

Are gauge symmetry transformations observable?

Katherine Brading[†] and Harvey R. Brown[‡]

ABSTRACT

In a recent paper in the *British Journal for the Philosophy of Science*, Kosso discussed the observational status of continuous symmetries of physics. While we are in broad agreement with his approach, we disagree with his analysis. In the discussion of the status of gauge symmetry, a set of examples offered by 't Hooft has influenced several philosophers, including Kosso; in all cases the interpretation of the examples is mistaken. In this paper we present our preferred approach to the empirical significance of symmetries, re-analysing the cases of gauge symmetry and general covariance.

1 Direct and Indirect Empirical Significance

The notion of symmetry that we are concerned with is defined with respect to the laws of motion. Given the laws, specified in terms of dependent and independent variables, a symmetry transformation is a transformation of these variables that preserves the explicit form of the laws. The issue we are interested in is the empirical status of such symmetry transformations. Galileo's famous ship experiment (Galileo [1967], pp. 186-8) provides an example of where (to an appropriate approximation) a symmetry transformation is both physically implementable and directly observable. The transformation is implemented via two empirically distinct scenarios of the ship at rest and in uniform motion with respect to the shore, and the symmetry is observed by noticing that, *relative to the cabin of the ship*, the phenomena inside the cabin do not enable us to distinguish between the two scenarios. Following Brown and Sypel ([1995]),¹ we maintain that the *direct empirical significance* of physical symmetries rests on the possibility of effectively isolated subsystems that may be actively transformed with respect to the rest of the universe.² This active transformation need not be physically implementable in practice (try boosting a planet, for example); the point is that we compare two empirically distinct possible scenarios at least theoretically, one containing the untransformed subsystem and one the transformed subsystem.

The example of Galileo's ship also illustrates that observing a symmetry involves *two observations*, as has been discussed by Kosso ([2000]). He writes (p. 85):

[†] Wolfson College, Oxford, OX2 6LE. email: katherine.bradings@wolfson.ox.ac.uk.

[‡] Faculty of Philosophy, 10 Merton Street, Oxford, OX1 4JJ. email: harvey.brown@philosophy.ox.ac.uk.

¹ See also Budden ([1997]).

² Notice that, since our objective is empirical significance, this goes beyond a purely mathematical active symmetry transformation. We discuss the case symmetries of the universe as a whole in the final paragraph of this section.

As long as one can claim to be able to observe that the transformation prescribed by a particular symmetry has taken place, and that the associated invariance held, then one can claim to be able to directly observe the physical symmetry in nature.

And he goes on (p. 87):

To observe the transformation is to observe both the unchanged reference and the changed system.

In other words, we first observe the transformation, which involves transforming a subsystem with respect to some reference that is itself observable, and we then observe that the symmetry holds for the subsystem (p. 86):

observation of a symmetry will always require two components: One must observe that the specified transformation has taken place, and one must observe that the specified invariant property is in fact the same, before and after.

In broad agreement with Kosso, we require that two conditions are met in order for a symmetry to have *direct empirical significance*:

1. *Transformation Condition*: the transformation of a subsystem of the universe with respect to a reference system must yield an empirically distinguishable scenario; and
2. *Symmetry Condition*: the internal evolution of the untransformed and transformed subsystems must be empirically indistinguishable.

Not everyone agrees with this account of the empirical significance of symmetry. For example, Morrison ([1995], p. 159) writes that ‘symmetries ... should not be viewed straightforwardly as concrete claims about the world’, and then goes on to claim that ‘Conservation laws provide the empirical component or manifestation of symmetries’.

Similarly, Sklar ([1996], p. 78) endorses the view that

the invocation of charges is simply another way of stating what the various symmetries are. There is a redundancy of the ‘theoretical structure’ on the ‘observational data’ to be explained.

However, even in cases where a given conservation law is connected to a certain symmetry (via Noether’s first theorem, say, of which more below), the conservation law does not exhaust the empirical manifestation of the symmetry.³ Invariance of the dynamical laws under spatial translation, for example, is directly manifested by the insensitivity of the dynamical evolution of systems to their location. Similarly, the invariance of the dynamical laws under boosts⁴ is directly manifested by Galileo’s ship experiment. A

³ See also Wigner’s discussions of the significance of symmetries, Wigner ([1967]).

⁴ In modern physics the Relativity Principle holds between inertial states of motion, related to one another by ‘boosts’ (the form of the transformation depending on whether we are in Newtonian or special relativistic physics).

symmetry of the laws manifests itself directly whenever an active interpretation of an effectively isolated subsystem can be operationally instantiated.⁵

Since symmetries are defined with respect to the laws they have not only direct empirical significance, but also *indirect empirical significance*, arising from properties of the laws that are connected with symmetries and that themselves have direct empirical significance. In similar vein, Kosso writes ([2000], p. 85):

Indirect evidence of the symmetry, by contrast, amounts to the observation of some consequence of the symmetry, but not of the transformation and invariance themselves.

For example, it is proven by Noether's theorem that every continuous, global symmetry is associated with a conservation law.⁶

This indirect empirical significance is independent of whether any active transformation of effectively isolated subsystems can be defined, and consists of claims of various different types, both actual and counterfactual; for example, if the laws didn't have this symmetry then we would be in a different universe, various empirical consequences of the symmetry wouldn't hold, and so forth.⁷ Notice also the bearing this has on the issue of the meaning of symmetries for the universe as a whole: while symmetries of the universe as a whole have no direct empirical significance, they do have indirect empirical significance via the laws of physics if these are assumed to hold for any isolated system.

