Symmetry, empirical equivalence, and identity

Simon Friederich

email@simonfriederich.eu

Universität Göttingen, Philosophisches Seminar, Humboldtallee 19, D-37073 Göttingen, Germany

Abstract:
The paper proposes a novel approach to the much discussed question of which symmetries have *direct empirical significance* and which do not. The approach is based on a development of a recently proposed framework by Hilary Greaves and David Wallace, who claim that, contrary to the standard folklore among philosophers of physics, *local* symmetries may have direct empirical significance no less than *global* ones. Partly vindicating the standard folklore, a result is derived here from a number of quite plausible assumptions, which states that local symmetries can indeed have no direct empirical significance. Ways to interpret the result are considered and possible morals are outlined.

Contents

1 Introduction

The importance of symmetries in physics has been a recurring topic in the philosophy of science in recent years.[1] According to a (first and imprecise) characterisation, symmetries are (or induce) mappings of a theory's state space onto itself which connect states that are in some sense "physically equivalent". Philosophical debates about symmetries often start from recognising that "physical equivalence" can have (at least) two different meanings in this context. Distinguishing between these two meanings by deciding which one applies in which case is perhaps the main challenge for the philosophical analysis of symmetries.

According to the first meaning of "physical equivalence", symmetries are descriptive redundancies in that any two states related by a symmetry transformation represent *one and the same* physical states of affairs in mathematically distinct ways. According to the second, symmetries operate between physically *distinct* states of affairs, but in such a way that there is no empirically detectable difference between states connected by symmetries for observers who can only make observations *inside* the region where the symmetry transformations operate.[2] How to classify symmetries in actual physical theories in terms of this distinction is a nontrivial task with respect to which there are controversial views. The debate is often formulated in terms of the question of which symmetries have "direct empirical significance" and which do not. Roughly speaking, those symmetries which connect physically identical states of affairs are (or correspond to) those which do *not* have any direct empirical significance, whereas those which operate between physically distinct states of affairs are (or correspond to) those which have some.

The distinction between symmetries which have direct empirical significance and those which do not is often linked to that between *global* and *local* symmetries. The received view—inasmuch as there is one in this debate—is that only global, but not local, symmetries can have direct empirical significance. Global symmetries, roughly speaking, are those which act in a "globally" uniform way on the whole of space-time (typically parametrised by a single real or discrete parameter), whereas local symmetries are defined in terms of real-valued functions $\chi(x)$ from the infinitely many ("local") points x of space-time to the reals. Examples of global symmetries include the Galileo transformations in Newtonian mechanics,

[1]See (Brading and Castellani [2003]) for a useful anthology.

[2]For the purposes of the present paper I will simply *assume* that the question of which state space automorphisms relate states between which there is no empirically detectable difference has been settled in advance. In practice, the task of determining the symmetries of a theory in the sense of singling out these automorphisms can be highly nontrivial. See (Belot [2011]) for a recent study of "symmetry" and "physical equivalence", which argues that we do not have any formal (i.e. purely mathematical) criterion for symmetries that goes well with the idea that symmetries are those state space automorphisms which operate between physically equivalent states.

the Poincaré transformations in special relativity (with the Lorentz transformations as a subgroup), and the constant shifts of the electrostatic potential in Coulomb electrostatics. Examples of symmetries which are counted as local symmetries include the gauge symmetries of the four-potential $A_\mu(x)$ in electrodynamics and its relatives in non-abelian gauge theories (such as those underlying our currently most successful theories of elementary particle physics) as well as the automorphisms of the metric- and mass/energy-configurations induced by coordinate diffeomorphisms in general relativity (referred to simply as "diffeomorphisms" in what follows).

Many philosophers of physics defend what may be called the standard view, namely, that certain global symmetries (though not necessarily all of them) have direct empirical significance, whereas local symmetries generally don't.[3] Hilary Greaves and David Wallace, however, have recently mounted a forceful challenge against this standard view in terms of a formal framework that is meant to allow to distinguish systematically between those symmetries which have direct empirical significance and those which don't, see (Greaves and Wallace [forthcoming]).

Greaves and Wallace mention three closely related aspects of symmetries and their role in theories which they regard as indicating problems for the standard view: first, any theory which has local symmetries must have global ones (i.e. globally constant ones) as a subgroup. If the local symmetries are parametrised in terms of space-time-dependent functions $\chi(x)$, the global symmetries correspond to the spatio-temporally constant functions $\chi(x) = \alpha$ with α constant. Greaves and Wallace note that 'it would be highly mysterious if global symmetries managed to have empirical significance while no other symmetries were around, but somehow lost this capacity once the full local group of transformations appeared as symmetries.' (Greaves and Wallace [forthcoming], p. 2) This observation seems to point to an inadequacy of the view that only global, but not local, symmetries can have direct empirical significance.

The second observation which, according to them, suggests a related problem for the received view is that theories which have a global symmetry are, historically, often replaced by theories where that symmetry is 'localised' (Greaves and Wallace [forthcoming], p. 2) in that the parameter labelling global symmetry transformations in the earlier theory is replaced by a space-time-dependent function in the later theory. For example, the group of constant potential shifts in Coulomb electrostatics is replaced by the group of local gauge transformations of the four potential A_μ in the full theory of electromagnetism. Greaves and Wallace argue

[3] See, for instance, (Kosso [2000]; Redhead [2002]; Brading and Brown [2004]; Lyre [2004]; Healey [2007], Healey [2009]), for such views, and (Maudlin [1998]) for a criticism, to which (Healey [1998]) replies.

that this would seem particularly troubling from the perspective of scientific realism: 'if the explanatory successes of old theories aren't reproduced in their successors, much of the justification for realism is undermined.' (Greaves and Wallace [forthcoming], p. 3)

The third observation made by Greaves and Wallace concerns the explanatory role of symmetries with respect to empirically indistinguishable situations. For example, the phenomenon that—as famously pointed out by Galilei– it is empirically impossible to distinguish from the inside of a ship's cabin whether the ship is at rest or at uniform motion with respect to the sea is explained in Newtonian physics in terms of global space-time symmetries (so-called "boosts"). These operate, in an "active" interpretation, between systems that are at uniform relative motion with respect to each other. The main claim brought forward by Greaves and Wallace with respect to scenarios where a symmetry helps explain an empirical equivalence of such sorts is that one can 'sketch several examples that seem, prima facie at least, to be perfect analogs of the Galileo-ship scenario for cases of local symmetry.' (Greaves and Wallace [forthcoming], p. 3) Other such scenarios include Faraday's cage and a version of the famous two-slit experiment they refer to as "'t Hooft's beam splitter". This latter setup is discussed in detail in Section 5 of this paper. According to Greaves and Wallace, these analogies are accommodated most naturally if one take the local symmetries used in these explanations to have direct empirical significance no less than the global ones.

