Empirical consequences of symmetries

Hilary Greaves[*]
David Wallace[†]

March 28, 2011

1   Introduction

There is something of a tension between two aspects of symmetry in physics. On the one hand, there is a widespread consensus[1] that 'two states of affairs related by a symmetry transformation are really just the same state of affairs differently described': that is, that if two mathematical models of a physical theory are related by a symmetry transformation, then those models represent one and the same physical state of affairs. This seems to give symmetry a purely formal role in physics: symmetries are not features of the world, but merely features of our method of describing the world.

On the other hand, it seems to be a matter of plain historical fact that the observable consequences of symmetries have guided physicists in their construction of theories ever since Galileo's ship thought experiment. It is *prima facie* mysterious how this fact is compatible with the account of symmetries given above.

The mystery is solved, or at least demoted from mystery to puzzle, once we recognise that a transformation can be a symmetry of a *subsystem* of the world without being a symmetry of the *whole* world.[2] As such, the transformation gives rise to a different state of affairs, but has no detectable consequences for measurements confined to the subsystem.

For example, Galileo famously noted that, provided the interior of a ship is suitably dynamically isolated from goings-on outside, experiments inside the cabin of a ship could not tell whether the ship was stationary or moving with uniform velocity with respect to the shore:

> Shut yourself up with some friend in the main cabin below decks on some large ship ... With the ship standing still, observe carefully [the results of various experiments you conduct inside the cabin. Then]

---

[*]Somerville College, Oxford, and Philosophy Faculty, University of Oxford
[†]Balliol College, Oxford, and Philosophy Faculty, University of Oxford
[1]See, e. g. , Saunders (2003) and references therein.
[2]This observation has been made by (amongst others) Brown and Sypel (1995), Kosso (2000), Brading and Brown (2004) and Healey (2009).

have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still. (Galileo 1967, pp. 186–7)

Here, the subsystem is the ship's cabin, and the transformation is a uniform boost of the ship relative to the shore. Restricted to the cabin, the transformation is a symmetry, and as such, has no detectable consequences as long as the cabin remains dynamically isolated from the exterior. But it is not a symmetry of the world as a whole: measurements in which the cabin interacts with the outside world (such as those made by opening the window and looking out) will in general have results which depend on whether or not the transformation has been carried out.

This line of thought *seems* to suggest an important distinction between "global" symmetry transformations, in which the same transformation must be carried out at each point of space and time in order for the transformation to be a symmetry (like translations and boosts in special relativity or global phase transformations in Klein-Gordon theory) and "local" symmetry transformations, in which the transformation may vary arbitrarily from point to point (as in the gauge symmetry of electromagnetism or the diffeomorphism symmetry of general relativity). *Prima facie*, any transformation which is a local symmetry of each subsystem of a system is also a local symmetry of the system as a whole, and hence does not relate different states of affairs to one another: hence, it seems (*prima facie*), we should expect that whilst global symmetries have direct empirical significance, local symmetries do not.

Brading and Brown (2004), Healey (2007), and Kosso (2000) all defend this view that local symmetries cannot have direct empirical significance. Brading and Brown, in fact, go further and argue that only global *spacetime* symmetries (as opposed to global internal symmetries like global phase transformations) can have direct empirical significance.

There are at least three reasons to be deeply puzzled by these claims, though. First: in any theory that has a 'local' symmetry group, the 'global' symmetries remain as a subgroup of that local symmetry group. (For example, in general relativity, the 'global' translations and boosts form a subgroup of the group of all diffeomorphisms.) It is therefore logically impossible that all global symmetries, but no local symmetries, can have direct empirical significance; and 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.

Secondly (and relatedly): very often in the history of physics, a theory with a 'global' symmetry is superseded by a more accurate theory in which that symmetry is 'localised': thus, for example, the (global) Poincaré symmetry group of a special-relativistic theory is extended to the (local) diffeomorphism group of general relativity, and the (global) group of constant shifts in electric potential in electrostatics is extended to the (local) group of 'gauge' transformations of

the four-potential in electrodynamics. It would be extremely odd (at best) if we had a theoretical explanation of an empirically observable symmetry while we stuck with the less accurate theory, but nothing of this explanation remained once we passed to the larger symmetry group in the more accurate theory: if for example, there was nothing we could say *from the perspective of general relativity* to explain Galileo's ship in terms of boost invariance. Scientific realists should regard this as not only odd, but also rather worrying: if the explanatory successes of old theories aren't reproduced in their successors, much of the justification for realism is undermined.

Thirdly: if the relationship between the boost invariance of Newtonian mechanics and the Galileo-ship thought experiment provides the defining example of the sense in which some theoretical symmetries may have 'direct empirical significance', we can (and will) sketch several examples that *seem*, *prima facie* at least, to be perfect analogs of the Galileo-ship scenario for cases of local symmetry.

Our purpose in this paper is to present and illustrate a new framework for thinking about the empirical consequences of symmetries which clears up these confusions. We will show that although a distinction related to the global/local distinction does determine whether or not a symmetry has direct empirical significance *in a particular empirical situation*, 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. In particular (we will argue), there are analogs of the Galileo-ship scenario for both diffeomorphism and gauge symmetries.

The structure of the paper is as follows. Section 2 introduces two examples — that of Faraday's Cage, and a modification of the two-slit experiment — that appear to be just such analogs of the Galileo-ship scenario for cases of local symmetry, and *registers* our disagreement with previous discussions (respectively, Healey (2009) and Brading and Brown (2004)) that claim to find relevant disanalogies, and hence to support the conventional wisdom in the face of such examples. To *justify* that disagreement requires us to develop a comprehensive theory and classification of the various circumstances under which one can perform a symmetry transformation on a subsystem without thereby performing a symmetry transformation on the universe, and to work out its application to a variety of examples; this task occupies the larger part of the paper. Section 3 begins to set up our framework, including the ways in which we regiment talk of symmetries, subsystems, environments, and boundary conditions governing compatibility between given subsystem and environment states. Section 4 applies this initial part of the abstract framework to the symmetries and subsystems of electrostatics, by way of illustrative example. Section 5 lays out the remainder of our framework, and defines the notions of 'interior' and 'boundary-preserving' symmetries, which we believe are crucial to an understanding of the phenomenon of Galileo-ship type scenarios. The next four sections apply the completed framework to four other theories. Section 6 discusses Newtonian gravity, and identifies a particular subgroup of that theory's symmetry group as corresponding to the 'Einstein's elevator' empirical symmetry. Up to this point,

our examples will have included only 'global' symmetries; sections 7–8 deal with 'local' symmetries. Section 7 discusses classical electromagnetism, and identifies the empirical symmetry of Faraday's cage as an empirical consequence of the local gauge symmetry of that theory. Section 8 discusses Klein-Gordon-Maxwell theory, and identifies the empirical symmetry in the modified two-slit setup that we first discuss in section 2 as an empirical consequence of the local gauge symmetry of *that* theory. Section 8 discusses general relativity, and identifies both Galileo's ship and Einstein's elevator as empirical consequences of particular subgroups of the diffeomorphism symmetry group of GR. Section 9 contains a compact presentation of our finished abstract framework, and a summary of our claims about the conditions under which symmetries can and cannot have empirical significance. Section 10 is the conclusion.

2   Analogs of Galileo's ship? Faraday's cage and t'Hooft's beam-splitter

Healey (2009) presents a puzzle for the 'conventional wisdom' about local and global symmetries; his puzzle will serve as a starting point for our discussion. To explain the puzzle — and to begin theorising about the empirical consequences of symmetries — we follow Healey in distinguishing between *empirical symmetries* and *theoretical symmetries*.

Theoretical symmetries are automorphisms of the class of mathematical models of a theory. For our purposes, we can *define* a theoretical symmetry to be a map $f : \Theta \to \Theta$ from the set Θ of models of a theory T to itself, such that any two models $m, f(m)$ related by that map represent the same physical state of affairs — the same possible world, if preferred — as one another. An example of a group of theoretical symmetries is given by the boost invariance of Newtonian mechanics: we work with a formulation of Newtonian mechanics that uses coordinate systems, and a transformation $\mathbf{x} \mapsto \mathbf{x} - \mathbf{v}t$ for the whole universe results merely in a redescription of the same physical state of affairs in a different inertial frame (that is, there is no distinction between 'active' and 'passive' transformations when the transformation is performed on the universe as a whole).[3]

An *empirical* symmetry is an altogether different sort of thing: one has an empirical symmetry only if one has two *physically* distinct states of the world, which (however) cannot be discriminated from one another by means of measurements confined to some specified subsystem. [4] An example of an empirical

---

[3]This view commands widespread if perhaps not universal agreement in physics and philosophy of physics; in this paper we assume it without further discussion. (For a positive case, see Saunders (2003).) Those unsympathetic to the view, however, can replace 'represent the same possible world' in our definitiion of 'theoretical symmetry' with 'represent an experimentally indistinguishable possible world'; *mutatis mutandis*, the rest of our discussion will go through unimpeded.