2 Global and local continuous symmetries

Continuous symmetries come in different varieties, one important distinction being between external and internal symmetries. External symmetries are symmetries of space and time: the symmetry transformations

⁵ The requirements for observing a symmetry in any particular case are, of course, subtle. Crucially, we must be able to prepare an 'effectively isolated subsystem' – i.e. a subsystem of the universe whose interactions with the rest of the universe are either negligible or irrelevant with respect to the evolution of the untransformed and transformed systems and the symmetry in question. This effectively isolated subsystem must then satisfy two further conditions. First, in order to observe that the untransformed and transformed systems are empirically distinct, we must interact with them. In the case of Galileo's ship, we can simply see that the ship is at rest (or in motion) with respect to the shore without disturbing the effective isolation of the ship – we are interacting with the ship via the electromagnetic field, but the phenomena inside the cabin still display the symmetry. More generally, we are required simply to make observations that distinguish empirically between the untransformed and transformed system, and this doesn't have to take place at the same time as the observations of empirical equivalence between the evolution of the two systems. For example, we could observe the phenomena inside the cabin of Galileo's ship when effectively isolated, and *then* interact the ship with another system in order to show that the state of motion of the untransformed ship differs from that of the transformed ship, where that interaction need not preserve the effective isolation of the ship. For example, the ship in motion may crash into rocks that the ship at rest stays a constant distance from. Second, we must be able to physically implement the active transformation. This does not mean that we have to take a single system and transform it, although this might be one way of implementing the transformation. Rather, it means being able to set up two systems whose initial conditions differ only by the symmetry transformation. If we can do all these things, then we can 'directly get our hands on' the associated symmetry.

⁶ This statement is subject to certain restrictions; see section 5, below. Note also that Noether proved the converse of her theorem.

⁷ Whenever a symmetry is 'observationally complete' (Kosso [2000], p. 88) it can have only indirect, and not direct, empirical significance.

are applied to the spatio-temporal degrees of freedom, and give rise to translations and rotations, for example. Internal symmetry transformations came into physics with the advent of quantum theory, in which additional (non-spatio-temporal) degrees of freedom are attributed to the system. Further details of this distinction can be found in Kosso's discussion ([2000], p. 84), for example.⁸ Kosso ([2000], p. 84-5) also distinguishes between global and local symmetries, but here he is not quite careful enough, and since this may be one source of problems later in his paper, we will take care to explain the distinction clearly here.

The terms 'global' and 'local' are used in philosophy of physics with a variety of meanings. One contrast is between global in the sense of applying to the universe as a whole versus local in the sense of applying to some region of the universe. Other examples come from debates concerning locality (and non-locality) in quantum theory, and discussions of the concept of locality pertinent to relativity theory. In the context of continuous symmetries, and in particular gauge symmetries, the contrast between global and local is none of these. Rather, the terms are used to mark the distinction between symmetries that depend on constant parameters (global symmetries) and symmetries that depend on arbitrary smooth functions of space and time (local symmetries). Examples of global symmetries are the familiar Galilean spacetime symmetries of spatial and temporal translations, spatial rotations, and boosts, along with internal global symmetries such as the global phase invariance of the Schrödinger equation for a free particle. The gauge symmetry of electromagnetism (an internal symmetry) and the diffeomorphism invariance in General Relativity (an external symmetry) are examples of local symmetries, since they are parameterized by arbitrary functions of space and time.

The sense of 'local' here is that intended by Weyl in his 1918 theory (Weyl [1918]) where he introduced the idea of local scale transformations. Weyl built a theory in which the lengths of two vectors are directly comparable only when the vectors are in the same place. If there is no direct way to compare the lengths of two spatiotemporally separated vectors, then fixing the length scale at one spacetime location does not fix the length scale at any other spacetime location. In contrast, a global length scale is a length scale which, once fixed at one spacetime point, is fixed everywhere. The same local/global distinction is intended by Yang and Mills ([1954], p. 192, our emphasis):

The conservation of isotopic spin is identical with the requirement of invariance of all interactions under isotopic spin rotation. This means that when electromagnetic interactions can be neglected, as we shall hereafter assume to be the case, the orientation of the isotopic spin is of no physical significance. The differentiation between a neutron and a proton is then a purely arbitrary process. *As usually conceived, however, this arbitrariness is subject to the following limitation: once one chooses what to call a proton, and what a neutron, at one space-time point, one is then not free to make any choices at other space-time points.*

⁸ For a review of the history of, and philosophical problems associated with, continuous symmetries, see Martin ([2003]). See also Castellani ([2002]).

It seems that this is not consistent with the localized field concept that underlies the usual physical theories. *In the present paper we wish to explore the possibility of requiring all interactions to be invariant under independent rotations of the isotopic spin space at all space-time points, so that the relative orientation of the isotopic spin at two space-time points becomes a physically meaningless quantity* (the electromagnetic field being neglected).

In other words, Yang and Mills intend to implement a local freedom in the specification of proton versus neutron, analogous to Weyl's implementation of a local freedom in the length scale; fixing what isotopic spin state constitutes a proton at one location does not determine all the proton-states everywhere else. Of course, once I have fixed the proton state at a given spacetime location relative to one nucleon, it is fixed for all nucleons at that spacetime location.

This global/local distinction is between 'the same at every point' and 'freely varying from point to point'. In terms of transformations, a global scale transformation changes the length scale at every point by the same amount, whereas for a local scale transformation the change in the length scale varies smoothly but otherwise arbitrarily from point to point. This is the sense of global and local that is at work in the terminology 'global gauge transformation' and 'local gauge transformation'.

The Weyl and Yang-Mills distinction clearly differs from that of applying a transformation to the universe as a whole versus applying it to a part of the universe (or, more generally, to an entire system versus a part of that system), and here we reserve the terms 'global' and 'local' for the former distinction. The crucial difference between these two distinctions can be seen by considering a system consisting of two sub-systems. A Weyl-type local scale transformation applied to the system allows for one subsystem to be freely rescaled relative to the other *only when the two subsystems are spatially separated*; it does not allow us to vary the scale of one subsystem with respect to the scale of the other when they are at the same location. How could it? Relative length at the same spacetime location is directly observable, this was where the Weyl argument began. In contrast, a scale transformation of one subsystem with respect to another is independent of whether those two subsystems are in the same place or not, and can indeed result in a change in the relative scale of the two sub-systems even at the same spacetime location.