Inspired by this challenge to the standard philosophers' view as regards symmetries, Greaves and Wallace develop a framework which is meant to permit a more systematic distinction between symmetries which have direct empirical significance and symmetries which don't. According to them, the main lesson from applying this framework to a wide range of examples is that 'nothing like the global/local distinction tells us which symmetries can in general have empirical significance—and in particular, it is false that local symmetries are in general unobservable.' (Greaves and Wallace [forthcoming], p. 3)

The aim of the present paper is 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. In conformity with the strategy proposed by Greaves and Wallace, the present development of their framework focuses on the symmetries of *subsystems* of the universe rather than on symmetries of the universe itself. Unlike their analysis, however, the present investigation applies the relation of representing the same physical state of affairs also to the *subsystem* states themselves, not only to the *universe* states. Using a small number of arguably rather plausible assumptions a result is derived which says that local symmetries—unlike global ones—are indeed without any direct empirical significance, more or less in agreement with the standard folklore and contrary to the claims made by Greaves and Wallace.

The structure of the remaining sections of this paper is as follows: Section 2 introduces the framework proposed by Greaves and Wallace and reviews their core claims. Section 3 elaborates on this framework by specifying the suggested assumptions concerning subsystem symmetries and the relation of representing the same physical state of affairs among subsystem states. Section 4 formulates the result that, given the assumptions specified in Section 3, local symmetries cannot have any direct empirical significance. Section 5 discusses an apparent contradiction between this result and a claim made by Greaves and Wallace in connection with the setup referred to as "'t Hooft's beam splitter". The paper closes in Section 6 with a brief summary and conclusion.

2 Greaves and Wallace on interior vs. non-interior symmetries

According to the leading interpretations of our most successful theories of space-time—the more sophisticated versions of relationalism and substantivalism—any two states of the universe as a whole which are connected by a space-time symmetry represent one and the same physical situation.[4] For example, two universe states u_1 and u_2 in special relativity which are linked by a Lorentz transformation (for example, a rotation or a boost) are regarded as representing one and the same global physical state of affairs, though described in different mathematical terms. If one accepts this perspective yet wants to maintain a nontrivial distinction between symmetries with and without direct empirical significance, one will therefore expect this distinction to hold between symmetries operating between the states of *subsystems* of the universe itself.[5]

The idea to focus on subsystem symmetries to elucidate matters of direct empirical significance is suggested already in (Brown and Sypel [1995]), taken up in (Kosso [2000]; Brading and Brown [2004]; Greaves and Wallace [forthcoming]), and adopted in the present paper as well. It goes well with the fact that in examples such as Galileo's ship the symmetries in question are in fact subsystem symmetries: in the case of Galileo's ship it is only the cabin of the ship, not the

[4]See (Baker [2011]) for detailed considerations in support of this claim.

[5]See (Guilini [1995]) for a defence of the view that symmetry transformations which act differently from the identity transformation at space-time-like infinity in a gauge theory are "*physical*" rather than "*gauge*", which might be taken to contradict the view mentioned here with respect to space-time symmetries. According to Healey, however, inasmuch as Giulini is right, his arguments do not establish that these transformations 'are empirical' (Healey [2007] p. 182) and do therefore not demonstrate that these symmetries have any direct empirical significance.

surrounding ocean, which is represented by a symmetry transformed state in the situation where the ship is moving as compared to where it is at rest.[6]

The framework suggested by Greaves and Wallace is meant to have a wide range of applications, both in quantum and classical contexts and for various different types of subsystems.[7] However, for the purposes of the present investigation it suffices to consider only classical theories and to assume that the subsystems under consideration are finite, compact space-time regions. The examples discussed by Greaves and Wallace can all be treated in that setting, so the question of whether or not the present considerations remain valid in the absence of this restriction need not bother us here.

When one analyses questions concerning the direct empirical significance of symmetries, one would like to be able to ask whether two *mathematically distinct* states represent one and the same *physical* state of affairs. It is therefore crucial to employ a *mathematical* (rather than physical) notion of state. The framework suggested by Greaves and Wallace is based on distinguishing between *subsystem* and *environment* (mathematical) states, where state spaces $\mathcal{S}$ and $\mathcal{E}$ are postulated for the subsystem and the environment, respectively. Elements $u \in \mathcal{U}$ of the universe state space $\mathcal{U}$ are assumed to be uniquely decomposable in terms of subsystem states $s \in \mathcal{S}$ and environment states $e \in \mathcal{E}$. The operation of combining a subsystem with an environment state is denoted by "$*$" (that is, $u = s * e$ is the universe state which arises from combining s and e).

Arbitrary pairs of subsystem and environment states s and e need not in general give rise to a well-defined universe state $u = s * e$. For example, if s and e denote field configurations of a finite space-time region and its environment in a classical field theory, they do not in general coincide on the subsystem boundary, and in case they do, their derivatives may not coincide. In these cases, their composition need not be well-defined (depending on whether higher derivatives are required to exist etc.). However, in those cases where the composition $u = s * e$ of states $s \in \mathcal{S}$ and $e \in \mathcal{E}$ *is* well-defined, it is assumed to be unique.

Greaves and Wallace introduce subsystem and environment symmetries σ_S and σ_E as restrictions of universe symmetries σ to the states spaces $\mathcal{S}$ and $\mathcal{E}$. They require that 'for all $s \in \mathcal{S}$, $e \in \mathcal{E}$, $\sigma(s * e) = \sigma|_S(s) * \sigma|_E(e)$ for some maps $\sigma|_S$ and $\sigma|_E$ [such that t]he symmetries Σ_S of $\mathcal{S}$ and Σ_E of $\mathcal{E}$ are just the sets of all such $\sigma|_S$ and $\sigma|_E$ respectively.' (Greaves and Wallace [forthcoming], p. 9) Using this formal machinery, the essential idea of their analysis is to consider arbitrary universe states $u = s * e$, apply some subsystem symmetry σ_S to the subsystem

[6]See (Healey [2007], Chapter 6) and (Healey [2009]) for a framework which proposes a related classification, but without the focus on subsystem states.

[7]Greaves and Wallace [forthcoming], pp. 8 f.

component s in u, and see whether the resulting universe state $u' = \sigma(s) * e$—if it exists! (see below for the case where it doesn't)—represents the same physical state of affairs as the original u. Their (unsurprising) suggestion is that σ_S does *not* have any direct empirical significance if, for all $s \in \mathcal{S}$ and $e \in \mathcal{E}$ for which $u = s * e$ and $u' = \sigma(s) * e$ are defined, u and u' represent the same physical state of affairs (in which case I shall write $u \sim u'$); and that σ_S does have direct empirical significance if for some $u = s * e$ and $u' = \sigma(s) * e$, these two universe states are physically distinct, i.e. $u \nsim u'$.

This simple idea is already sufficient to distinguish between the two following types of subsystem symmetries $\sigma_S \in \Sigma_S$ in terms of whether they have any direct empirical significance:

First, there are those symmetries σ_S where for all $s \in \mathcal{S}$ and $e \in \mathcal{E}$ such that both $s * e$ and $\sigma_S(s) * e$ are defined the universe states $s * e$ and $\sigma_S(s) * e$ 'represent the same possible world as one another.' (Greaves and Wallace [forthcoming], p. 11) In other words, these are the symmetries for which $s*e \sim \sigma_S(s)*e$ for all $s \in \mathcal{S}$ and $e \in \mathcal{E}$ where both $s * e$ and $\sigma_S(s) * e$ are defined. In line with the considerations just presented, Greaves and Wallace argue that these symmetries do not have any direct empirical significance. Since these symmetries differ from the identity transformation only in the *interior* of the subsystem S (and reduce to the identity transformation on its boundary), they refer to them as "interior symmetries":

> 'Since performing an interior symmetry transformation on the subsystem state and leaving the environment state alone results in a redescription of the same possible world, such a subsystem transformation does not lead to a *distinct* situation, hence no (nontrivial) empirical symmetry is associated with such transformations.' (Greaves and Wallace [forthcoming], pp. 11 f.)