[4]Healey proceeds at this point by introducing talk of sets of 'situations' (to be thought of as possible physical states of affairs in some proper part of the universe, and sharply contrasted with mathematical models of theories), and identifying empirical symmetries with

symmetry is the boost of Galileo's (physical) ship by a given uniform velocity **v**. This empirical symmetry 'corresponds to' the theoretical boost invariance of Newtonian mechanics in the sense that it follows from the assumptions (i) that Newtonian mechanics adequately describes the Galileo ship scenario, (ii) that Newtonian mechanics is boost invariant and (iii) that the ship's cabin is effectively dynamically isolated[5] from the environment relative to which it is boosted, that experiments confined to the cabin cannot distinguish between the stationary and moving states of the ship.

2.1   Faraday's cage

Healey supports the conventional wisdom that 'local' symmetries cannot have direct empirical significance in this sense: that 'there can be no analogue of the Galilean ship experiment for local gauge transformations' (Brading and Brown 2004, p.657). If this is the case, his puzzle is that the following case of *Faraday's cage* seems, *prima facie* at least, to *be* a perfect analog in all relevant respects.

Faraday's observation was that, if one constructs a cage from conducting material and places a static charge on the walls of the cage, then experiments confined to the interior of the cage cannot detect the presence or magnitude of that charge. Physically, this is because, in the limit in which the material is *perfectly* conducting, charge placed on its exterior will redistribute itself under the influence of electrostatic repulsion, and the equilibrium distribution is such that the net electric field at any point inside the cage — summing contributions from from the charges everywhere on the surrounding wall — is zero.

Constructing such a cage and performing the experiment, Faraday writes

> I went into this [charged] cube and lived in it, but though I used lighted candles, electrometers, and all other tests of electrical states, I could not find the least influence upon them. (Maxwell 1881, p.53)

The puzzle arises because, in electromagnetic theory, charging the cage has the effect of *raising its electric potential* — but shifts of electric potential are elements of a *local* symmetry group of electromagnetism. Hence Healey's puzzle: It *seems* that the empirical symmetry between Faraday's charged and uncharged cages is a consequence of the theoretical gauge symmetry of electromagnetism, in precisely the same way as that in which the empirical symmetry between Galileo's stationary and moving ships is a consequence of the boost invariance of Newtonian mechanics. The challenge to an advocate of the conventional wisdom is therefore: 'why doesn't Faraday's cube provide a *perfect* analogue of Galileo's ship for a class of local gauge transformations of classical electromagnetism?'(Healey 2009, p.701; emphasis in original).

---

automorphisms of such sets of situations. We won't need this development.

[5]"Dynamical isolation" is of course approximate in practice, since no proper subsystem can in practice be *perfectly isolated*. It is also relative to the observation capabilities of relevant observers (if Galileo had included a GPS tracker in his list of the cabin's accoutrements, things would have been rather different).

Healey's own view is that, the striking *prima facie* analogy between Faraday's cages and Galileo's ships notwithstanding, there is in fact an important disanalogy between the two, and (furthermore) it is a disanalogy that upholds the conventional wisdom: he concludes from it that '[l]ocal gauge symmetry is *merely* a theoretical symmetry of classical electromagnetism; and [implies] no corresponding empirical gauge symmetry' (Healey 2009, p.710; emphasis added). We disagree: there *is* a disanalogy between Galileo's ship and Faraday's cage that is of some relevance to the empirical status of the theoretical symmetries involved, but it is not the disanalogy that Healey picks out. Further, we will argue, the relevant disanalogy that there *is* has nothing to do with the global/local distinction, and does not imply that the gauge symmetry of electromagnetism has less (or less direct) empirical significance than the boost invariance of Newtonian mechanics.[6]

2.2 t'Hooft's beam-splitter

A similar puzzle is considered by Brading and Brown (2004), in discussion of t'Hooft (1980). They consider the following variant on the two-slit experiment. A beam (say, of electrons) is split in two. The two halves of the beam follow disjoint paths until they are recombined some time later. The resulting interference pattern is observed by having the beam stopped by a fluorescent screen. On each run of the experiment, the experimenter may or may not insert a 180 degree phase-shifter into the upper, but not the lower, path. It is observed that the interference pattern depends nontrivially on whether or not the phase-shifter is present.

Our interest in this setup[7] arises from the idea that one can treat the upper half of the beam as subsystem (analogous to Galileo's ship), and the lower half as environment (analogous to Galileo's shore). We know that by performing experiments on the upper half of the beam (subsystem) *alone*, we would not be able to detect the presence or absence of the phase-shifter; it is only when we allow the upper half to interact with the lower half (environment) that we can detect whether or not the phase-shift has been performed.

But this setup can be modelled in a theory that represents the state of the electron via a complex scalar field, ψ, coupled to an electromagnetic field that

---

[6]The present paper is devoted to laying out our preferred framework for thinking about the empirical significance of symmetries, and drawing out its consequences for various cases. If our analysis is correct, it follows that Healey's must be incorrect, but in this paper we do not attempt to say where.

[7]This was not t'Hooft's interest. t'Hooft introduced the setup as a purported empirical illustration of the fact that 'gauge transformations' on the wavefunction alone—transformations of the form

$$\psi'(x) = e^{i\theta(x)}\psi(x) \qquad (1)$$
$$A'_\mu(x) = A_\mu(x) \qquad (2)$$

—are not symmetries of the theory. That it is an empirical illustration of that fact is contested by Brading and Brown.

is represented by a four-potential A_μ (see section 8).[8] The following 'gauge transformations' are symmetries of the theory:

$$
\begin{aligned}
\psi &\mapsto e^{i\chi}\psi, \\
A_\mu &\mapsto A_\mu + \partial_\mu\chi,
\end{aligned}
\tag{3}
$$

where, again, χ is an arbitrary smooth function on spacetime (but now the transformation must modify *both* ψ and A_μ in the manner stated, with common χ, in order to be a symmetry).

But 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). Hence Healey's question arises again for this case: Why doesn't the modified two-slit experiment provide a *perfect* analog of the Galileo-ship thought experiment, vis-a-vis the notion of direct empirical significance, for local gauge symmetry?

Brading and Brown agree with Healey: the analogy is superficial, and the empirical result is not — despite appearences — a direct consequence of the theoretical symmetry. In this case as in the last, we disagree: indeed, our view is that t'Hooft's beam-splitter, even more so than Faraday's cage, is a *direct* analog of Galileo's ship. To defend this view requires us to lay out our preferred framework for thinking about the empirical consequences of symmetries; it is to this task that we now turn.

3 A framework for symmetries I — systems and subsystems

If the observable consequences of a theory's symmetry transformations arise from performing those transformations on subsystems of a system, one might think that a careful analysis of this process ought to start by giving explicit definitions of 'subsystem' and 'environment'. The language of differential geometry, in particular, seems especially well suited to this ("given a model $\langle \mathcal{M}, g, \phi \rangle$ of the theory, a subsystem of that model is an open subset of $\mathcal{M}$ together with the restrictions of g and ϕ to that subset, such that ..."). But this sort of explicit move falls foul of the open-ended variety of subsystems that physics provides us with. The stream of neutrinos passing through your body at this moment, for instance, is a system interacting only very weakly with you despite its spatial coincidence: a fact which allows us (in principle!) to demonstrate the Lorentz symmetry group empirically by performing a Lorentz boost on you but not the neutrinos. A single electron can with profit be treated as comprising two systems — specified by the spin and spatial degrees of freedom — which in

---

[8]The complex scalar field is generally taken to be a one-particle wavefunction, and may be so taken in our discussion. However, the empirical results discussed would also occur in a possible world containing a classical charged field, in which our analysis would also apply; one of us has argued (Wallace 2009a) that for this reason there is nothing essentially quantum about the AB effect.

many circumstances can be treated as effectively isolated. And in unitary quantum mechanics, 'system' and 'environment' become branch-relative concepts in a sense that is not easily characterised in terms of spacetime regions.

For this reason, we choose instead to adopt a more axiomatic approach: we define a framework into which all the examples above (and, hopefully, any others) can fit, and we take this framework *implicitly* to define 'subsystem' and 'environment' for present purposes. We suppose firstly that there exists a set $\mathcal{S}$ of states of the subsystem, and another set $\mathcal{E}$ of states of the environment. These 'states' are intended to be mathematical objects: the mathematical entities representing the physical state of system and environment, not those physical states themselves. In a field-theoretic context, for instance, states of $\mathcal{S}$ might be specified by the field values on some subset $\mathcal{N}$ of the theory's underlying manifold $\mathcal{M}$, and states of $\mathcal{E}$ would be specified by the field values on the rest of the manifold.

It would be natural to define states of the total system as ordered pairs $\langle s, e \rangle$ of states $s \in \mathcal{S}$ and $e \in \mathcal{E}$, and we shall do just this. But there are three important caveats.

Firstly, in doing so we make the assumption that knowing the state of the subsystem and its environment suffices to specify the state of the total system. There is a *prima facie* worry that this assumption will exclude such phenomena as quantum entanglement between 'subsystem' and 'environment'. However, this worry is misguided: our assumption is one about the mathematical formalism with which we describe the theory and not about the theory itself. For example, it is true for Yang-Mills gauge theories in the connection formalism but not in the holonomy formalism; it is true for quantum mechanics in (Deutsch and Hayden's construal of) the Heisenberg picture but not in the Schrödinger picture.[9]

Secondly, the idea that subsystem and environment can be treated as distinct and isolated — so that $\langle s, e \rangle$ and $\langle s', e \rangle$, can both be treated as possible states of the total system — is at best an idealisation. Forces in the physical universe are not normally so cooperative as *entirely* to screen off any system from another; the gravitational field of an arbitrary planet in Andromeda has a non-zero dynamical effect on your body. But in a great many situations of physical significance, the idealisation is highly accurate.