3 Gauge symmetry

Following an article in the *Scientific American* by G. 't Hooft ([1980]), several authors (Auyang [1995]; Mainzer [1996]; and Kosso [2000]) have made a set of claims concerning gauge symmetry, for which the evidence they cite involves variations on the familiar 'two-slits experiment'.

In the case of local gauge symmetry, the examples concern electrons (the matter fields) and the electromagnetic field (the gauge fields). Two claims are made:

- (1) local gauge transformations of the matter fields alone (i.e. local phase transformations) are not symmetry transformations;
- (2) local gauge transformations of the matter fields plus the gauge fields are symmetry transformations having indirect *but not direct* empirical significance.

Both these claims are correct, but the evidence cited in each case is mistaken, and this means that the accompanying interpretation of local gauge symmetry, including the understanding of its empirical status, is wrong.

There is also a problem in the case of global gauge symmetry. The claim here is that a global phase transformation is a symmetry transformation, and that it has *direct* empirical significance. However, the evidence cited is, once again, mistaken, showing that the understanding of global gauge symmetry is flawed. We discuss global gauge symmetry later, but first the local case.

3.1 Local gauge symmetry

3.1.1. Discussion of the first claim. Consider first the claim (1) above, that local gauge transformations of the matter fields alone are not symmetry transformations. This is true. The Schrödinger equation for the free electron is:

$$\frac{-\hbar^2}{2m} \nabla^2 \psi(x, t) = i\hbar \frac{\partial}{\partial t} \psi(x, t), \quad (1.1)$$

and this is not invariant under a local phase transformation of the wavefunction for the electron,

$$\psi \rightarrow \psi' = \psi e^{iq\theta(x, t)}, \quad (1.2)$$

where $\theta(x, t)$ is an arbitrary function of space and time. This is not in question; the problem arises when 't Hooft, Auyang, Mainzer and Kosso appeal to a variation on the two-slits experiment with electrons, as described by 't Hooft in his 1980 article, as experimental evidence of the failure of invariance. However, this experiment *cannot* be interpreted as the implementation of a local phase transformation (1.2), as we will now discuss.

Let the wavefunction associated with the beam of electrons be Ψ , where for each closed loop between source and screen

$$\Psi = \frac{1}{\sqrt{2}} (\psi_I + \psi_{II}), \quad (1.3)$$

ψ_I and ψ_{II} being associated with paths through the first and second slits respectively. An interference pattern is produced at the screen as a result of the phase relations between ψ_I and ψ_{II} at each point along the screen. Call this interference pattern A. If we now insert a phase shifter into the path of only one component of the wavefunction, say ψ_I , we obtain a new interference pattern, say pattern B. According to 't Hooft ([1980], p. 98), this experiment constitutes evidence that a local gauge transformation of the matter fields alone is not a symmetry transformation:

a theory of the electron fields alone, with no other forms of matter or radiation, is
not invariant with respect to a ... local gauge transformation.

In other words, inserting the phase shifter into the path of one component of the wavefunction is to be regarded as actively implementing a local gauge transformation, and there are empirical consequences of this transformation: the change in the interference pattern is to be viewed as a failure of local gauge

symmetry. Kosso ([2000]) uses 't Hooft's example in his discussion of the empirical and observational status of local gauge symmetry. He writes (p. 95):

Passing the one beam through a half-wave plate, for example, while leaving the other beam alone, amounts to a local phase shift, that is, a gauge transformation. In this way, the active gauge transformation is observable.

Similarly, Auyang ([1995], p. 57) concludes that:

In the two-slits experiment, the interference pattern changes with a local phase change. This means that the free electron Lagrangian is *not* invariant under local phase transformations. The failure of invariance is unsatisfactory...

We believe that this analysis of the experiment is mistaken, and that the insertion of a phase-shifter into one path of the two-slits experiment cannot be interpreted as a local phase transformation. Introducing a phase shifter into one path results in new wavefunction, Ψ'' , (related to a new ray in Hilbert space):

$$\Psi = \frac{1}{\sqrt{2}}(\psi_I + \psi_{II}) \rightarrow \Psi'' = \frac{1}{\sqrt{2}}(\psi_I e^{i\Lambda} + \psi_{II}), \quad (1.4)$$

where Λ is a constant, and the components of the wavefunction have been transformed by

$$\begin{aligned} \psi_I &\rightarrow \psi_I'' = \psi_I e^{i\Lambda} \\ \psi_{II} &\rightarrow \psi_{II}'' = \psi_{II}. \end{aligned} \quad (1.5)$$

Thus, the relative phase of the two components – i.e. of ψ_I with respect to ψ_{II} – is changed by a constant amount (i.e. by the same amount at every point on the screen, and with no time variation). We use the new wavefunction, Ψ'' , to calculate the new interference pattern. Now compare this with a local phase transformation of Ψ , where we interpret 'local' in the Weyl sense, which is the standard sense in which the terminology 'local gauge transformation' is used. A local gauge transformation of the wavefunction has the following form:

$$\Psi = \frac{1}{\sqrt{2}}(\psi_I + \psi_{II}) \rightarrow \Psi''' = e^{i\theta(x,t)} \frac{1}{\sqrt{2}}(\psi_I + \psi_{II}) = e^{i\theta(x,t)} \Psi. \quad (1.6)$$

This new wavefunction, Ψ''' , differs from that obtained by inserting a phase shifter into one component of the wavefunction: Ψ''' and Ψ'' are not the same. This is the point, in a nutshell. A local gauge transformation of the matter fields (i.e. a local phase transformation) has the form (1.6), and this differs from a relative phase transformation which has the form (1.4). The experiment described by 't Hooft realises a relative phase transformation, and not a local phase transformation.