Greaves and Wallace argue that local gauge transformations in gauge theories and space-time diffeomorphisms in general relativity which differ from the identity transformation only in the interior of the subsystem belong to these class of symmetries and hence have no direct empirical significance.

The second type of symmetries σ_S in the classification proposed by Greaves and Wallace are those where, for some $s \in \mathcal{S}$ and $e \in \mathcal{E}$ where $s * e$ and $\sigma_S(s) * e$ are defined, $s * e \nsim \sigma_S(s) * e$. For these symmetries, in other words, universe states $s * e$ and $\sigma_S(s) * e$ represent physically distinct states of affairs and the symmetry σ_S is therefore taken to have direct empirical significance. This is in accordance with the "transformation condition" for symmetries having direct empirical significance suggested by Brading and Brown, which says that 'the transformation of a subsystem of the universe with respect to a reference system must yield an empirically distinguishable scenario' (Brading and Brown [2004] p. 646) for the whole universe.

Greaves and Wallace argue that the symmetry which has direct empirical significance 'in this case will be *purely relational*, in the sense that the intrinsic properties of both subsystem and environment separately are entirely unaffected, and it is only the relations between the two [...] that are changed by the transformation.' (Greaves and Wallace [forthcoming], p. 13, the emphasis is due to Greaves and Wallace.) Since there are subsystem and environment states in Newtonian mechanics and special relativity where the subsystem is effectively isolated from its environment, the space-time transformations such as translations, rotations, and boosts as restricted to proper subsystems of the universe belong to this type. Greaves and Wallace do not stop here, however, but make the further claim that even certain *local* symmetry transformations in gauge theories and diffeomorphisms in general relativity as restricted to subsystems belong to this class and thus have direct empirical significance. This claim will be critically examined in Section 5 when discussing the setup referred to as "'t Hooft's beam splitter".

Symmetries of a third type pose at least prima facie problems for the approach suggested by Greaves and Wallace: namely, those subsystem symmetries where, if σ_S differs from the identity transformation, for all pairs of $s \in \mathcal{S}$ and $e \in \mathcal{E}$, either $s * e$ or $\sigma_S(s) * e$ is undefined. This is indeed the typical case in theories with local symmetries, but it holds also in Coulomb electrostatics.[8] Unfortunately, most symmetries with respect to which it is controversial whether they have direct empirical significance belong to this class. Even though their framework gives no clear-cut verdict in these cases, Greaves and Wallace argue that σ_S has direct empirical significance here nevertheless, since there typically is some suitably chosen environment state $e' \neq e$ for which $s * e \nsim \sigma_S(s) * e'$. Their main motivation for this move seem to be their worries concerning the standard view that only global symmetries can have direct empirical significance, which were outlined in Section 1. So, even though their stipulation that σ_S has direct empirical significance in these cases is well-motivated, it is by no means mandatory. In particular, if one acknowledges that a physical difference exists between $s * e$ and $\sigma_S(s) * e'$, one may always regard it as arising from a physical contrast between e and e', in which case there would be no compelling reason for regarding σ_S as having any direct empirical significance. This move could only be blocked if Greaves and Wallace came up with a well-motivated recipe for obtaining e' from e such that the physical difference between $s * e$ and $\sigma_S(s) * e'$ could not longer be blamed on the choice of e'.

The problem of deciding whether σ_S has any direct empirical significance in these cases arises of course directly from the fact that, for any combination of $s \in \mathcal{S}$

[8] See (Greaves and Wallace [forthcoming], pp. 18-20 and p. 14).

and $e \in \mathcal{E}$, either $s * e$ or $\sigma_S(s) * e$ is undefined. In the following section, I suggest an approach which allows us to overcome this problem. Its core idea is to apply the relation of representing the same physical state of affairs not only to universe states but also to subsystem states. The next section formulates a few simple and arguably very plausible assumptions as regards which subsystem states represent the same physical state of affairs. Based on these assumptions, a result can be *derived*—and need not be stipulated—which allows us to answer the question whether the symmetries for which $s * e$ and $\sigma_S(s) * e$ are not both defined have any direct empirical significance for many cases of interest. As it turns out, the answer is (typically) negative.

3 Elaborating on the Greaves/Wallace framework

Greaves and Wallace presuppose that a sharp distinction exists between pairs of universe states which represent physically distinct states of affairs and pairs of universe states which represent identical physical states of affairs. They do not apply this distinction to pairs of *subsystem* states and would perhaps be unwilling to do so. However, at least prima facie there seems to be no reason for withholding this distinction from *subsystem* states and, therefore, to accept that well-defined subsystem physical states of affairs exist no less than well-defined universe states of affairs.

It may seem, however, as if well-defined identities for subsystem physical states of affairs could not help us much in our investigation of the direct empirical significance of symmetries, simply because, just as for universe states, any two subsystem states which are related by a subsystem symmetry would have to be regarded as representing exactly the same subsystem physical state of affairs. This would make it trivial that *all* subsystem symmetries do not have any direct empirical significance—surely an unwanted result.

All the *intrinsic* physical properties of a subsystem are, by definition, the same in two symmetry-related subsystem states. However, unlike for universe states, for subsystem states one may consider the relations to systems *external* to the subsystem. And indeed, from the perspective of these external systems there may well be a fact of the matter as to which of two symmetry-related states correctly describes the subsystem. (Consider, for example, an observer at the shore, who is looking at Galileo's ship. For her, whether or not the ship is at rest or in motion surely makes a physical difference.) Furthermore, this physical difference, if it exists, may well be regarded as grounded in a physical difference in (what one may call) the *extrinsic* physical properties of the subsystem at issue.[9] If one takes these extrinsic

[9] The notion of an extrinsic property is of course an intricate one. Here its only purpose is to ges-

properties into account as well, the relation of representing the same physical state of affairs may well become non-trivial with respect to subsystem states. In fact, as I shall argue, it can then be used in a natural way to define what it means for a subsystem symmetry to have direct empirical significance.

In what follows, I shall use the symbol "$\sim$" to denote the relation of representing the same physical state of affairs for universe as well as for subsystem states. We can now *define* what it means for a subsystem symmetry to have direct empirical significance, simply by saying that the symmetry has direct empirical significance if and only if it connects at least one pair of subsystem states between which the relation $\sim$ does *not* hold. To sum up (where "DES" stands for "direct empirical significance"):

Assumption (DES)
A subsystem symmetry σ_S has direct empirical significance iff $\sigma_S(s) \nsim s$ for some $s \in \mathcal{S}$.