Finally, and of most direct importance for this paper, note that one should not assume that just any old pair $\langle s, e \rangle$ is a possible state of the total system. In general, there are boundary conditions which must be satisfied in order for a given subsystem-environment pair to be a solution of the theory's dynamical equations. In the case where the total system is a field theory, these will be boundary conditions in the differential-equation sense: requirements that the fields and their derivatives match up on the boundary. In other cases they may take a different form: in Heisenberg-picture quantum mechanics, for instance,

---

[9]For more information on the different formalisms for Yang-Mills theories, see Healey (2007) and references therein; for discussions of locality in different formulations of quantum theory, see Deutsch and Hayden (2000) and Wallace and Timpson (2007).

the boundary condition is the requirement that the operators of system and environment commute.

For these reasons, we represent the states of the total system by a set $\mathcal{U}$ which is a *subset* of the Cartesian product $\mathcal{S} \times \mathcal{E}$ (which subset will depend on the details of the theory under study). We write $s * e$ to be the combination of s and e — thus, $s * e$ is equal to $\langle s, e \rangle$ when $\langle s, e \rangle \in \mathcal{U}$ and is undefined otherwise.

In this abstract formalism, the boundary conditions themselves are best represented as sets of states. Given a subsystem state s, the boundary condition B_s of that state is the set of all $e \in \mathcal{E}$ such that $s * e$ is defined. And conversely, the boundary condition C_e of any *environment* state e is the set of all $s \in \mathcal{S}$ such that $s * e$ is defined.

What of the symmetries? In keeping with the principle that (mathematically, at least) symmetries are defined in the first place for the total system, we suppose that the *theoretical symmetries* of the theory are represented by a set Σ of 1:1 maps from $\mathcal{U}$ onto itself, closed under inverses and composition. We require that these maps restrict uniquely to well-defined 1:1 maps of $\mathcal{S}$ and $\mathcal{E}$ onto themselves: this amounts to the requirement 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 $\sigma|_E$. The theoretical 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.

4   An example: Coulombic electrostatics

At this point, it will be helpful to see how these formal ideas play out in a concrete example. Over the course of this paper we will develop several such examples, but our first — and simplest — is *Coulombic electrostatics*, the theory of charged particles interacting electrostatically (applicable in the real world in a low-velocity classical regime where radiation and magnetic fields can be neglected).

To be precise, we wish to consider the theory of some set of charged particles moving in a static potential field, together with (optionally) some background conducting surfaces whose evolution in time is fixed *ab initio*. The dynamics are given by the Coulomb force law

$$\ddot{\mathbf{x}}_i = -\frac{q_i}{m_i} \nabla \Phi(\mathbf{x}_i) \tag{4}$$

(where $\mathbf{x}_i, q_i$, and m_i are the position, charge and mass of the ith particle), by the Poisson equation

$$\nabla^2 \Phi(\mathbf{x}) = 4\pi \rho(\mathbf{x}) = 4\pi \sum_i q_i \delta(\mathbf{x} - \mathbf{x}_i) \tag{5}$$

and, in the presence of background conductors, by the boundary condition that $\nabla \Phi$ is normal to — and so Φ is constant on — any conducting surface, and by some elastic collision law whereby charged particles bounce elastically off surfaces. (Readers allergic to delta functions should give each particle a small finite radius and adjust (5) accordingly.) We assume that not all particles have

the same charge-to-mass ratio (the significance of this assumption will become clear in section 6).

What are the total-system symmetries of the theory? There is an internal (that is, not spacetime) symmetry

$$\Phi(\mathbf{x}, t) \longrightarrow \Phi(\mathbf{x}, t) + f(t); \quad \mathbf{x}_i \longrightarrow \mathbf{x}_i. \tag{6}$$

In the absence of background conductors, spatial and temporal translations, velocity boosts, and rotations are also all symmetries. Background conductors, in general, break this symmetry.

There are two natural candidates for the role of isolated subsystem. The first is any collection of charged particles located far from the remaining charged particles and from the conducting surfaces. Even if the system and/or environment have a net charge, the effects of the charge of each on the other will be negligible due to the great distance between them. The second is any collection of charged particles located entirely within some conducting surface. The boundary conditions, in either case, are the same: the subsystem and environment potentials must match up on the spatial boundary of the subsystem.

There are two ways to formalise the subsystem/environment split. First, suppose we are interested only in internal symmetries (in particular, as we will be when considering subsystems that are isolated by being contained within a closed conducting surface). In that case, if (intuitively) our subsystem is localised within region W, and contains p particles, then we will specify states of the subsystem by a potential function on W (at all times) together with the trajectories of the particles in the specified set P ($|P| = p$), all satisfying the Coulomb law and the Poisson equation; and the state of the environment is given by $N - p$ trajectories in, and a potential function on, $R^4 - W$ (where N is the total number of particles in the universe). States of the total system are given by superimposing the two; provided the two potential functions agree on the boundary of W, the result will be a solution to the equations of motion of the total system.

Things are slightly messier when we wish to consider spacetime symmetries in conjunction with a subsystem/environment split that is itself spatiotemporal in character, since then, in general, the transformation we wish to consider will not leave invariant the 'subsystem region' $W \subset \mathbb{R}^4$. To deal with this, we will represent the subsystem by the trajectories of the particles in the set P within, and one potential function defined on, the whole of the spacetime; the environment will be represented similarly. Still, we take some particular region $W \subset \mathbb{R}^4$ to be part of the characterisation of the subsystem/environment split. For given states of subsystem and environment, the corresponding state of the universe is given by superimposing the trajectories of the subsystem and environment particles, and setting the potential function equal to the subsystem function within W and to the environment function outside W. If the particles of the subsystem are contained within W and far from its boundary, and the particles of the environment are contained within $\mathbb{R}^4 - W$ and again are far from its boundary, then the potential functions of both system and environment

10

will be very close to constant on the boundary of W. As such, the result of combining the subsystem and environment states as above will be a solution of the dynamical equations whenever the potential functions take the same constant value on the boundary of W.[10]

5 A framework for symmetries II — the relationship between theoretical and empirical symmetries

We now return to the general problem of constructing a framework in which to study the notion of 'direct empirical significance'—that is, to study the relationship between theoretical and empirical symmetries. Although we have defined (theoretical) symmetries in terms of the total system, we are more directly interested, for the purposes of this task, in the symmetries of the subsystem $\mathcal{S}$. We will find it useful to distinguish two classes of subsystem symmetries. As we observed at the start of this paper, a subsystem symmetry can have direct empirical significance only if, roughly, the operation of acting on the subsystem with that subsystem symmetry and 'leaving the environment alone' is not a symmetry of the whole system. Accordingly, we first define:

Interior symmetries. A subsystem symmetry $\sigma_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ is *interior* iff $\sigma_{\mathcal{S}} * \mathbb{I}$ is a symmetry of the whole system: that is, iff for all subsystem states $s \in \mathcal{S}$ and all environment states $e \in \mathcal{E}$ such that $s * \mathcal{E}$ is defined,

1. the state $\sigma_{\mathcal{S}}(s) * e$ is also defined, and
2. $s * e$ and $\sigma_{\mathcal{S}}(s) * e$ represent the same possible world as one another.

We write $I_{\mathcal{S}}$ for the group of interior symmetries of a given subsystem $\mathcal{S}$.

Examples of spacetime symmetries that are interior include the diffeomorphisms used in the Hole Argument[11] (treating the 'hole' as the subsystem region W). Examples of internal symmetries that are interior include gauge transformations which coincide with the identity on a neighbourhood of the boundary of the subsystem region. (Note that the only sense in which a given *universe* symmetry can be said to be 'interior' is subsystem-relative (viz. that its restriction to some specified subsystem is interior in the sense of the above definition); although some universe symmetries — notably the nontrivial global symmetries — are not interior relative to any subsystem.)

Since performing an interior symmetry transformation on the subsystem state and leaving the environment state alone results in a redescription of the

---

[10]There are other symmetries for whose analysis even this more complicated definition would be inadequate. If we were interested in *permutation invariance*, for instance, our practice of taking the subsystem to be specified in part by *some particular collection of particles* would be inadvisable. Plausibly, for any (informally) given subsystem/environment split, there is always a most general way of formalising that split, adequate to dealing with all symmetries at once, and coming at the cost of increased complication.

[11]See Norton (2008) and references therein.

same possible world, such a subsystem transformation does not lead to a *distinct* situation, hence no (nontrivial) empirical symmetry is associated with such transformations.[12]

We now define a collection of subgroups of the group of subsystem symmetries, each of which includes the interior transformations: the group of those subsystem symmetries that 'leave the boundary conditions alone' when acting on some specified subsystem state $s \in \mathcal{S}$.

Boundary-preserving symmetries. A subsystem symmetry $\sigma_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ *preserves the boundary of $s \in \mathcal{S}$* iff $s, \sigma(s)$ are elements of the same boundary condition as one another: that is, iff for all $e \in \mathcal{E}$ and all $s * e \in \mathcal{U}$, we have $\sigma_{\mathcal{S}}(s) * e \in \mathcal{U}$.