If more needs to be said in order to make the point convincing, then the easiest way to see that the wavefunctions Ψ''' and Ψ'' differ is to focus on what happens at the screen, and in particular on the relative phase of ψ_I with respect to ψ_{II} at each point along the screen, since this is what gives rise to the interference pattern. The expectation values associated with position for Ψ''' and Ψ are the same, *and so both give rise to the same interference pattern* (pattern A). The fact that the interference pattern is unchanged is readily seen by noticing that the *relative phase* of ψ_I with respect to ψ_{II} at any point x_1 is

unchanged, even though the *overall* phase of the new wavefunction Ψ'' differs from that of the original wavefunction Ψ point by point. The wavefunction Ψ'' , on the other hand, gives rise to a *different* interference pattern (pattern B). The change in the interference pattern is due to the change in the relative phase of ψ_I with respect to ψ_{II} at each point along the screen. A *local gauge transformation*, such as (1.6), will not achieve this. Local gauge freedom is the freedom to vary the overall phase of the wavefunction from point to point, but it is *not* the freedom to vary the phase of ψ_I with respect to ψ_{II} at a single spacetime point. Under a local gauge transformation the phase of ψ_I at some point on the screen will be changed by the same amount as the phase of ψ_{II} at that same point. What we need in order to change the interference pattern is a *relative phase transformation* of ψ_I with respect to ψ_{II} at each point on the screen.

One source of confusion might perhaps be that in spatial regions where ψ_I and ψ_{II} *do not overlap*, a local gauge transformation (1.6) and a relative phase transformation (1.4) can be made to coincide by choosing (i) $\theta(x, t) = 0$ in (1.6) throughout that part of the spatial region occupied by ψ_{II} , and (ii) $\theta(x, t) = \Lambda$ wherever ψ_I is present. However, an interference pattern occurs only where ψ_I and ψ_{II} overlap, and clearly these conditions on $\theta(x, t)$ cannot be met in such a region.

This may leave a puzzle: how can it be that (1.6) and (1.4) coincide at some time $t = 0$, when the wavepackets associated with ψ_I and ψ_{II} do not overlap, but that when we evolve the wavefunction forward we reach different predictions for the interference pattern at the screen? The reason is that Ψ'' and Ψ''' obey different dynamics: Ψ'' satisfies the free Schrödinger equation but Ψ''' does not. In more detail, the situation is as follows. The two wavefunctions Ψ and Ψ'' give rise to distinct interference patterns, due to the relative phase transformation of the components of Ψ (see (1.4)). Nevertheless, both Ψ and Ψ'' satisfy the same dynamical equations, the Euler-Lagrange equations associated with the Lagrangian for a free complex scalar field, i.e. the free Schrödinger equation (1.1). The two wavefunctions Ψ and Ψ''' , on the other hand, give rise to the *same* interference pattern, but Ψ''' does not satisfy the free Schrödinger equation. In order to find equations of motion for Ψ''' we have to introduce an additional dependent variable, which (for empirical reasons) we associate with the electromagnetic potential; this is the so-called gauge principle in action (see Brown [1999]). The Schrödinger equation for a single spinless particle moving in an electromagnetic potential is:

$$\frac{-\hbar^2}{2m} \left(\nabla - \frac{q}{c} \mathbf{A} \right)^2 \psi + q\phi\psi = i\hbar \frac{\partial}{\partial t} \psi \quad (1.7)$$

where $\mathbf{A}(x, t)$ is the vector potential and $\phi(x, t)$ is the scalar potential. This equation is not invariant under local phase transformations of the wavefunction ψ alone, but it is invariant under local gauge transformations of the form

$$\left. \begin{aligned} \psi &\rightarrow \psi' = \psi e^{i\frac{q}{\hbar c}\chi} \\ \mathbf{A} &\rightarrow \mathbf{A}' = \mathbf{A} + \nabla\chi \\ \phi &\rightarrow \phi' = \phi - \frac{\partial\chi}{\partial t} \end{aligned} \right\} \quad (1.8)$$

where $\chi = \chi(x, t)$ is an arbitrary smooth function of space and time. Even in the case where there is no electromagnetic field present, a local gauge transformation involves not only a local phase transformation of the matter fields, but also a local gauge transformation of the electromagnetic potential. This leads to the different predictions when we evolve forwards from the region where ψ_I and ψ_{II} do not overlap to the region where they do.

In short, then, the experiment described by 't Hooft constitutes a relative phase transformation and *not* a local gauge transformation of the matter fields.

3.1.2 Discussion of the second claim. In our discussion of the above case we introduced the full local gauge transformations (1.8) in which the electromagnetic potentials $\mathbf{A}(x, t)$ and $\phi(x, t)$ are transformed along with the matter fields $\psi(x, t)$. This brings us to the second of the claims made by 't Hooft and co., i.e., that local gauge transformations of the matter fields plus the gauge fields are symmetry transformations having indirect *but not direct* empirical significance. Once again, the claim is correct: first, the transformations (1.8) are symmetry transformations of the equation (1.7), and second, these transformations have indirect but not direct empirical significance. We will explain the reason why this latter part of the claim is true once we have discussed the evidence cited by 't Hooft and co., since their account of why this claim holds depends on that evidence.

According to 't Hooft ([1980, p. 98]), introducing electromagnetism into our theory allows us to achieve local gauge symmetry: 'describing the two fields together the local symmetry can be extended to both of them.' This is something that can be observed, he claims, using the Aharonov-Bohm solenoid. The Aharonov-Bohm set-up is described in various textbooks, including Ballentine ([1990], section 11-4). A long solenoid is placed behind the two slits and the magnetic field is contained entirely within the solenoid, in the 'shadow' region behind the two slits, where the wavefunction does not pass, and furthermore it is shielded from electrons. Nevertheless, when the electromagnetic field is switched on, a change in the interference pattern results, and the field strength can be chosen so that we get pattern B. We can also (as Auyang does) insert a phase-shifter into one beam so that the relative phase shift due to the Aharonov-Bohm solenoid is compensated for by the phase-shifter, and we recover interference pattern A. How should these further scenarios be interpreted? Following 't Hooft, Mainzer ([1996], p. 423) writes:

If the phase is shifted ... behind only one slit, then the interference pattern changes. ...

Thus a local change of the interference pattern is realised experimentally. To 'restore the symmetry', a force field must be found which compensates for the local change of the phase shift.