In other words, according to (DES), σ_S does *not* have any direct empirical significance iff $\sigma_S(s) \sim s$ for all $s \in \mathcal{S}$.

How do facts as regards which mathematically distinct states represent identical *physical* states relate to each other for pairs of universe states on the one hand and pairs of subsystem states on the other? A natural answer to this question, which essentially reproduces the claims made by Greaves and Wallace as regards the first two types of symmetries discussed in the previous section, is as follows: As explained, Greaves and Wallace claim that σ_S does *not* have any direct empirical significance if $s * e \sim \sigma_S(s) * e$ for all s and e where both $s * e$ and $\sigma_S(s) * e$ are defined, and that σ_S does have direct empirical significance if $s * e \nsim \sigma_S(s) * e$ for some s and e where both $s * e$ and $\sigma_S(s) * e$ are defined. Using (DES), we can translate this into the statement that σ_S does not have any direct empirical significance just in case $s * e \sim \sigma_S(s) * e$ for all $s \in \mathcal{S}$ and $e \in \mathcal{E}$ where both $s * e$ and $\sigma_S(s) * e$ are defined. Since this should hold for arbitrary subsystem symmetries σ_S, it seems natural to generalise it as follows (where "SUL" stands for "subsystem-universe link"):

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.

Assumption (SUL) allows us to connect facts about the $\sim$-relation for universe states with facts about it for subsystem states. It is natural to ask whether we can say

ture at the possibility of defining a more ambitious relation of physical equivalence among subsystem states in addition to the trivial one according of which symmetry-related states are by definition equivalent.

anything as to how facts about the $\sim$-relation are related for *different* subsystems. I shall argue that there is a further plausible assumption we can make, which allows us to answer this question:

Consider a subsystem S and decompose it into two non-overlapping sub-subsystems S_1 and S_2. Now consider the state spaces $\mathcal{S}_1$ and $\mathcal{S}_2$ of these sub-subsystems. The question I propose to ask is whether for all pairs of sub-subsystem states $s_1,\ s_1' \in \mathcal{S}_1$ and $s_2,\ s_2' \in \mathcal{S}_2$ such that $s_1 \nsim s_1'$ (and either $s_2 \sim s_2'$ or $s_2 \nsim s_2'$) the inequivalence $s_1 * s_2 \nsim s_1' * s_2'$ must obtain. In other words, the question is whether the complete *physical* state of affairs of the subsystem S determines uniquely the *physical* states of affairs of the sub-subsystems S_1 and S_2.

Arguably, on a reading of "physical states of affairs" that conforms to our intuitive ("pre-theoretic") understanding of what constitutes "physical states of affairs" the question should be answered affirmatively: if s_1' is obtained from s_1 by means of a change in *physical* situation (where "change" is meant metaphorically, not as a process in time), we cannot *in general* "compensate" for this change by altering s_2 to a physically distinct state s_2' with the result that, in the end, the original universe physical state of affairs $s_1 * s_2$ is reproduced in the form $s_1' * s_2'$. The main reason for why it is arguably plausible that $s_1 * s_2$ is physically different from $s_1' * s_2'$ unless $s_1 \sim s_1'$ and $s_2 \sim s_2'$ lies in the fact that, in order for them to be physically identical, replacing $s_1 * s_2$ by $s_1' * s_2'$ must have *no net effect* on the overall physical universe state *for arbitrary environment states e* with which $s_1 * s_2$ and $s_1' * s_2'$ may be combined.

To understand how important the arbitrariness of the environment state e is in this context (and how much it bears on the plausibility of the assumption I am going to make), consider what may seem to be the scenario where the claim just made seems most likely to fail, namely, a pair of equally large spherical sub-subsystems S_1 and S_2 one of which can be seen as the "mirror-image" of the other. It may seem that, with respect to these systems, a "state swapping", i.e. the choice $s_1 = s_2'$ and $s_2 = s_1'$, would lead to identical physical states of affairs $s_1 * s_2$ and $s_1' * s_2' = s_2 * s_1$ for the combined state, contrary to the claim just made. However, even in this highly special case there exist countless environment states e which disturb the envisaged mirror-symmetry between S_1 and S_2 (one of them might represent, for example, a human agent who points with a finger at S_1 and with no finger at S_2). In fact, for all environment states e which do not exhibit the same symmetry as the two spherical systems together, one has $s_1 * s_2 * e \nsim s_2 * s_1 * e$ and therefore $s_1 * s_2 \nsim s_2 * s_1$.

To sum up these considerations, I suggest the following assumption (where "MAH" stands for "modest anti-holism"):

Assumption (MAH)

For all $s_1,\ s_1' \in \mathcal{S}_1$ and $s_2,\ s_2' \in \mathcal{S}_2$,
if $s_1 * s_2 * e \sim s_1' * s_2' * e$ for all $e \in \mathcal{E}$ for which $s_1 * s_2 * e$ and $s_1' * s_2' * e$ are defined, then $s_1 \sim s_1'$ and $s_2 \sim s_2'$.

This assumption is an "anti-holism" inasmuch as it rejects the holistic idea that one and the same physical state of affairs of a subsystem $S = S_1 \cup S_2$ can be "reduced" (or "decomposed") in more than one way to physical states of affairs of its (sub-) subsystems. I have already argued for why it merits the predicate "modest" as well.

To derive the announced result about the (lack of) direct empirical significance of local symmetries, we now need an interesting general feature of theories which have interior symmetries. According to Greaves and Wallace, these include gauge theories such as Maxwell electrodynamics and theories with diffeomorphism invariance such as general relativity. In all these cases, the interior symmetries form a subclass of a larger class of symmetries (consisting of local gauge symmetries or generic diffeomorphisms), which, for the sake of brevity, will be uniformly referred to as the "*local*" symmetries of the theory in what follows. The feature of such theories which I have in mind is the following (with the label "Ext" standing for "extendability"):

Assumption (Ext)
Any local symmetry σ_S defined on the subsystem state space $\mathcal{S}$ can be extended to an interior symmetry defined on the state space $\mathcal{V}$ of a larger subsystem $V \supset S$.

In other words, according to (Ext), for any σ_S on $\mathcal{S}$ there exists a symmetry σ_V on $\mathcal{V}$ which does not have any direct empirical significance and which reduces to σ_S on $\mathcal{S}$.

To see why (Ext) holds for theories with local symmetries (under the assumption, made by Greaves and Wallace, that the local symmetries include interior ones which differ from the identity), consider as an example Maxwell electrodynamics as formulated in terms of the four-vector potential A_μ. It is invariant under local gauge transformations of the form

$$A_\mu(x) \mapsto A_\mu(x) + \partial_\mu \chi(x)\,, \tag{1}$$

where $\chi(x)$ is a (differentiable) real-valued function of the space-time variable x. Now, if S is a finite and compact space-time volume and $\chi(x)$ is a function that is defined *within* S and induces a symmetry transformation σ_S, it is always possible to extend $\chi(x)$ in such a way *outside* S that its value becomes *constant* (say zero) towards the boundary of a larger subsystem V which contains S as a proper subsystem and remains at that constant for the "outward" rest of space-time. Thus, the resulting extended symmetry transformation σ_V (which acts on the state space associated with V) acts differently from the identity only in the *interior* of V, but not on its boundary and outside of it. By the assumptions made by Greaves and Wallace, the symmetry transformation σ_V is interior. Since no special assumptions about $\chi(x)$ *inside* S were made, this establishes (Ext) for arbitrary local subsystem symmetries σ_S of the form (1) (assuming that subsystem symmetries which reduce to the identity at the subsystem boundary are indeed interior symmetries with respect to that subsystem).

