).

Acknowledgement

Thanks to Nic Teh, Geoff Hall, Ellen Lehet, Ben Middleton, and Jeremy Steeger for their very helpful comments. Also, I want to thank the referees for giving me extremely valuable feedback.

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*.

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. (2019a). 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.

Gomes, H. (2019b). Holism as the significance of gauge symmetries [preprint] url: http://philsci-archive.pitt.edu/id/eprint/196.

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.

Murgueitio Ramírez, S. and Teh, N. (2019). Abandoning Galileo's ship: The quest for non-relational empirical significance [preprint].

Teh, N. J. (2015). Galileo's gauge: Understanding the empirical significance of gauge symmetry. *Philosophy of Science*, 83(1):93–118.