Auyang describes the situation as follows ([1995], p. 56):

The effect of the local phase shift is compensated by the potential A_μ of the electromagnetic field. ... At some value of A_μ , the original interference pattern is recovered. This demonstrates that the interacting electron-electromagnetic field system can be made invariant under local phase transformations.

As we have seen, the local gauge symmetry of the Schrödinger equation (1.7) depends on the introduction of a new dependent variable, which we associate with the electromagnetic potential. *However*, as Auyang herself points out (p. 58), the Lagrangian associated with the Schrödinger equation (1.7) is locally gauge invariant even when the electromagnetic fields happen to vanish – the potential could be a fixed flat background potential with which no dynamical field is associated, and we would still have a locally gauge invariant theory. Local gauge invariance alone does not require non-zero electromagnetic fields (see Brown [1999], especially p. 53). Therefore, the presence of a non-zero electromagnetic field cannot be necessary for the phenomena to exhibit local gauge symmetry. To put the point another way, the result of inserting a phase-shifter into one path in the above experiment can be described by the locally gauge invariant equation (1.7), just as well as when the A-B solenoid is added. Local gauge symmetry is independent of whether there are electromagnetic fields present.

It is true that, in the scenario described, a combined transformation of matter field (via a phase shifter) and electromagnetic potential (via the A-B solenoid) gives us empirical invariance. But this transformation of the matter field and of the electromagnetic potential is not a local gauge transformation. What is shown by this experiment is that both the phase-shifter and the Aharonov-Bohm solenoid can be used to introduce a change in the relative phase of ψ_I with respect to ψ_{II} , and hence that the electromagnetic field strength can be *carefully chosen* so that the specific relative phase change due to a given phase-shifter is cancelled out by the relative phase change induced by that particular current through the Aharonov-Bohm solenoid.

In short, the analysis of the above experiments offered by 't Hooft and co. must be mistaken: no local gauge transformation of a system of matter and gauge fields can bring an electromagnetic field into (or out of) existence; no change in the interference pattern can result from a local gauge transformation.

What, then, of the claim that local gauge symmetry has only indirect but not direct empirical significance? Kosso ([2000]) has argued that local gauge symmetry can be observed only indirectly on the grounds that the insertion of a phase shifter into one path results in a violation of the symmetry, and the symmetry can be restored only by adding a new force (electromagnetism) to our theory. In this he again follows 't Hooft, and similar arguments are given by Auyang ([1995], p. 57-8) and Mainzer ([1996], p. 421-4). As we have seen, this analysis and explanation of the status of local gauge symmetry cannot be right. The reason why local gauge symmetry has indirect but no direct empirical significance lies elsewhere.

Consider first the question of whether local gauge symmetry has direct empirical significance: i.e. is there a 'Galilean ship' analogue for local gauge transformations? Consider an effectively isolated system

of matter plus gauge fields, and transform this relative to some other system, such that the resulting scenario is empirically (= observationally = measurably, in this case) distinct? The measurable quantities are the relative phase relations between components of the wavefunction of a single system, and the electromagnetic field strengths. The latter are unchanged by a local gauge transformation (only the potentials, not the field strengths, transform), and so no empirical consequences can arise. In the former case, the overall phase of one system relative to another is not measurable, and so again there can be no observable consequences of transforming one system with respect to another. The only remaining option is to consider a region where the wavefunction can be decomposed into two spatially separated components, and then to apply a local gauge transformation to one region (i.e. to the component of the wavefunction in that region, along with the electromagnetic potential in that region) and not to the other. But then either the transformation of the electromagnetic potential results in the potential being discontinuous at the boundary between the ‘two subsystems’, in which case the relative phase relations of the two components are undefined (it is meaningless to ask what the relative phase relations are), or the electromagnetic potential remains continuous, in which case what we have is a special case of a local gauge transformation on the entire system – and this of course brings us back to where we started – such a transformation has no observable consequences. The most fundamental point is perhaps this: in theories with local gauge symmetry, the matter fields are embedded in a gauge field, and the local symmetry is a property of both sets of fields *jointly*. Because of this, we cannot apply a symmetry transformation to one subsystem of matter fields independently of some other subsystem – we must also transform the gauge fields in order for the transformation to be a symmetry transformation, and both subsystems of matter fields are embedded in the same gauge fields. Thus, a transformation applied to one subsystem will involve the other subsystem, even if only because the transformation of the gauge field goes smoothly to the identity. In conclusion, there can be no analogue of the Galilean ship experiment for local gauge transformations, and therefore local gauge symmetry has only *indirect* empirical significance (being a property of the equations of motion). We will have more to say about this indirect empirical significance in section 5, below.

3.2 Global Gauge Symmetry

Return to the original two-slits experiment, and now insert identical phase shifters into each path. The result will be an unchanged interference pattern. According to 't Hooft, this demonstrates the global gauge invariance of the electron field; Auyang, Mainzer and Kosso agree. Kosso ([2000], p. 83) writes:

This invariance is easily observed, and the experiment as a whole amounts to an observation of the internal, global symmetry in nature.

However, the experiment described does *not* constitute an observation of global gauge symmetry, and this is because it does not meet the first of the two conditions required by Kosso himself: we have not observed that a transformation has taken place. Mathematically, inserting the phase shifters corresponds to the following transformation:

$$\begin{aligned}\Psi &= \frac{1}{\sqrt{2}}(\psi_I + \psi_{II}) \\ \rightarrow \Psi' &= \frac{1}{\sqrt{2}}(e^{i\theta}\psi_I + e^{i\theta}\psi_{II}) = e^{i\theta} \frac{1}{\sqrt{2}}(\psi_I + \psi_{II}) = e^{i\theta}\Psi.\end{aligned}\tag{1.9}$$