Evidently, the assumption (Ext) can be avoided for a theory by denying that it has any interior symmetries (other than perhaps the identity), which do not have any direct empirical significance in the first place. For example, on a non-standard view of general relativity according to which diffeomorphisms which differ from the identity only in the interior of a finite space-time volume (the "hole" of the *hole argument*) may still relate physically *distinct* states of affairs, the assumption (Ext) does not hold. Non-standard views aside, however, the assumption (Ext) seems quite plausible with respect to the theories formulated in terms of local symmetries. In the following section I shall highlight what follows from assuming (Ext) in conjunction with (DES), (SUL), and (MAH).

4 The result

After the conceptual groundwork of the previous sections I can now formulate the following result concerning the (lack of) direct empirical significance of local symmetries:

Proposition Given (DES), (SUL), (MAH), and (Ext), local symmetries do not have any direct empirical significance.

The proposition is easily demonstrated as follows:

Proof. Let σ_S be a local symmetry defined on a subsystem state space $\mathcal{S}$. By (Ext), we can extend it to an interior symmetry σ_V on $\mathcal{V}$, the state space associated with a larger subsystem $V \supset S$. Assume $V = S \cup S'$ with disjoint space-time regions S and S' and let $\mathcal{S}'$ be the state space associated with S'. Now, the fact that σ_V is interior means that for arbitrary states $s \in \mathcal{S}$ and $s' \in \mathcal{S}'$ for which $s * s'$ is defined:

$$\sigma_V(s * s') \sim s * s' \,. \tag{2}$$

Using (SUL), one obtains that for all $e \in \mathcal{E}$ for which $\sigma_V(s * s') * e$ and $s * s' * e$ are defined

$$\sigma_V(s * s') * e \sim s * s' * e \,. \tag{3}$$

Decomposing σ_V in terms of its restrictions σ_S (the symmetry we were originally interested in) and $\sigma_{S'}$ on the state spaces $\mathcal{S}$ and $\mathcal{S}'$, we find that for all $e \in \mathcal{E}$ for which $\sigma_V(s * s') * e$ and $s * s' * e$ are defined:

$$\sigma_S(s) * \sigma_{S'}(s') * e \sim s * s' * e \,. \tag{4}$$

Using (MAH), it follows that

$$\sigma_S(s) \sim s \,. \tag{5}$$

Since s was chosen arbitrary, Eq. (5) holds for all $s \in \mathcal{S}$. By (DES), this means that σ_S doesn't have any direct empirical significance. □

The result that local symmetries cannot have any direct empirical significance is in direct opposition to the claim made by Greaves and Wallace that subsystem local symmetries which do not reduce to the identity on the subsystem boundary *do* have direct empirical significance. Let us briefly discuss the prospects for evading the result by dropping one of the assumptions (DES), (SUL), (MAH), or (Ext)—either by rejecting it or by saying that it does not apply to the theories under consideration.

(DES) may be rejected in either of two ways: either by saying that the relation of "representing the same physical states of affairs" (i.e. the relation "$\sim$") cannot be meaningfully applied to subsystem states; or by accepting that it applies to subsystem states while disputing that it has anything to do with the question of which subsystem symmetries have direct empirical significance. The first option—which rejects the relation of representing the same physical state of affairs for subsystem states—is perhaps the one Greaves and Wallace themselves would choose. I see no reason to doubt its coherence, but it comes at the price of denying that subsystem physical states of affairs have well-defined identities at all.

The second option of rejecting (DES) is to accept the $\sim$-relation between subsystem states and to dispute that it has anything to do with the question of which subsystem symmetries have direct empirical significance. This seems to be the option of choice for those who believe in well-defined subsystem physical states of affairs, while holding that any two symmetry-related subsystem states represent the same physical state of affairs. The drawback of this position, as explained above, is that it ignores the possibility that symmetry-related subsystem states may differ in their *extrinsic* physical properties. Therefore, I would not recommend this option as very attractive.

Rejecting (SUL) while preserving (DES) is a logically viable option, but I see no strong reasons for it. If one accepts the $\sim$-relation as a non-trivial relation between subsystem states, then (SUL), in conjunction with (DES), immediately reproduces the plausible verdicts by Greaves and Wallace on the first two types of symmetries they consider, while being tacit about those of the more controversial third type. Therefore, there seem to be very good reasons for accepting it (unless, of course, one rejects (DES) as well).

Rejecting (MAH) while preserving (DES) is a logically viable option as well, but it commits one to a strong holism. It implies that the physical state of a universe subsystem S does not uniquely determine the physical state of its sub-subsystems $S_1 \subset S$ and $S_2 \subset S$, so that two pairs $s_1, s_1' \in \mathcal{S}_1$ and $s_2, s_2' \in \mathcal{S}_2$ of pairwise physically distinct states (i.e. $s_1 \nsim s_1'$ and $s_2 \nsim s_2'$) can combine to yield one and the same "whole", i.e. one and the same physical state of affairs represented by $s_1 * s_2 \sim s_1' * s_2'$. This is implausible, as argued above, due to the abundance of environment states which allow to distinguish between the two systems S_1 and S_2, which means that one cannot simply reproduce the state $s_1 * s_2$ in the form $s_1' * s_2' = s_2 * s_1$.

As already remarked, whether one accepts the assumption (Ext) for some theory depends on whether one holds that the "local symmetries" of the theory include any interior ones, which do not have any direct empirical significance. Greaves and Wallace claim explicitly with respect to theories such as Maxwell electrodynamics and general relativity that they do have such interior symmetries. This seems to make it difficult for them to avoid the assumption (Ext). In contrast, those who are willing to accept the extreme view that *all* symmetries in these theories have direct empirical significance will see no reason to accept (Ext) in the first place. Such a view seems quite daring, however, as it implies that any distinct configuration of the gauge potentials in a gauge theory and any distinct coordinate configuration in general relativity represents a distinct physical states of affairs. Proponents of this view (and probably also of slightly weakened versions of it) clearly have to address such challenges as the hole argument.

Independent reasons for denying (Ext) in its full generality for some theory may perhaps arise if the underlying space-time has nontrivial topological structure. However, I am unable to address this suggestion at present and recommend it as an interesting topic for future work. Fortunately, none of the scenarios discussed by Greaves and Wallace is set in a space-time with nontrivial topological structure. Arguably, with respect to the scenarios they discuss the four assumptions (DES), (SUL), (MAH), and (Ext) are all quite attractive.

5 Problems with 't Hooft's beam splitter?

In this section I discuss a more specific point of contradiction between the result stated in the previous section and a more specific claim made by Greaves and Wallace. Readers who are willing to accept the result of the previous section without worrying about its correctness may directly turn to the concluding section.