In general, we expect—although we do not assume—that a subsystem symmetry will be boundary-preserving *on all states* just in case it is an interior symmetry. (For example, in the case of local internal symmetries in a field theory, the only way for a subsystem symmetry $\sigma_{\mathcal{S}}$ to be boundary-preserving on all states is for it to coincide with the identity transformation in a neighbourhood of the subsystem boundary, in which case it is the restriction of the symmetry $\sigma_{\mathcal{S}} * \mathbb{I}$ to the subsystem $\mathcal{S}$, and hence is an interior symmetry.)

The more interesting case is that of subsystem symmetries *that are boundary-preserving on some set of subsystem states that arise when the subsystem is isolated from the environment, and in general not otherwise.* For an example of this phenomenon, consider again the case of boosting a spatially isolated collection P of particles. Recall (from section 4) that we take a state of the subsystem to be given by the trajectories of the particles in P and an associated potential function (viz. that component of the total potential function that is due to the particles in P), all defined on the whole of spacetime. The motivation for picking out the particular particle collection P as a subsystem is that *we are considering* universe states with the feature that P is spatially isolated (i.e. with the feature that the particles in P are contained within some region W and far from its boundary, while the particles of the environment are contained within $\mathbb{R}^4 - W$ and again are far from its boundary). The boundary conditions are then given by matching the subsystem and environment potential functions on the boundary of W. For those states in which the conditions that motivated picking out P as a subsystem in the first place hold, in the vicinity of the boundary there are no trajectories and potentials are flat, and hence Poincaré transformations that are 'small' enough to preserve the fact that the particles are far from the boundary of W will preserve the boundary conditions. But it is a merely contingent fact that this collection of particles is isolated: having thus defined our subsystem $\mathcal{S}$, there are subsystem states $s \in \mathcal{S}$ according to which the particles are very close to ∂W, and the boundary conditions of *those* states will not be preserved even by 'small' Poincare transformations.

---

[12]In the terminology of Kosso (2000): the performance of an interior symmetry transformation satisfies the Invariance Condition, but not the Transformation Condition.

Many of the notions we are using, of course, apply only approximately in practice. For example, no collection of particles is ever *entirely* isolated from its environment: however far we place the environment particles from the subsystem boundary, the environmental potential function on that boundary will not be *exactly* the same as it would have been had the environment particles not existed. The salient aspect of this failure of exact isolation for our project is the following: symmetries will rarely if ever *exactly* preserve the boundary conditions of our subsystem states of interest, and (hence) if the subsystem state s and environment state e jointly solve the equations of motion, in general $\sigma_{\mathcal{S}}(s)$ and e strictly speaking will not. However, if the subsystem is approximately isolated from its environment, then $\sigma_{\mathcal{S}}$ will *approximately* preserve the boundary of s, in the sense that there is another state $s' \in \mathcal{S}$ 'close' to $\sigma_{\mathcal{S}}(s)$, whose boundary condition is the same as that of s. In such circumstances we will simply write $\sigma_{\mathcal{S}}(s) * e \in \mathcal{U}$, understanding $\sigma_{\mathcal{S}}(s) * e$ to denote the result of combining s' and e.[13]

The question now is: when do subsystem symmetries have direct empirical significance? Well: as we noted at the outset, this should occur when there is some transformation of the universe that is not a symmetry of the universe, but whose restriction to the subsystem *is* a symmetry of that subsystem. Accordingly, in the terminology just defined:

1. Interior symmetries cannot have direct empirical significance, since combining an interior symmetry of the subsystem with the identity on the environment gives a universe symmetry transformation.

2. Non-interior symmetries that are boundary-preserving on some subset $\mathcal{P} \subseteq \mathcal{S}$ of states of the subsystem can have direct empirical significance when performed on subsystem states in that subset $\mathcal{P}$: they are the restrictions of (universe) symmetries to the subsystem $\mathcal{S}$, and combining them with the identity transformation on the environment gives a transformation that preserves the satisfaction of the dynamical equations (by boundary-preserving-ness), but need not be a (universe) symmetry transformation (by non-interiority). The empirical symmetry 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 (for example, the relative position, orientation or velocity) that are changed by the transformation.[14]

3. Non-interior symmetries that are *not* boundary-preserving on the subsystem states of interest can have direct empirical significance, but, in

---

[13]In order to conform to the requirement that subsystem and environment states determine total-system state uniquely, we will have to suppose that some particular choice of which 'close' solution to pick has been specified, but the details of how this is done will not be of physical consequence.

[14]In the terminology of Kosso (2000), case (1) corresponds to the case where the Transformation condition cannot be met, case (2) to the case where the Transformation and Invariance conditions are both met.

order to realise this significance, the environment state must be altered. Firstly, in order to generate a dynamically possible state of the universe at all, we require some transformation $e \mapsto e'$ of the environment with the feature that $e' \in B_{\sigma(s)}$. There is always, of course, a 'cheap' way of finding such an e': simply perform $\sigma_{\mathcal{E}}$ on e, where $\sigma_{\mathcal{E}}$ is the restriction to the environment of the *same* universe symmetry transformation whose restriction to the subsystem we are performing on s. This cheap trick, however, fails to generate a distinct physical state of the universe, and hence does not lead to a nontrivial empirical realisation of the subsystem symmetry in question. We can find a dynamically possible state of the universe whose restriction to the subsystem is $\sigma_{\mathcal{S}}(s)$ *but that represents a distinct possible world from* $s * e$ if we can find an environment state e' that differs physically from e (i.e., in general, whose difference from e could be detected by means of measurements performed on the environment alone), while satisfying the condition $e' \in B_{\sigma(s)}$. In this case, the universe states $s * e, \sigma_{\mathcal{S}}(s) * e'$ will be physically distinct, but the difference will not be detectable by means of measurements confined to the subsystem — that is, we will have a nontrivial empirical symmetry associated with the non-boundary-preserving theoretical subsystem symmetry $\sigma_{\mathcal{S}}$.[15]

In the case of Coulombic electrostatics, for instance:

1. There are no interior symmetries (except the trivial case: the identity).

2. In the absence of background conducting surfaces, the spacetime symmetries — the boosts, rotations and translations — are all non-interior symmetries that are boundary-preserving when the system is spatially isolated from the environment. These correspond to Galileo-ship type empirical symmetries.

3. The potential shifts are non-interior symmetries but are not boundary-preserving on any states. The associated empirical symmetry is that of Faraday's cage; the intrinsic change to the environment state that we implement is the placement of an electric charge on the exterior of the cage.[16]

---

[15]There is another cheap trick that avoids the problem mentioned above: first, perform $\sigma_{\mathcal{E}}$ on e; second, make some small and irrelevant, but intrinsic, further change to the state of the environment. The task of demarcating transformations $e \mapsto e'$ that appear to correspond in some interesting way to the theoretical symmetry of interest, on the one hand — such as, arguably, the physical implementation of a potential shift by placing a charge on the exterior of Faraday's cage — and implementations of this modified cheap trick, on the other, lies beyond the scope of this paper. It is the project of stating when a given empirical symmetry with the above structure really deserves to be counted as 'associated to' the non-boundary-preserving theoretical subsystem symmetry in question.

[16]Healey (2009) agrees that the Faraday cage empirical symmetry would correspond to potential shifts *in a world that was exactly described by electrostatics*. Hence, *so far*, our treatment does not disagree with his. The disagreement arises because we hold that the Faraday-cage effect *still* corresponds to the potential shifts once the underlying theory is changed from electrostatics (in which the potential shifts are a global symmetry) to electromagnetism (in which they are local). We discuss this latter case in section 7 below.

Actually, it is potentially a little misleading to speak of *a given non-interior symmetry* having direct empirical significance. The reason is that, for any subsystem symmetry σ of $\mathcal{S}$, any environment transformation $\rho : \mathcal{E} \to \mathcal{E}$ [17] and any *interior* subsystem symmetry j of $\mathcal{S}$, 'performing $\sigma * \rho$' and 'performing $j \cdot \sigma * \rho$' amount physically to the same thing as one another. So there is no empirical significance to the fact that it is σ *as opposed to, say,* $j \cdot \sigma$ that is performed on the subsystem. This motivates the following definition: we say that two subsystem symmetries σ_1, σ_2 are *equivalent* iff they differ by an interior symmetry (that is, iff there exists $j \in I$ such that $\sigma_1 = j \cdot \sigma_2$). Empirical symmetries correspond $1 : 1$, not to non-interior (theoretical) symmetries themselves, but to *equivalence classes of non-interior symmetries* under this equivalence relation: that is, to elements of the quotient group $\Sigma_{\mathcal{S}}/\mathcal{I}_{\mathcal{S}}$. We will, however, continue to write of individual theoretical symmetries 'having direct empirical significance' where confusion will not result.