Hence, Ψ and Ψ' differ mathematically by an overall global phase.⁹ However, this overall global phase is of no empirical significance: physically, Ψ and Ψ' represent exactly the same quantum mechanical state, indistinguishable in every way. This means that a global gauge transformation cannot be used to create an empirically distinguishable scenario. Perhaps it will be objected that the insertion of the phase shifters is enough to ensure that the transformation of Ψ to Ψ' is physically implemented. But it is not the means by which the alleged transformation is carried out that guarantees that we have a physical transformation, it is the resulting empirically distinct scenario. Another objection might take the form of a counterfactual: if global gauge symmetry were violated then inserting the phase shifters would lead to an empirically different outcome, but it doesn't, so observing no change is equivalent to observing global gauge symmetry directly. But this will not serve our needs: if global gauge symmetry were violated, then we would have evidence that inserting a phase shifter changes the physical state of the system; in the absence of this, we have no evidence that a transformation has been physically implemented. Criterion (1) for direct empirical significance cannot be met, and so the empirical significance of global gauge symmetry is solely indirect.¹⁰

We end by noting as different possible approach to the direct observation of a global phase transformation. It might be tempting to interpret the relative phase transformation (1.4) as a global phase transformation of ψ_I with respect to ψ_{II} , and thereby claim that *global* phase transformations are directly observable in that experiment, rather than in the case where we insert phase shifters into both paths. The crucial issue here is whether the two components of the wavefunction, ψ_I and ψ_{II} , can be interpreted as representing genuine subsystems of Ψ . Our position is that only Ψ represents a physical system, with ψ_I representing one (basis-dependent) component of the wavefunction Ψ . (If we kill the second component of Ψ (i.e. ψ_{II}), then ψ_I represents a physical system, but now the interference pattern between ψ_I and ψ_{II} cannot be observed, and we are back to square one.) The point is made particularly vivid by considering a single electron passing through the two slits: on our view there is only one system here, described by Ψ , and the components ψ_I and ψ_{II} do not represent subsystems of the electron. The same general point holds

⁹ Notice that this conclusion rests on the assumption that the electron beam closely approximates a plane wave. If we were to consider wavepackets instead, then the insertion of suitably chosen dispersive phase shifters would lead to a delay in the time of arrival of the wavepacket at the screen, and hence we would be able to empirically distinguish the case with the phase shifters present from the ones without. However, in the case of wavepackets, not only is it extremely difficult in practice to preserve the form of the wavepacket as it passes through the phase shifter, but also the resulting wavefunction is not related to the original wavefunction by a global phase transformation. The resulting time delayed wavefunction $\Psi(x, t + \tau)$ is only equivalent to $e^{i\theta}\Psi(x, t)$ for the case of the plane wave.

¹⁰ This conclusion requires the qualification that we are assuming the absence of a further interaction that does not respect the symmetry in question.

even on an ensemble interpretation, and – at least in the absence of further argument – it seems that ψ_I and ψ_{II} , cannot be interpreted as representing genuine subsystems.

4 Space-time symmetries

Global space-time symmetries, such as the spatial translations and rotations of particle mechanics, have an active interpretation in terms of subsystems of the universe; furthermore, these active transformations are implementable in experiments such as the Galilean ship experiment. This much is uncontroversial. The question arises of whether *local* space-time symmetries – arbitrary co-ordinate transformations that leave the explicit form of the equations of motion unaffected – also have an active interpretation. As in the case of local gauge symmetry, it has been argued in the literature that the introduction of a force is required to ‘restore’ local symmetry (see Rosen [1990] and Kosso [2000]). In the case of arbitrary co-ordinate transformations, the force invoked is gravity. Once again, we believe that the arguments (though seductive) are wrong, and that it is important to see why.

Kosso’s discussion of arbitrary coordinate transformations is analogous to his argument with respect to local gauge transformations. He writes ([2000], p. 89):

Observing this symmetry requires comparing experimental outcomes between two reference frames that are in variable relative motion, frames that are relatively accelerating or rotating.

He goes on:

One can, in principle, observe that this sort of transformation has occurred. ... just look out of the window and you can see if you are speeding up or turning with respect to some object that defines a coordinate system in the reference frame of the ground.

Then:

Now do the experiments to see if the invariance is true. Do the same experiments in the original reference frame that is stationary on the ground, and again in the accelerating reference frame of the train, and see if the physics is the same. One can run the same experiments, with mechanical forces or with light and electromagnetic forces, and observe the results, so the invariance should be observable.

And he concludes:

But when the experiments are done, the invariance is not directly observed. Spurious forces appear in the accelerating system, objects move spontaneously, light bends, and so on. ... The physics is different.

In other words, if we place ourselves at rest first in an inertial reference frame, and then in a non-inertial reference frame, our observations will be distinguishable. For example, in the non-inertial reference frame objects that are seemingly force-free will appear to accelerate, and so we will have to introduce extra, ‘spurious’, forces to account for this accelerated motion. The transformation described by Kosso is clearly not a symmetry transformation. Despite that, his claim appears to be that if we move to General Relativity

this transformation becomes a symmetry transformation. In order to assess this claim, let's begin by considering Kosso's experiment from the point of view of classical physics.

Suppose that we describe these observations using Newtonian physics and Maxwell's equations. We would not be surprised that our descriptions differ depending on the choice of coordinate system: arbitrary coordinate transformations are not symmetries of the Newtonian and Maxwell equations of motion as usually expressed. Nevertheless, we are free to re-write Newtonian and Maxwellian physics in generally covariant form. *But notice:* the arbitrary coordinate transformations now apply *not just* to the Newtonian particles and the Maxwellian electromagnetic fields, *but also* to the metric, and this is necessary for general covariance.

Kosso's example is given in terms of passive transformations – transformations of the coordinate systems in which we re-coordinatise the fields. In the Kosso experiment, however, we re-coordinatise the matter fields *without* re-coordinatising the metric field. This *is not* achieved by a mere coordinate transformation in generally covariant classical theory: a passive arbitrary coordinate transformation induces a re-coordinatisation of not only the matter fields but also the metric. The two states described by Kosso are *not* related by an arbitrary coordinate transformation in generally covariant classical theory. Further, such a coordinate transformation applied to only the matter and electromagnetic fields is not a symmetry of the equations of Newtonian and Maxwellian physics, regardless of whether those equations are written in generally covariant form.