As already mentioned, for most of the local subsystem symmetries σ_S which, according to Greaves and Wallace, have direct empirical significance there are no pairs of states $s \in \mathcal{S}$ and $e \in \mathcal{E}$ for which both $s * e$ and $\sigma_S(s) * e$ are defined. According to them, in these cases "in order to realise [the direct empirical significance of σ_S] the environment state must be altered" (Greaves and Wallace [forthcoming], p. 14), that is, one must change the environment state e to another state e' with the result that an inequivalence $s * e \nsim \sigma_S(s) * e'$ holds. As remarked, with respect to these cases Greaves and Wallace essentially *stipulate* that this inequivalence $s * e \nsim \sigma_S(s) * e'$ indicates the direct empirical significance of σ_S. They do not *derive* the result that σ_S has direct empirical significance from any of the assumptions of their framework. Therefore, there is no consistency problem between what they directly *establish* and the result stated in the previous section, which says that, given (DES), (SUL), (MAH), and (Ext), σ_S does *not* have any direct empirical significance.

However, with respect to a more narrow class of local symmetries Greaves and Wallace claim that there *are* subsystem and environment states $s \in \mathcal{S}$ and $e \in \mathcal{E}$ such that for a subsystem symmetry σ_S both $s * e$ and $\sigma_S(s) * e$ are defined and represent physically *distinct* states of affairs. According to them, in other words, the inequivalence $s * e \nsim \sigma_S(s) * e$ holds for these local symmetries. Let us assume that this is true. Then, for appropriate states $s \in \mathcal{S}$ and $e \in \mathcal{E}$, it follows from (SUL) that $s \nsim \sigma_S(s)$ and further, from (DES), that σ_S has direct empirical significance, contradicting the result presented in the previous section and establishing an inconsistency between the assumptions involved.

An example of a theory which, according to Greaves and Wallace, has local symmetries σ_S with the property that $s * e \nsim \sigma_S(s) * e$ is what they call "Klein-Gordon-Maxwell electrodynamics". In this theory, the four-vector potential A_μ as familiar from Maxwell electrodynamics is coupled to a complex scalar field ψ. The Lagrangian of the theory is given by

$$\mathcal{L} = (\partial_\mu\psi - iqA_\mu\psi)^*(\partial^\mu\psi - iqA^\mu\psi) - m^2\psi^*\psi + \mathcal{L}_{\mathrm{EM}} \,. \tag{6}$$

Here $\mathcal{L}_{\mathrm{EM}}$ is the Lagrangian of Maxwell electrodynamics without matter. The Lagrangian $\mathcal{L}$ itself is invariant under local gauge transformations of the form

$$\psi(x) \mapsto \exp\left(iq\chi(x)\right)\psi(x)\,, \tag{7}$$

$$A_\mu(x) \mapsto A_\mu(x) + \partial_\mu\chi(x)\,. \tag{8}$$

The argument brought forward by Greaves and Wallace for the claim that some of these local symmetry transformations (conceived as subsystem symmetries) have direct empirical significance is based on a setup that realises a version of the celebrated two-slit interference experiment. In this version of the two-slit setup, often referred to as "'t Hooft's beam splitter"[10], a phase-shifter is inserted behind the upper, but not the behind the lower, of the two slits. The essential empirical observation is that the interference pattern that emerges on a screen behind the slits 'depends nontrivially on whether or not the phase-shifter is present.' (Greaves and Wallace [forthcoming], p. 6) Greaves and Wallace interpret this fact as demonstrating that subsystem local gauge transformations have direct empirical significance in this case.

The crucial move in their argument is to treat the upper half-beam as a subsystem in the sense of their framework and to argue that 'the phase-shifter precisely implements a transformation $\psi \mapsto e^{i\theta}\psi$ on the upper half-beam while leaving the electromagnetic potential unchanged (as gauge transformations with constant χ do).' (Greaves and Wallace [forthcoming], p. 7) Their suggestion is that, if the situation where the phase-shifter is absent corresponds to some state $s*e$, the situation where the phase-shifter is present corresponds to some state $\sigma_S(s) * e$, where σ_S is some transformation $\psi \mapsto e^{i\theta}\psi$ with constant θ. As they point out, it is possible to let both $s * e$ and $\sigma_S(s) * e$ be well-defined in this case by choosing ψ such that it vanishes in a neighbourhood of the boundary which separates the subsystem and the environment. Since the physical state of affairs depends on whether or not the phase-shifter is present, they conclude that σ_S has direct empirical significance in this case.

However, their assumption that inserting the phase-shifter "implements a transformation $\psi \mapsto e^{i\theta}\psi$ on the upper half-beam" without having an impact on the physical state of the *environment* is untenable. If the upper half-beam plays the role of the subsystem S and the rest of the setup 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 $e \nsim e'$. The physical situation of the screen is different, and this must be accounted for by a physical difference between these states e and e'. Insisting that the state of the *lower* half-beam, excluding the screen, is unchanged by inserting a phase shifter in the upper half-beam does not help, since the physical difference between the two situations of the screen must still be accounted for by the environment state. So, the situation where the phase-shifter is present and the one where it is absent are *not* represented by states $s * e \nsim \sigma_S(s) * e$ with a fixed environment state e.[11]

[10] The same setup is also discussed in (Brading and Brown [2004] and Lyre [2004]), who maintain that it does *not* provide an example of a situation where a local symmetry has direct empirical significance.

[11] Essentially the same point is made by Brading and Brown, who emphasise that '*an interference pattern occurs only where ψ_I and ψ_{II} overlap, and clearly these conditions [i.e. a non-boundary*

In response to this argument, Greaves and Wallace may claim that, in order to realise the physical difference between ψ and $e^{i\theta}\psi$, the screen behind the setup need not be *actually* present, but only *hypothetically*.[12] The idea in this case would be that the difference between the two interference patterns on the screen is a mere *reflection* of a preceding physical difference between ψ and $e^{i\theta}\psi$ which had been there all along in the absence of the screen. The purpose of invoking the hypothetical screen, this argument goes, is to make the physical difference between ψ and $e^{i\theta}\psi$ *vivid* and *transparent*, not to include the screen as an actual ingredient of the environment system.

In more formal terms, this argument takes the following form: Consider, first, a situation where the two beams have vanishing overlap and can be described by the field configurations $(\psi_I, A_{\mu,I})$ and $(\psi_{II}, A_{\mu,II})$, respectively. Consider, second, a situation where they are described by $(e^{i\theta}\psi_I, A_{\mu,I})$ and $(\psi_{II}, A_{\mu,II})$ with $\theta \neq 0,\ 2\pi$. The question we would like to answer is whether the two situations are physically distinct or identical. However, as long as there is no overlap between the two beams (i.e. no region where both ψ_I and ψ_{II} are non-zero), there is no interference pattern and therefore no *manifest* physical difference between the two situations. The suggested idea is to consider a modification of both situations by introducing a *hypothetical* region of overlap ("the screen"), to arrive at two *manifestly* physically different situations, and to conclude from that difference that a physical difference between $(\psi_I, A_{\mu,I})$ and $(\psi_{II}, A_{\mu,II})$ had been there all along.