6   Newtonian gravity

We now begin illustrating these ideas for theories other than Coulombic electrostatics. Our first example is another theory with only global symmetries, and actually it looks pretty similar to electrostatics: it is the theory of some set of massive particles moving in a Newtonian gravitational field, so that the dynamics are given by Newton's law

$$\ddot{\mathbf{x}}_i = -\nabla\Phi(\mathbf{x}_i), \tag{7}$$

where $\mathbf{x}_i$ is the position of the ith particle, and by the Poisson equation

$$\nabla^2\Phi(\mathbf{x}) = 4\pi\rho(\mathbf{x}) = 4\pi\sum_i m_i\delta(\mathbf{x} - \mathbf{x}_i). \tag{8}$$

Note that these equations are identical to the equations (4)–(5) of electrostatics, *except that* in the gravitational case, the 'charge-to-mass ratio' is always unity, hence is the same for all particles. This difference, as is well known, leads to interesting and important effects. The point of the present section is to show how such effects are predicted in our framework; the case also serves as a second example of a global symmetry that is non-boundary-preserving, and hence gives rise to an empirical symmetry that is not purely relational.

As in Coulombic electrostatics, we will take the subsystem to correspond to some fixed spatial region W, and we will take a subsystem state to be specified by the trajectories of some (fixed) subset of particles together with a potential function over all space, such that they jointly satisfy the dynamical equations; an environment state is specified by the trajectories of the residual particles,

---

[17]Not necessarily an environment symmetry. In the cases we are interested in, ρ is *either* the identity transformation on $\mathcal{E}$ (hence trivially an environment symmetry), or is a transformation whose effect is to tweak the environment state just enough to render it compatible with the transformed subsystem state (for example, by placing a charge on the outside of the cage).

together with another potential over all space. A given subsystem state is compatible with a given environment state if (a) the subsystem particle trajectories are confined to the subsystem region W; (b) the environment particle trajectories are confined to the complement of W; (c) the potential functions agree on the boundary of W.[18]

In particular, if the subsystem particles are at all times inside W and sufficiently far removed from the environment particles then subsystem and environment will be compatible iff their potential functions have the same behaviour at spatial infinity (since, in that case, their behaviour on the boundary of W is given by their behaviour at spatial infinity).

The symmetries of electrostatics—Galilean transforms and time-dependent, spatially independent potential shifts—are also symmetries of Newtonian gravitation. The Galilean transforms lead to exactly the same, purely relational, empirical symmetry in both cases. In principle, the potential shifts lead to non-relational empirical symmetries, but in practice, the absence of conducting surfaces makes this of little interest.[19]

There is, however, a third group of symmetries, arising from the above mentioned difference between the equations of electrostatics and those of Newtonian gravity. This corresponds to applying a spatially independent but otherwise arbitrary acceleration $\mathbf{k}(t)$ to all the particles, beginning at some time t_0. In closed form, this takes a particle at point $\mathbf{x}(t)$ at time t to

$$\mathbf{x}'(t) = \mathbf{x}(t) + \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{0}^{\tau_1} \mathrm{d}\tau_2\, \mathbf{k}(\tau_2). \tag{9}$$

If in addition to this transformation of the particle positions we also replace Φ with Φ', where

$$\Phi'(\mathbf{x}', t) = \Phi(\mathbf{x}, t) + \mathbf{k}(t) \cdot \mathbf{x}, \tag{10}$$

then the resultant system still satisfies the equations of motion; the transformation is a symmetry.

What, if anything, is the empirical significance of this third group of symmetries? Restricted to any subsystem, it is not boundary-preserving on any state, so it leads to no purely relational empirical symmetries. As in the case of potential shifts in electrostatics, however, failure to preserve boundary conditions will not prevent our symmetry from having empirical realisations. We merely have to find distinct states e, e' of the environment, such that $s * e \in \mathcal{U}$ and such that the boundary conditions of e and e' are related to one another by the subsystem symmetry in question. The transformation $s * e \mapsto \sigma_{\mathcal{S}}(s) * e'$ will then be an empirical realisation of our symmetry; if e, e' represent intrinsically distinct physical states of the environment then this transformation will lead to a representation of a distinct possible world, but it follows from the fact that

---

[18]As in our discussion of Coulombic electrostatics, in practice these conditions will only approximately be met.

[19]One counter-example is the empirical symmetry between a system of particles far out in space and the same system in the interior of a hollow sphere of massive particles: as is well known, the gravitational potential of such a sphere is constant within the sphere.

σ_S is a subsystem symmetry that the difference will not be detectable within the subsystem.

It is easy to find such e, e' in the case in which the subsystem and environment are spatially isolated from one another, i.e. the subsystem and environment particles are both far from the boundary of the 'subsystem region' W. The boundary conditions are the requirements that the values and derivatives of the subsystem and environment gravitational potentials match on the boundary ∂W of the subsystem region; thus, since the untransformed and transformed subsystem potentials Φ_s, Φ'_s are related as in (10), the necessary and sufficient condition for matching is that the untransformed and transformed environment potentials Φ_e, Φ'_e are also related by

$$\Phi'_e(\mathbf{x}, t) = \Phi_e(\mathbf{x}, t) + \mathbf{k} \cdot \mathbf{x}. \tag{11}$$

That is, the condition is that the transformed state of the environment differs from the untransformed state by the addition of a uniform gravitational field of strength $\mathbf{k}$ on ∂W.

The prediction, in other words, is that in Newtonian gravity, a system floating freely in space behaves identically, with respect to its internal processes, to a system freely falling in a uniform gravitational field. The empirical symmetry associated to the symmetry (10) is Einstein's elevator, the thought experiment which led Einstein to the equivalence principle.

It is worth emphasizing that the conditions required for realisation of this empirical symmetry are easy to find in nature; sophisticated technology is not required. Since the gravitational field is smooth, it will count as "approximately uniform" in *any* given spacetime region, for the purposes of a sufficiently small subsystem (for example, a human being jumping from a diving board in the Earth's gravitational field). More formally, the condition is realised when, in some solution to the dynamical equations, there is some spacetime tube of width L for which

1. At any time t, the potential within the tube has sufficiently small second derivative to be approximated as

$$\Phi(\mathbf{x}, t) \simeq \Phi(\mathbf{x}_0, t) + \nabla\Phi(\mathbf{x}_0, t) \cdot (\mathbf{x} - \mathbf{x}_0) \tag{12}$$

2. The centre of the tube is accelerating with acceleration, at time t, equal to $-\nabla\Phi(\mathbf{x}_0, t)$.

This in turn will occur whenever the tube is a width-L tube surrounding the trajectory of a test particle moving freely under gravity and the tidal effects in this tube can be neglected (i. e. when any significant massive bodies are distance $\gg L$ from the tube). In this circumstance, the boundary conditions of the tube match those of the subsystem up to an overall acceleration, so there will be some symmetry transformation of the subsystem which allows it to be superimposed into the tube.

7 Local symmetries that are not boundary-preserving: Classical electromagnetism and Faraday's cage

We now wish to consider theories with local symmetries: that is, roughly speaking, symmetries specified by functions on spacetime, rather than by global parameters. (We shall have no need for a more precise definition of "local".[20]) One of the simplest examples of such a theory is classical electrodynamics.

States in electrodynamics are given by the trajectories x_i of a collection of charged particles (with the ith particle having charge q_i), and by an electromagnetic four-potential, A_μ, defined on spacetime. The dynamics are given by Maxwell's equations and by the Lorentz force law.

Recall that this theory is the locus of our original disagreement with Healey: all parties agree that the Faraday Cage phenomenon corresponds to global potential shift in Coulombic electro*statics*, but we and not Healey think this continues to hold once the potential-shift symmetry is 'localised'.

Our treatment of this case should, by now, be rather easily predictable. As before, we first list the theory's symmetries. There are global translations, rotations, and Galilean boosts. In contrast to electrostatics and Newtonian gravity, however, the theory of electrodynamics also has a *local* symmetry: the group of 'gauge transformations' of the form

$$A_\mu \to A_\mu + \partial_\mu \chi, \tag{13}$$

where χ is an arbitrary smooth real-valued function.

We next consider what might count as an 'isolated subsystem' in this theory. The natural answers, as in the case of Coulombic electrostatics, are:

- Any collection of charged particles located far from the remaining charged particles and from any conducting surfaces (in which case the subsystem region W is taken to be some region such that, in the states we are interested in, our subsystem particles' trajectories are contained well within, and the environment particles' trajectories are well outside, W).

- Any collection of charged particles located entirely within some conducting surface (in which case W is taken to be the region strictly enclosed by the conducting surface).

Again as in the case of electrostatics, the details of how we represent subsystem and environment states depend on whether our interest is in spacetime symmetries or only in interior symmetries:

---

[20]The defender of the conventional wisdom, of course, must believe that some more precise account can be given, since *she* believes that the global/local distinction has some fundamental importance for the notion of direct empirical significance. On *our* view, on the other hand, it would not be surprising if nothing more than a rough-and-ready characterisation of the global/local distinction could be given. As it turns out, though, our analysis will suggest one: cf. footnote 21.

- If we wish to discuss spacetime symmetries, we specify subsystem and environment states by giving the relevant particles' trajectories in, and a four-potential function A_μ on, the whole of the spacetime; for given states (in this sense) of the subsystem and environment, the corresponding state of the universe (if it exists) is obtained by superimposing the trajectories of the subsystem and environment particles, and by grafting the environment potential outside W and the subsystem potential inside W.

- If our interest is only in internal symmetries (as, in the present section, it is), a simpler definition suffices: again as in section 4, we define the subsystem four-potential only on W, and the environment potential only on $M - W$.