With this in mind, let's turn our attention to General Relativity. Suppose that we use General Relativity to describe the above observations. Kosso suggests that in General Relativity the observations made in an inertial reference frame will indeed be related by a symmetry transformation to those made in a non-inertial reference frame. He writes ([2000], p. 90):

The invariance can be restored by revising the physics, by adding a specific dynamical principle. This is why the local symmetry is a dynamical symmetry. We can add to the physics a claim about a specific force that restores the invariance. It is a force that exactly compensates for the local transform. In the case of the general theory of relativity the dynamical principle is the principle of equivalence, and the force is gravity. ... With gravity included in the physics and with the windows of the train shuttered, there is no way to tell if the transformation, the acceleration, has taken place. That is, there is now no difference in the outcome of experiments between the transformed and untransformed systems. The force pulling objects to the back of the train could just as well be gravity. Thus the physics, all things including gravity considered, is invariant from one locally transformed frame to the next. The symmetry is restored.

This analysis mixes together the equivalence principle with the meaning of invariance under arbitrary coordinate transformations in a way which seems to us to be confused, with the consequence that the account of local symmetry in General Relativity is mistaken.

Einstein's field equations are covariant under arbitrary smooth coordinate transformations. However, as with generally covariant Newtonian physics, these symmetry transformations are transformations of the matter fields (such as particles and electromagnetic radiation) *combined with* transformations of the metric. Kosso's example, as we have already emphasised, re-coordinatises the matter fields *without* re-coordinatising the metric field. So, the two states described by Kosso are *not* related by an arbitrary coordinate transformation even in General Relativity. We can put the point vividly by locating ourselves at the origin of the coordinate system: I will always be able to tell whether the train, myself, and its other contents are all freely falling together, or whether there is a relative acceleration of the other contents relative to the train and me (in which case the other contents would appear to be flung around). This is completely independent of what coordinate system I use – my conclusion is the same regardless of whether I use a coordinate system at rest with respect to the train or one that is accelerating arbitrarily. (This coordinate independence is, of course, the symmetry that Kosso sought in the opening quotation above, but his analysis is mistaken.)

What, then, of the equivalence principle? The Kosso transformation leads to a physically and observationally distinct scenario, and the principle of equivalence is not relevant to the difference between those scenarios. What the principle of equivalence tells us is that the effect in the second scenario, where the contents of the train appear to accelerate to the back of the train, may be due to acceleration of the train in the absence of a gravitational field, or due to the presence of a gravitational field in which the contents of the train are in free fall but the train is not. Mere coordinate transformations cannot be used to bring real physical forces in and out of existence.

It is perhaps worthwhile briefly indicating the analogy between this case and the gauge case considered above. Active arbitrary coordinate transformations in General Relativity involve transformations of both the matter fields and the metric, and they are symmetry transformations having no observable consequences. Coordinate transformations applied to the matter fields alone are no more symmetry transformations in General Relativity than they are in Newtonian physics (whether written in generally covariant form or not). Such transformations do have observational consequences. Analogously, local gauge transformations in locally gauge invariant relativistic field theory are transformations of both the particle fields and the gauge fields, and they are symmetry transformations having no observable consequences. Local phase transformations alone (i.e. local gauge transformations of the matter fields alone) are no more symmetries of this theory than they are of the globally phase invariant theory of free particles. Neither an arbitrary coordinate transformation in General Relativity, nor a local gauge transformation in locally gauge invariant relativistic field theory, can bring forces in and out of existence: no generation of gravitational effects, and no changes to the interference pattern.

5 Direct and indirect empirical significance again

In so far as internal global symmetries and local symmetries are perfect symmetries (i.e., there are no other interactions that fail to respect the symmetry in question), they have no direct empirical significance, only indirect empirical significance. These symmetries are properties of the associated laws of motion, and

therefore have consequences for the behaviour of systems described by these laws. These consequences can be vividly highlighted using three theorems that are derivable from a mathematical problem posed by Emmy Noether (Noether [1918]).¹¹ Noether's first theorem is the most famous of these theorems, and it connects *global* symmetries (both external and internal) with conservation laws.¹² In the same paper Noether also proved a second theorem associated with *local* symmetries; and, with Noether's assistance, Klein derived results that are related to a third theorem which is again associated with local symmetries. We call this third theorem the Boundary theorem, for reasons associated with its derivation.

Noether's First Theorem relates the global phase symmetry of the Schrödinger equation for a free particle (1.1) to the condition that normalization must be time-independent. In the case of local symmetry, Noether's second theorem and the Boundary theorem demonstrate that the restrictions on the possible form of a theory with a given local symmetry are very dramatic. For example, in any theory for which (1.8) are symmetry transformations, we can use Noether's second theorem to show that not all the equations of motion are independent of one another:¹³ this leads to an underdetermination in the theory which may be removed by opting for a theory in which the dynamical fields act non-locally (see Belot [1998] for a detailed discussion of this problem in electromagnetism and of the options available).¹⁴ Using the Boundary theorem we can show that the matter fields act as the source of the electromagnetic fields, and what the form of the coupling is between the matter fields and the electromagnetic potential; we can also derive the form of a current that is conserved when either the matter field equations or the electromagnetic field equations are satisfied. These and other results of the Noether and Boundary theorems are discussed in Brading and Brown [2003]; a discussion of these theorems in relation to general covariance and General Relativity can be found in Brown and Brading [2002]. What they show is that the requirement of local symmetry imposes extremely strong restrictions on the possible form of a theory, and it is this that gives local symmetry its indirect – but potentially very powerful – empirical significance.

6 Conclusion

By way of brief conclusion then, the central point is an important distinction between continuous global space-time symmetries and all other continuous symmetries. Global space-time symmetries have a special status, both theoretically and practically: theoretically, they have an active interpretation in the sense that a

¹¹ This problem can be posed for all theories that can be given a Lagrangian formulation. See Brading and Brown ([2003]).