Now let us try to write down the field configurations for two physically different situations in the (hypothetical) case where a region of overlap is present such that the physical difference between them appears as a consequence of the physical difference between ψ_I and $e^{i\theta}\psi_I$.

The first situation, where the phase-shifter is absent, —call it "situation A"— can be taken to be one where the upper half-beam is described by $(\psi_I, A_{\mu,I})$, the lower by $(\psi_{II}, A_{\mu,II})$, and the overlap region by a superposition of the two:

$$\Big(\psi_{I,A}(x) + \psi_{II,A}(x),\ A_{\mu,A}\Big) \tag{9}$$

with $\psi_{I,A}$ and $\psi_{II,A}$ in the overlap region chosen as smooth continuations of ψ_I and ψ_{II} in the upper and lower half-beams.

If a phase-shifter is inserted in the upper half-beam, however, a physically different situation arises—call it "situation B"—, where the upper half-beam can be described by $(e^{i\theta}\psi_I, A_{\mu,I})$, the lower (unchanged) by $(\psi_{II}, A_{\mu,II})$, and the overlap

preserving local gauge transformation which acts only on the state of the upper half-beam] on [sic] cannot be met in such a region.' (Brading and Brown [2004] p. 653. The emphasis is due to Brading and Brown.)

[12] I would like to thank two anonymous referees for suggesting this possible move.

region by the superposition

$$\left(e^{i\theta}\psi_{I,A}(x) + \psi_{II,A}(x),\ A_{\mu,B}\right) \tag{10}$$

with the same $\psi_{I,A}$ and $\psi_{II,A}$ as before.

For non-trivial values of the constant θ the two field configurations (9) and (10) are indeed physically different—at least if the plausible assumption is made that $|\psi|^2 = \psi^*\psi$ corresponds to a physical quantity such as the matter or charge density. The suggested argument now appeals to the physical difference between these two situations A and B, which appears *in the presence of an overlap region*, and sees it as reflecting a physical difference between the two situations described by $(\psi_I, A_{\mu,I})$ and $(e^{i\theta}\psi_I, A_{\mu,I})$ in the upper half-beam—both with *and without* the region of overlap. The idea here is that the physical difference between the situations A and B in the presence of an overlap region arises precisely as a dynamical consequence of the physical difference between $(\psi_I, A_{\mu,I})$ and $(e^{i\theta}\psi_I, A_{\mu,I})$ in the upper half-beam and thereby testifies to the fact that this difference has been there all along.

Contrary to this suggestion, however, the physical difference between the situations A and B should *not* be taken to reflect a physical difference between the upper half-beam configurations $(\psi_I, A_{\mu,I})$ and $(e^{i\theta}\psi_I, A_{\mu,I})$. To see this, it suffices to note that the field configuration $(e^{i\theta}\psi_I, A_{\mu,I})$ can also describe the situation A, and the field configuration $(\psi_I, A_{\mu,I})$ can also describe the situation B in the upper half-beam, while the field configuration of the lower half-beam remains $(\psi_{II}, A_{\mu,II})$ in both situations.

To be specific, the situation A can alternatively be described by the field configuration $(e^{i\theta}\psi_I, A_{\mu,I})$ in the upper half-beam, $(\psi_{II}, A_{\mu,II})$ in the lower half-beam, and by the gauge-transformed

$$\left(e^{i\theta\lambda(x)}\left(\psi_{I,A}(x) + \psi_{II,A}(x)\right),\ A_{\mu,A} + \frac{\theta}{q}\,\partial_\mu\lambda(x)\right) \tag{11}$$

in the region of overlap, where $\lambda(x)$ is a smooth real-valued function with $\lambda(x) = 1$ at the boundary between the upper half-beam and the region of overlap and $\lambda(x) = 0$ at the boundary between the lower half-beam and the region of overlap.

Similarly, situation B can alternatively be described by the field configuration $(\psi_I, A_{\mu,I})$ in the upper half-beam, $(\psi_{II}, A_{\mu,II})$ in the lower half-beam, and by the gauge-transformed

$$\left(e^{-i\theta\lambda(x)}\left(e^{i\theta}\psi_{I,A}(x) + \psi_{II,A}(x)\right),\ A_{\mu,B} - \frac{\theta}{q}\,\partial_\mu\lambda(x)\right) \tag{12}$$

in the region of overlap. The point of writing down these alternative field configurations is of course that they describe exactly the same situations A and B as before, but now with reversed ascriptions of ψ_I and $e^{i\theta}\psi_I$ to the upper half-beam.

The upshot of these considerations is that one should not take the difference between the two interference patterns in the situations A and B to reflect a physical difference between ψ_I and $e^{i\theta}\psi_I$. In fact, *both* situations A and B can equally well be described in terms of either ψ_I or $e^{i\theta}\psi_I$ for the upper half-beam. From the point of view based on (DES), (SUL), (MAH), and (Ext), developed here, this makes perfect sense: Both ψ_I and $e^{i\theta}\psi_I$ describe the same physical state of affairs, and the acknowledged physical difference between the situations A and B in the overlap region concerns *merely* the environment state, not the state of the upper half-beam.

In other words, while the interference pattern dynamically reflects whether or not the *phase shifter* is present, it does not dynamically reflect whether the state of the upper-half beam is ψ_I or $e^{i\theta}\psi_I$. The phase shifter and the screen are both features of the environment (excluding both beams), and the dynamical correlation between them, according to the present analysis, is a feature which concerns *only* the environment in that the physical states of the beams are the same in both situations. Whether or not this means that the physical situation on the screen depends *non-locally* on the presence of the phase shifter is a complicated question that I wish to leave aside in the context of the present investigation, for its (by no means obvious) answer requires a detailed investigation of the specific dynamical properties of the theory under considerations.[13] To conclude, the hypothetical presence of a region where the half-beams overlap does not give us any reason to regard ψ_I and $e^{i\theta}\psi_I$ as representing different subsystem physical states of affairs.

Finally, it should be noted that the present argument for the physical identity of ψ_I and $e^{i\theta}\psi_I$ applies only in virtue of the fact that the symmetry (7) is a *local* rather than a *global* one. To see this, consider, as an example, a version of Galileo's ship where a screen, located at the shore, registers different interference pattern of the water waves, depending on whether the ship is at rest or in constant relative motion with respect to the shore. Do the different interference patterns on the screen correspond one-to-one to physically different, symmetry related states s and (its boosted analogon) $\sigma_S(s)$ of the ship in this setup? The answer depends crucially of whether we assess the problem in the framework of a theory which contains only global symmetries (such as special relativity) or also local ones (such as general relativity). If the theory has only global symmetries, once some coordinate system has been chosen (with the shore at rest, say), the question of whether the ship is in s or in $\sigma_S(s)$ has a determinate answer; if it has also local symmetries, both s and $\sigma_S(s)$ can be chosen to describe the ship both at rest and in uniform motion, and the physical difference between the two situations, as it manifests itself in the two

[13]See (Vaidman [2012]) for an illuminating discussion of the conceptually related Aharonov-Bohm effect, which lends some plausibility to the idea that a non-local influence need *not* be assumed.

different interference patterns on the screen, is accounted for entirely by a physical difference in the states used to describe the (complete) environment of the ship. This goes well with the conclusion of the above discussion of 't Hooft's beam splitter, where the symmetry (7) is also local. So, claiming that the local symmetries in the case of 't Hooft's beam splitter lack any direct empirical significance does not commit one to the view that global symmetries do not have any direct empirical significance either.