This is the first theory we have studied which has nontrivial interior symmetries: if a smooth function $\chi_S : A \to \mathbb{R}$ vanishes on a neighborhood of the boundary of W, then it defines a gauge transformation which is an interior symmetry of the subsystem: it is the restriction to the subsystem of the universe symmetry transformation $\chi : M \to \mathbb{R}$ defined by

$$\chi(x) = \begin{cases} \chi_S(x), & x \in W \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

But not all gauge symmetries of the subsystem are interior symmetries: indeed, if χ_S does not vanish on the boundary, then defining χ in this way leads to a discontinuity on the boundary of W, which violates the requirement that χ is a smooth function. This is not mere mathematical pedantry: a discontinuity in χ generates a delta-function potential on the boundary, and so (obviously!) does not leave the environment state unchanged.

For the purposes of identifying and classifying the empirical symmetries in physical situations well-described by this theory, we need to identify which non-interior (theoretical) subsystem symmetries are and are not boundary-preserving, relative to subsystem states that satisfy the conditions for the subsystem to be isolated. The answers are that

- the Poincaré transformations, as usual, are boundary-preserving on states satisfying the conditions for spatial isolation, and give rise to Galileo-ship type empirical symmetries;

- gauge transformations (except those that vanish in a neighbourhood of the subsystem boundary, hence are 'interior' and have no empirical significance) are non-boundary-preserving; those whose restriction to the subsystem boundary ∂A coincide with (the restrictions of) constant potential shifts correspond to the Faraday cage empirical symmetry (where, as in our discussion of electrostatics, the environment state must be intrinsically changed in order to obtain a nontrivial empirical realisation of the symmetry).

The shift from the global 'gauge' group of electrostatics to the local group of electrodynamics thus has no great effect on the empirical status of the potential shifts: its only effect is that, in the case of the local group, it is strictly speaking the equivalence classes of gauge transformations that tend to the same constant value as one another on the subsystem boundary (elements of $\Sigma_{\mathcal{S}}/\mathcal{I}_{\mathcal{S}}$), rather than the individual gauge transformations themselves (elements of $\Sigma_{\mathcal{S}}$), that have empirical significance.[21]

Our conclusion is thus (*pace* Healey) that the Faraday cage effect *is* a direct consequence of the local theoretical gauge symmetry of classical electromagnetism in just the same way that other empirical symmetries are direct consequences of global theoretical symmetries, and in just the same way that Faraday's cage *itself* is a consequence of the global gauge symmetry of electrostatics. Our argument is simply that the general framework that we think correctly captures the phenomenon of 'direct empirical significance', viz., the framework developed in this paper, straightforwardly entails this result. Since the symmetry in question is non-boundary-preserving, and hence the corresponding empirical symmetry not purely relational, however, the case to which Faraday's cage is directly analogous is that of Einstein's elevator, rather than Galileo's ship.

8 Local boundary-preserving symmetries: Klein-Gordon-Maxwell gauge theory and t'Hooft's beam-splitter

We turn now to another theory with 'local' symmetries: Klein-Gordon-Maxwell theory, the theory of a complex scalar field minimally coupled to a U(1) connection. Here our task is to justify the claim we made in section 2: that t'Hooft's beam-splitter is a *direct* analog of Galileo's ship, more so even than Faraday's cage.

Mathematically speaking, states of Klein-Gordon-Maxwell theory are specified by a four-potential $A_\mu \equiv (\mathbf{A}, \Phi)$ and a complex scalar function ψ, and the dynamics are generated by the Lagrangian density

$$\mathcal{L} = ((\partial_\mu - iqA_\mu)\psi)^* \, (\partial_\mu - iqA_\mu)\,\psi + m^2\psi^*\psi + \mathcal{L}_{EM}, \tag{15}$$

where $\mathcal{L}_{EM}$ is the standard Lagrangian of classical electrodynamics in the absence of matter. This theory can be understood either as a classical field theory (albeit one not realised in nature) or, as is more usual, the semiclassical theory

---

[21] This point can be ignored when only global symmetries are present, since in these cases there are no nontrivial interior symmetries (i.e. $\mathcal{I}_{\mathcal{S}} = \{\mathbb{I}\}$), and hence $\Sigma_{\mathcal{S}}/\mathcal{I}_{\mathcal{S}}$ is canonically isomorphic to $\Sigma_{\mathcal{S}}$. In fact, we can take the existence of nontrivial interior symmetries to *define* what it is for a symmetry to be 'local' rather than 'global' (relative to a given subsystem/environment split); this may be preferable to the problematic talk of 'symmetry groups parametrised by arbitrary functions of spacetime' that is usually used in attempts to define 'local'.

of a quantized charged scalar particle (a pion, for instance) interacting with a classical background electromagnetic field.

The symmetries of this theory include Poincaré boosts, translations and rotations, which lead to Galileo's ship-type empirical consequences as usual. More interestingly for our purposes, it also has the 'gauge' symmetry

$$\psi(\mathbf{x}, t) \quad \mapsto \quad \exp(-iq\chi(\mathbf{x}, t))\psi(\mathbf{x}, t) \tag{16}$$

$$A_\mu(t, \mathbf{x}) \quad \mapsto \quad A_\mu(t, \mathbf{x}) + \partial_\mu \chi(t, \mathbf{x}), \tag{17}$$

where χ is an arbitrary smooth function. It is easily verified that such joint transformations leave invariant the Lagrangian (15), and hence all equations of motion derived therefrom.

The natural choice of isolated subsystem is a system of matter and electromagnetic fields far from other matter concentrations and from radiation. We specify a subsystem by giving a set $W \subset \mathbb{R}^4$, and the states of subsystem and environment respectively by giving scalar fields and connection on W and $\mathbb{R}^4 - W$, satisfying the dynamical equations in each case. A subsystem state is compatible with an environment state iff the fields of the two agree on some neighbourhood of the boundary of W.

As in the case of electrodynamics, a subsystem symmetry transformation, given by a smooth real-valued function χ_S on the subsystem region W, need not be interior; it will be interior only if χ_S vanishes in a neighbourhood of the boundary ∂W of the subsystem region. In particular, suppose that χ_S is constant, but non-zero, on a neighborhood of the boundary. Then (given that we are considering states in which the subsystem is isolated, and so (in particular) ψ is zero on the boundary) the transformation defined by χ_S leaves the boundary conditions of our states of interest invariant: that is, it defines a symmetry that is boundary-preserving on those states. But χ_S is not an interior symmetry: it does not leave invariant the boundary conditions of *other* subsystem states, in which ψ is *non*zero on the boundary of W. Any two subsystem symmetries are (by definition) equivalent when they differ by an interior symmetry: so χ_S and χ_S' define equivalent symmetries iff $\chi_S - \chi_S'$ vanishes on a neighborhood of the boundary. Assuming that the boundary is connected, this entails that an equivalence class of these symmetries is specified by a real number: the constant value of χ_S on the boundary. A natural choice of representatives for these equivalence classes are the constant functions, sometimes called the *global* gauge transformations.

According to the general theory of empirical symmetries developed in sections 3 and 5, this means that 'local' subsystem gauge transformations that are constant in a neighbourhood of the subsystem boundary correspond, for subsystem states in Klein-Gordon-Maxwell theory satisfying the conditions for spatial isolation, to purely relational realizations of the gauge symmetry; and we might ask what, if any, are the empirical consequences of these realizations. But the answer to this question is clear: at least mathematically speaking, the theory of one particle in an electromagnetic field, described in section 2.2 and there asserted adequate to model t'Hooft's modified beam-splitter, just is the Klein-

Gordon-Maxwell theory discussed in the present section. Hence the empirical symmetry obtaining between the unshifted and shifted half-beams in t'Hooft's beam splitter is a direct empirical consequence of local gauge symmetry, in the same way that Galileo's ship is a direct empirical consequence of boost invariance.

Our analysis is in conflict with Brading and Brown (2004): as we have noted, they argue that interferometry experiments of this kind are not empirical consequences of either global or local symmetries. They reason as follows:

> [We could] consider a region where the wavefunction can be decomposed into two spatially separated components, and then ... 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. (Brading and Brown (2004, p. 656))

Brading and Brown's argument is essentially a more careful version of the argument-sketch we gave in the introduction for why local symmetries might be thought to have no empirical consequences; and, being more careful, it allows us to see why that argument fails. It assumes, tacitly, that if some smooth transformation φ of the fields on a (connected) region $A \cup B$ restricts to gauge transformations φ_A and φ_B on the fields on A and B respectively, then the overall transformation must be a gauge transformation. This would indeed be the case if the physically given transformation φ were itself a *function from spacetime to the gauge group*, with $\varphi_A = \varphi|_A$ and $\varphi_B = \varphi|_B$: for, in that case,

$$\varphi(x) = \begin{cases} \varphi_A(x), & x \in A, \\ \varphi_B(x), & x \in B, \end{cases} \tag{18}$$

would just be the gauge transformation in question. The key to seeing why this argument fails is noting that what is given, when we are given the pre- and post-transformed states of the universe, is *not* a function from spacetime to the gauge group, but merely the effect of whatever transformation is being performed on the *particular* pre-transformation (universe) state $(\psi, A_\mu) \in \mathcal{U}$. And if this particular ψ happens to vanish on the overlap region $A \cap B$, then nothing about the corresponding gauge transformations φ_A, φ_B can be 'read off' from their effects on the wavefunction in that region. It is therefore possible that the universe transformation being performed might correspond to the effect of (say) some constant gauge transformation φ_A in W, and a *different* constant

22

gauge transformation φ_B in B, so that there is no way of patching φ_A and φ_B together to obtain a single smooth function from the whole of spacetime to the gauge group. This possibility is exploited by the scenario under discussion.[22]

General relativity

Our final example is another "local" theory: the General Theory of Relativity, with or without matter fields present. We formulate the theory in the traditional way, in terms of a metric on a manifold; the theoretical symmetries of the theory include all the diffeomorphisms of the manifold. (We ignore any internal symmetries of the matter fields.)