¹² In fact, Noether's theorems *do not* apply to all continuous symmetries of the Euler-Lagrange equations, and the connection depends on satisfaction of further conditions. These being met, in field theory the theorem connects a linear combination of Euler derivatives with a divergence expression, such that if all the fields on which the Lagrangian depends satisfy Euler-Lagrange equations, the divergence expression vanishes. This can then be converted into a conservation law, subject to suitable boundary conditions being satisfied. See Brading and Brown ([2003]).

¹³ More accurately, Noether's second theorem tells us that not all the Euler derivatives are independent of one another.

¹⁴ For the general connection between Noether's second theorem and indeterminism, see Barbashov and Nesterenko ([1983], section 7), and Brading and Brown ([2003]). See also Earman ([2002]) and Redhead ([2002]) for further philosophical discussion.

symmetry transformation applied to a subsystem of the universe yields an empirically distinct scenario; and, furthermore, instances of these active transformations are implementable in practice through the use of effectively isolated subsystems. Neither global internal symmetries, nor local symmetries of either variety, have even a theoretical active interpretation of this kind.

Acknowledgements

We would like to thank Elena Castellani, Steve Lidia, Nick Maxwell, Simon Saunders and David Wallace for useful discussions. K.A.B. thanks Wolfson College, Oxford, for their support.

References

- Auyang, S. Y. [1995]: *How is Quantum Field Theory Possible?*, Oxford: Oxford University Press.
- Ballentine, L. E. [1990]: *Quantum Mechanics*, New Jersey: Prentice-Hall.
- Barbashov, B. M., and Nesterenko, V. V. [1983]: ‘Continuous Symmetries in Field Theory’, *Fortschr. Phys.* **31**, pp. 535-67.
- Barbour, J. [1989]: *Absolute or Relative Motion?* vol. 1, *The discovery of dynamics*, Cambridge: Cambridge University Press.
- Belot, G. [1998]: ‘Understanding Electromagnetism’, *British Journal for the Philosophy of Science* **49**, pp. 531-55.
- Brading, K. A., and Brown, H. R. [2003]: ‘Noether’s Theorems, Gauge Symmetries and General Covariance’, in preparation.
- Brown, H. R. [1999]: ‘Aspects of Objectivity in Quantum Mechanics’, in *From Physics to Philosophy*, ed. J. Butterfield and C. Pagonis, Cambridge: Cambridge University Press, 1999, pp. 45-70. E-print: PITT-PHIL-SCI00000223.
- Brown, H. R., and Brading, K. A. [2002]: ‘General Covariance from the Perspective of Noether’s Theorems’, *Diálogos* **79**, pp. 59-86.
- Budden, T. [1997]: ‘Galileo’s Ship and Spacetime Symmetries’, *British Journal for the Philosophy of Science* **48**, pp. 483-516.
- Castellani, E. [2002]: ‘Symmetry, Quantum Mechanics, and Beyond’, *Foundations of Science* **7**, pp.181-96. E-print: PITT-PHILSCI 00000573.

Earman, J. [2002]: 'Gauge Matters', *Philosophy of Science* **69**, pp. S209-S220. E-print: PITT-PHIL-SCI00000070.

Galileo [1967]: *Dialogue Concerning the Two Chief World Systems*, English translation, second revised edition, Berkeley: University of California Press.

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

Lee, T. D. [1971]: 'Symmetry Principles in Physics', in A. Zichichi (ed.), *Elementary Processes at High Energies*, Part A, pp. 306-19, New York: Academic Press.

Mainzer, K. [1996]: *Symmetries of Nature*, Berlin: Walter de Gruyter.

Martin, C. [2003]: 'Continuous Symmetry and the Foundations of Physics', in K. A. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press.

Miller, A. I. [1981]: *Albert Einstein's Special Theory of Relativity: emergence (1905) and early interpretation (1905-1911)*, Reading MA: Addison-Wesley Publishing Company Inc.

Morrison, M. [1995]: 'The New Aspect: Symmetries as Meta-Laws - Structural Metaphysics', in Weinert (ed.), 1995, pp. 157-87.

Noether, E. [1918]: 'Invariante Variationsprobleme', *Nachr. d. Konig. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse*, pp. 235-57. English translation: Tavel, M. A. [1971]: 'Noether's theorem', *Transport Theory and Statistical Physics* **1**, pp. 183-207.

O'Riifeartaigh, L. [1997]: *The Dawning of Gauge Theory*, Princeton Series in Physics, Princeton: Princeton University Press.

Redhead, M. [2002]: 'The Interpretation of Gauge Symmetry', in M. Kuhlmann, H. Lyre, and A. Wayne (eds.), *Proceedings of the Conference Ontological Aspects of Quantum Field Theory*, Bielefeld, Oct. 11-13, 1999, World Scientific (Singapore), forthcoming. Reprinted in K. A. Brading and E. Castellani (eds.), 2003, *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press.

Rosen, J. [1990]: 'Fundamental Manifestations of Symmetry in Physics', *Foundations of Physics* **20**, 283-307.

Sklar, L. [1996]: 'Saving the Phenomena', in D. Papineau, ed., 1996, *The Philosophy of Science*, Oxford: Oxford University Press, pp. 61-81. Reprinted from *Philosophical Topics* **13**, 1982, pp. 49-72.

't Hooft, G. [1980]: 'Gauge Theories and the Forces Between Elementary Particles', *Scientific American* **242**, pp. 90-166.

Weyl, H. [1918]: 'Gravitation and Electricity', English translation in O'Riifeartaigh, 1997, pp. 24-37, page numbers refer to this translation.

Wigner, E. [1967]: *Symmetries and Reflections*, Bloomington: Indiana University Press.

Yang, C. N., and Mills, R. L. [1954]: 'Conservation of Isotopic Spin and Isotopic Gauge Invariance', *Phys. Rev.* **96**, pp. 191-5. Reprinted in O'Riifeartaigh, 1997, pp. 186-96.