6 Summary and conclusion

The present paper has proposed a development of a recently suggested framework by Hilary Greaves and David Wallace to distinguish between symmetries which have direct empirical significance and symmetries which do not. A result has been presented, based on four rather plausible assumptions, according to which subsystem local symmetries do not have any direct empirical significance, contrary to the claims made by Greaves and Wallace.

The assumptions, to recapitulate, are the following: first, that the question of direct empirical significance for subsystem symmetries is the same as that of which symmetries connect physically identical subsystem states of affairs (assumption (DES)); second, that two subsystem states s and s' correspond to the same physical state of affairs just in case their combinations with arbitrary environment states e (if defined) yield identical universe states of affairs (assumption (SUL)); third, that the combined physical state of affairs of two subsystems determines uniquely the physical states of the individual subsystems (assumption (MAH)); and, fourth, that local symmetries defined on subsystem state spaces can always be extended to interior symmetries (i.e. symmetries without any direct empirical significance) on the state spaces of larger subsystems (assumption (Ext)). None of the assumptions is indisputable—in particular, the assumption (Ext) may not be adequate for some theories with local symmetries—, but all four can be motivated very well for the scenarios considered by Greaves and Wallace.

An implication of the Greaves/Wallace framework together with (DES), (SUL), (MAH), and (Ext) is that subsystem symmetries in theories with only *global* symmetries may have direct empirical significance, whereas their ("globally constant") counterparts in theories which have *local* symmetries do not. For example, from the perspective of an interpretation which accepts the assumption (Ext) for diffeomorphisms in general relativity, subsystem-wise globally constant diffeomorphisms do not have any direct empirical significance, whereas boosts as employed in the Newtonian explanation of Galileo's ship may. According to Greaves and Wallace, such a perspective is odd, since 'it would be highly mysterious if global symmetries managed to have empirical significance while no other symmetries were around, but somehow lost this capacity once the full local group of transformations appeared as symmetries.' (Greaves and Wallace [forthcoming], p. 2) In response to these qualms, those who defend the assumptions (DES), (SUL), (MAH), and (Ext) may respond as follows:

It is true that, from the perspective of an account which accepts the Greaves/-Wallace framework together with (DES), (SUL), (MAH), and (Ext), it appears that subsystem global symmetries "lose" their direct empirical significance once the *global* symmetry group of the earlier theory is replaced by the *local* symmetry group of the later one. However, there is nothing odd or "mysterious" about this formal "loss", for all that it means is that what appear to be two physically *distinct* yet empirically *indistinguishable* subsystem situations in the earlier theory turns out to be one *single* physical subsystem situation in the later theory (which, however, can be related in physically distinct ways to its environment). As a matter of fact, it is difficult to imagine a more elegant explanation of the empirical equivalence![14] So, far from being worrisome for the scientific realist (as suggested by Greaves and Wallace, see Greaves and Wallace [forthcoming], p. 3.), the *loss* of direct empirical significance due to the "localisation" of global symmetries in the switch from one theory to another can in fact be seen as a *gain* in our explanatory resources.

Acknowledgements

I would like to thank Holger Lyre for fueling my interest in symmetries and for alerting me of the work of Graves and Wallace. I am particularly grateful to Nazim Bouatta for many fruitful meetings dedicated to the understanding of symmetries and the Greaves/Wallace framework. Furthermore, I would like to thank two anonymous referees of the *British Journal for the Philosophy of Science* for

[14] In (Healey [2009], Section 5), Richard Healey presents an account of how empirical equivalences between physical situations ("empirical symmetries") are sometimes accounted for as "strong" in an earlier theory and as "perfect" in a later one. For example, as Healey points out and as famously noted by Einstein, pre-relativistic Maxwell electrodynamics gives conceptually different accounts of electrodynamic processes, depending on which systems are regarded as rest and which are regarded as moving, while the predictions it makes are empirically equivalent. Special relativity accounts for this empirical equivalence by revealing that, as Healey puts it, 'distinct but symmetrically related situations are in fact duplicates of one another.' (Healey [2009] p. 709) The present account, based on the Greaves/Wallace framework together with (DES), (SUL), (MAH), and (Ext), permits to see the replacement of the global symmetries of the earlier theory by the local symmetries of the later one in an analogous light: what appeared to be physically distinct, yet empirically equivalent, subsystem situations from the point of view of the earlier theory is revealed to be one single subsystem situation from the point of view of the later one.

their very useful comments.

References

Baker D. J., [2011]: ‘Broken Symmetry and Spacetime’, *Philosophy of Science*, **78**, pp. 128-48.

Belot, G. [2011]: ‘Symmetry and Equivalence’, in *Philosophy of Science Assoc. 22nd Biennial Mtg. (Montréal, QC) > PSA 2010 Contributed Papers*.

Brading, K. and Brown, H. R. [2004]: ‘Are Gauge Symmetry Transformations Observable?’, *British Journal for the Philosophy of Science*, **55**, pp. 645-65.

K. Brading and E. Castellani (eds) [2003]: *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press.

Brown, H. and Sypel, R. [1995]: ‘On the Meaning of the Relativity Principle and other Symmetries’, *International Studies in the Philosophy of Science*, **9**, pp. 23553.

Giulini, D. [1995]: ‘Asymptotic Symmetry Groups of Long-ranged Gauge Configurations’, *Modern Physics Letters*, **10**, pp. 2059-70.

Greaves, H. and Wallace, D. [forthcoming]: ‘Empiricial Consequences of Symmetries’, <philsci-archive.pitt.edu/8906/>.

Healey, R. [1998]: ‘Quantum Analogies: A Reply to Maudlin’, *Philosophy of Science*, **66**, pp. 440-7 (1998).

Healey, R. [2007]: *Gauging What’s Real: The Conceptual Foundations of Contemporary Gauge Theories*, New York: Oxford University Press.

Healey, R. [2009]: ‘Perfect Symmetries’, *British Journal for the Philosophy of Science*, **60**, pp. 697-720.

Lyre, H. [2004]: ‘Holism and Structuralism in $U(1)$ Gauge Theory’, *Studies in History and Philosophy of Modern Physics*, **35**, pp. 643-70.

Kosso, P. [2000]: ‘The Empirical Status of Symmetries in Physics’, *British Journal for the Philosophy of Science*, **51**, pp. 81-98.

Maudlin, T. [1998]: ‘Healey on the Aharonov-Bohm Effect’, *Philosophy of Science*, **65**, pp. 361-8.

Redhead, M. [2002]: 'The Interpretation of Gauge Symmetry', in M. Kuhlmann, H. Lyre and A. Wayne (eds), 2002, *Ontological Aspects of Quantum Field Theory*, Singapore: World Scientific.

Vaidman, L. [2012]: 'Role of Potentials in the Aharonov-Bohm Effect', *Physical Review A*, **86**, pp. 040101(1-4)(R).