As usual, we have some freedom over what to pick out as a 'subsystem'. For our current purposes, we can proceed in much the same way that we did for the analysis of spacetime symmetries in sections 4–8. Thus: subsystem and environment are each specified by giving the fields on all of some (common) underlying manifold M, but the subsystem is defined by some particular region $W \subset M$.[23] A given subsystem state s and environment state e are *dynamically compatible* iff the metric and matter fields of s and e coincide in some neighbourhood of ∂W; subsystem and environment are mutually *isolated* if (further) the matter fields vanish, and the metric is flat, on such a neighbourhood. We obtain 'subsystem symmetries' by allowing diffeomorphisms (of the entire man-

---

[22]Brading and Brown raise further problems for our analysis: they claim that it is not legitimate to view one half of the split beam as a *subsystem*, which of course is precisely what we are doing. Explaining why they think this is illegitimate, Brading and Brown (writing Ψ, Ψ_I, Ψ_{II} for (respectively) the overall system wavefunction and the components of that wavefunction with support in the region of the upper and lower half-beams, so that $\Psi = \Psi_I + \Psi_{II}$) write:

> 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. ...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 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.

We find this unpersuasive. For one thing, though the theory is quantum-mechanical, it is mathematically identical to a classical theory in which systems can be understood in precisely the way required; for another, the whole problem can be modelled perfectly effectively using Fock-space methods if we wish to understand the two halves of the beam as two quantum-mechanical systems in the usual sense. But more importantly, 'subsystem' is just a word, to be used in whatever way is most perspicuous, and in *this* case (as is not the case in, say, analysis of entanglement), the most perspicuous definition of 'subsystem' does indeed seem to treat the two components as distinct subsystems.

[23]Arguably, this definition fails to take into account certain particular features of general relativity: notably (a) the possibility of topology change, so that a subsystem may not have a fixed topology invariant across all states of that system, and (b) the fact that a given subset of the bare manifold does not correspond to anything physical in the absence of the (matter and metric) fields. We ignore these issues for simplicty (see, however, Wallace (2009b).)

ifold M) to act only on the subsystem state. Applying our usual classification of such subsystem symmetries, we see that:

- The *interior* symmetries of $\mathcal{S}$ are those diffeomorphisms that coincide with the identity diffeomorphism on a neighbourhood of ∂W.

- Diffeomorphisms that coincide with a Poincare transformation on a neighbourhood of ∂W are non-interior, but are boundary-preserving on states that meet the condition for isolation. (As usual for spacetime symmetries, we also need to require that these transformations are not large enough to move the non-flat part of the region W across the boundary.) These correspond empirically to boosts, rotations and translations. (Note that these operations are only definable in GR where the system being transformed lies in flat spacetime — i. e., when it is an isolated subsystem in our sense.) So the Galileo-ship empirical symmetry can be seen as a direct empirical consequence of the diffeomorphism symmetry of the theory.

- *Arbitrary* diffeomorphisms are in general non-interior, and do not preserve the boundary of any states. However, recall, from our general framework, that this need not prevent them from having direct empirical significance: given *any* two subsystem states $s, \sigma_{\mathcal{S}}(s)$ that are related by a subsystem symmetry, we can find an empirical realisation of this symmetry by finding intrinsically different environment states e, e', such that $s * e \in \mathcal{U}$ and the boundary conditions of e, e' are related to one another by the same symmetry (so that $\sigma_{\mathcal{S}}(s) * e' \in \mathcal{U}$ also).

In fact, Einstein's elevator can be seen as an empirical realistation of certain of the non-interior diffeomorphism symmetries of the theory. For in general relativity, if u_1 and u_2 are solutions of the equations such that for some region $W \subset M$:

1. γ_1, γ_2 in W are geodesics according to u_1, u_2 respectively;

2. spacetime is flat on lengthscales $\sim L$ around γ in u_1 and around γ_2 in u_2;

3. tubes S_1 around γ_1 and S_2 around γ_2, each of width L, are both contained in W,

then there will be some diffeomorphism d, taking the surface of S_1 to that of S_2, such that $d_* u_1|_{\partial W} = u_2|_{\partial W}$. Thus u_1, u_2 restrict to environment states e_1, e_2 respectively whose boundary conditions are indeed related by a diffeomorphism, and hence for any subsystem state s compatible with e_1, there is a subsystem-symmetry-transformed counterpart $d_* s_1$ that is compatible with e_2.

Which subsystem states are compatible with environment states that satisfy (1)–(3)? The answer is essentially the same as in Newtonian gravity: any isolated system of matter and metric — some collection of material bodies, say, or a collection of black holes interacting in some complex manner, — of effective size L, so that outside a tube of spatial diameter $\sim L$ the metric is approximately flat and the matter fields approximately vanish. (If the isolated system is not

strongly self-gravitating, the second condition will entail the first.) Any such system is compatible with e_1, and hence with e_2. Hence, any such system will behave identically in any spacetime region whose curvature can be neglected on scales of $\sim L$. And this, of course, is just the (strong) equivalence principle.[24]

9   Summary

We have developed our framework a little at a time, in order to display the motivations for and applications of each of its elements. Here we give a summary of the finished framework itself, for convenience of reference.

- Let $\mathcal{U}$ be the set of dynamically allowed states for the universe (where 'state' is intended in the mathematical, rather than the physical, sense).

- Let Σ be the set of symmetries $\sigma : \mathcal{U} \to \mathcal{U}$ of the universe .

- Let $\mathcal{S}$ and $\mathcal{E}$ be the sets of dynamically allowed states of (respectively) subsystem and environment, and $\pi_{\mathcal{S}} : \mathcal{U} \to \mathcal{S}, \pi_{\mathcal{E}} : \mathcal{U} \to \mathcal{E}$ the projections from universe states onto (respectively) system and environment states.

- Require that the structure $(\mathcal{U}, \mathcal{S}, \mathcal{E}, \pi_{\mathcal{S}}, \pi_{\mathcal{E}})$ satisfies the following conditions:

  A1. For each $s \in \mathcal{S}, e \in \mathcal{E}$, there is at most one $u \in \mathcal{U}$ such that $\pi_{\mathcal{S}}(u) = s$ and $\pi_{\mathcal{E}}(u) = e$. (The states of subsystem and environment jointly determine the state of the total system.)

  A2. For all $\sigma \in \Sigma$ and all $u, u' \in \mathcal{U}$, if $\pi_{\mathcal{S}}(u) = \pi_{\mathcal{S}}(u')$ then $\pi_{\mathcal{S}}(\sigma(u)) = \pi_{\mathcal{S}}(\sigma(u'))$; similarly for $\pi_{\mathcal{E}}$. (The projections from the total system to the subsystem/environment respect symmetries.)

- With these requirements in place, we make the following further definitions and observations:

  - Define the compatibility relation between $\mathcal{S}$ and $\mathcal{E}$ as follows: s is compatible with e iff for some $u \in \mathcal{U}$, $\pi_{\mathcal{S}}(u) = s$ and $\pi_{\mathcal{E}}(u) = e$. (This is intended to be read as 'the subsystem state s and the environment state e are dynamically compatible' — i.e. (for field equations) their fields approximately match on the boundary between them.)

  - Define a partial function $*$ from $\mathcal{S} \times \mathcal{E}$ onto $\mathcal{U}$ as follows: if there exists $u \in \mathcal{U}$ such that $\pi_{\mathcal{S}}(u) = s$ and $\pi_{\mathcal{E}}(u) = e$, then $s * e = u$; otherwise $s * e$ is not defined. (A1 ensures that this does indeed well-define $s * e$ on the desired domain.)

---

[24]Note that this analysis of the equivalence principle does not in any way require that the spacetime *interior* to the subsystem be approximately flat: the equivalence principle, like the relativity principle, applies to strongly self-gravitating systems, not just systems in which gravity can be neglected (something which we also saw to be the case in Newtonian graviation. See Wallace (2009b) for further discussion.

- For arbitrary $s \in \mathcal{S}$ (resp. $e \in \mathcal{E}$), define the *boundary condition of* s (respectively, of e), B_s (resp. C_e), to be the set of all $e \in \mathcal{E}$ (resp. $s \in \mathcal{S}$) such that $s * e$ is defined.

- Define the restriction $\sigma_{\mathcal{S}}$ (resp. $\sigma_{\mathcal{E}}$ of a given total-system symmetry σ to the subsystem (resp. to the environment) as follows: for all $s \in \mathcal{S}$, $\sigma_{\mathcal{S}}(s) = \pi_{\mathcal{S}}(\sigma(u))$, where $u \in \mathcal{U}$ is such that $\pi_{\mathcal{S}}(u) = s$. (A2 ensures that $\sigma_{\mathcal{S}}(s)$, thus defined, is independent of the choice of representative u.) Write $\Sigma_{\mathcal{S}}$ (resp. $\Sigma_{\mathcal{E}}$) for the set of all subsystem (resp. environment) symmetries.

- We are now in a position to define the notions of an 'interior' and a 'boundary-preserving' subsystem symmetry:

  - Let $\mathcal{I}_{\mathcal{S}} \subset \Sigma_{\mathcal{S}}$ be the set of 'interior' symmetries of $\mathcal{S}$; that is, the set of subsystem symmetries $\sigma_{\mathcal{S}} \in \Sigma_{\mathcal{S}}$ such that $\sigma_{\mathcal{S}} * \mathbb{I} \in \Sigma$.

  - For any $s \in \mathcal{S}$, let $\mathcal{BP}_{\mathcal{S}}(s)$ be the set of subsystem symmetries that 'preserve the boundary of s', i.e. the set of subsystem symmetries $\sigma_{\mathcal{S}}$ such that $s, \sigma_{\mathcal{S}}(s)$ lie in the same boundary condition C_e as one another. (Clearly, for any $s \in \mathcal{S}$ we have $\mathcal{I}_{\mathcal{S}} \subset \mathcal{BP}_{\mathcal{S}}(s)$; equally clearly, the interior symmetries are just those symmetries that are boundary-preserving on *all* states $s \in \mathcal{S}$.)

We take *any* $(\mathcal{U}, \mathcal{S}, \mathcal{E}, \pi_{\mathcal{S}}, \pi_{\mathcal{E}})$ satisfying (A1) and (A2) to define a legitimate decomposition of the universe into 'subsystem' and 'environment'. Without further restricting the nature of the 'subsystem', we wish to indulge in talk of the 'intrinsic properties of the subsystem', and so forth; such talk is to be cashed out as follows. (1) The 'intrinsic properties of the subsystem' must depend on $\pi_{\mathcal{S}}(u)$ alone (and not on any other features of u). (2) If some transformation $u \mapsto u'$ does not change which possible world is being represented (i.e., if u and u' are related to one another by a universe symmetry), then the intrinsic properties of the subsystem (resp. environment) according to u must be the same as those according to u'. (3) an 'experiment confined to the subsystem' is an experiment whose outcomes (or probabilities for outcomes) depend on the intrinsic properties of the subsystem alone.

We are now in a position to talk about the conditions under which the intrinsic properties of the subsystem might have *changed*, without having said anything (else) about what those intrinsic properties *are*. Specifically, the following are now direct consequences of our definitions:

- Interior symmetries have no direct empirical significance: if $\sigma_{\mathcal{S}} \in \mathcal{I}(\mathcal{S})$, then for all $s * e \in \mathcal{U}$, $\sigma_{\mathcal{S}}(s) * e$ is defined *but* represents the same possible world as $s * e$.

- If $\sigma_{\mathcal{S}}$ is a non-interior symmetry that is boundary-preserving on some state $s \in \mathcal{S}$, then, for all $e \in \mathcal{E}$ such that $s * e$ is defined, the states $s * e$, $\sigma_{\mathcal{S}}(s) * e$ in general represent different possible worlds from one another *but*, given that the two subsystem states are related by a subsystem symmetry,

then (by (1) and (2)) the transformation induces no *intrinsic* change to the subsystem (and of course induces none to the environment). Hence, the difference will not be detectable by experiments confined to either subsystem or environment alone: all that has been changed are *relations* between the subsystem and environment.

- Quite generally, if $\sigma_{\mathcal{S}}$ is a non-interior symmetry of the subsystem $\mathcal{S}$ (not necessarily boundary-preserving on any salient collection of subsystem states), still any two environment states e, e' whose boundary conditions are related to one another by σ_s are dynamically compatible with the same *intrinsic* states of the subsystem as one another. Formally: $\forall \sigma_{\mathcal{S}} \in \Sigma_{\mathcal{S}}, e, e' \in \mathcal{E}, \forall s \in \mathcal{S}$, if $C_{e'} = \sigma_{\mathcal{S}}(C_e)$ then $\forall s \in \mathcal{S}, s * e \in \mathcal{U} \Rightarrow \sigma_s(s) * e' \in \mathcal{U}$. If e, e' are *intrinsically* different states of the environment, then the transformation $s * e \mapsto \sigma_{\mathcal{S}} * e'$ is a change in physical state of the universe, and one that is *not* purely relational. The empirical symmetry will nevertheless 'correspond to' the theoretical subsystem symmetry $\sigma_{\mathcal{S}}$, insofar as (and *only* insofar as) there is a principled connection between the subsystem symmetry transformation $\sigma_{\mathcal{S}}$ and the environment transformation $e \mapsto e'$. We have not investigated the possibilities for such 'principled connections'.

10   Conclusions

We began this paper with two pieces of conventional wisdom: that the direct empirical consequences of symmetry transformations arise when only a subsystem is transformed, and that because of this, global but not local symmetries have direct empirical significance. In developing an appropriate framework to explicate the first of these, we have found fault with the latter. The perspicuous distinction for analysis of phenomena on the pattern of Galileo's ship is subsystem-relative, and is between symmetries of the subsystem in question that are *interior* and those that are non-interior, rather than between restrictions of global and restrictions of local symmetries to the subsystem: it is the interior symmetries of a given subsystem that cannot have empirical consequences when performed on that subsystem. The second piece of conventional wisdom is simply false: an element of a given 'local' symmetry group would be unable to have empirical significance *simpliciter* only if its restrictions to *all* subsystems were interior symmetries of those subsystems, but no transformation (other than the identity, of course) has this property. The relationship between the interior/exterior and the global/local distinction is this: only when the group of theoretical symmetries is 'local' will there *be* any interior symmetries, distinct from the identity, for subsystems that are identifiable as subregions of spacetime.

We are now in a position to resolve the issues raised by the three objections to which the claim that 'global symmetries have empirical significance but local symmetries do not' gave rise, at the start of this paper.

Our first objection was that, since the global symmetries form a subgroup of

the local symmetry group, it is logically impossible that all global and no local symmetries can have empirical significance. The short way with this objection, in the light of our analysis, is to note that we have rejected the offending claim that local symmetries cannot have empirical significance. However, we should also check that no analogous paradox recurs for our new set of claims. Indeed it does not. The reason lies in a structural difference between the original problematic claims and our replacements: rather than holding that all elements of a ('local') symmetry group have one property while elements of some subgroup thereof have a contradictory property, we hold that (for any given subsystem) there is a subgroup of subsystem symmetry transformations (the 'interior' ones) that cannot have empirical significance, and that it is elements of the *quotient* of the larger group by this subgroup that are candidates for correspondence to physical operations. There is thus no object of which we assert both that it does, and that it does not, have some given property.

We will return to our second objection in a moment. Our third objection was more concrete: it was that several particular cases, for example those of Faraday's cage and t'Hooft's beam-splitter, seemed (at least) to be perfect analogs of the Galileo-ship thought-experiment for various local symmetries: that is, cases in which *local* symmetries corresponded to empirical effects in just the same way that the boost invariance of Newtonian mechanics corresponds to the Galileo-ship effect. Our framework classifies the various cases we have discussed as follows:

Theory	Theoretical symmetries	Symmetry group global or local?	Boundary preserving?	Corresponding empirical symmetries
All flat-spacetime theories (e.g. Newtonian mechanics, Coulombic electrostatics, Newtonian gravity)	Galilean/Poincare transformations	Global	Yes	Galileo's ship
General relativity	Asymptotically Poincare transformations	Local	Yes	Galileo's ship
Quantum mechanics with only global phase symmetry	Constant phase shifts	Global	Yes	t'Hooft's beam-splitter
Quantized Klein-Gordon-Maxwell theory	Asymptotically constant gauge transformations	Local	Yes	t'Hooft's beam-splitter
Theories with a 'global' potential shift symmetry (e.g. Coulombic electrostatics, Newtonian gravity)	Spatially constant potential shifts	Global	No	Faraday's cage (or gravitational analog)
Theories with a 'local' potential shift symmetry (e.g. Classical electrodynamics, general relativity)	Asymptotically constant potential shifts	Local	No	Faraday's cage (or gravitational analog)
Newtonian gravity	Spatially constant accelerations	Global	No	Einstein's Elevator
General relativity	Diffeomorphisms not vanishing on the boundary of a system	Local	No	Einstein's elevator

Examining this list, we see that there is no particularly important distinction between theories with a global symmetry group and those with a local symmetry group, from the point of view of the connection between theoretical

and empirical symmetries. Every 'global' theory and the corresponding 'local' theory will entail exactly the same empirical symmetries, with constant transformations in the former case replaced by equivalence classes of transformations in the latter. (Healey might just as well have posed his puzzle in terms of the ability of general relativity to explain Galileo's ship: the association of Galileo's ship more closely with global symmetries, and Faraday's cage more closely with local symmetries, is spurious.) This in turn resolves our *second* original puzzle, about whether explanatory power is lost when we 'localise' a symmetry, in a way that (in particular) should worry scientific realists: it is not.

There *is*, however, a principled distinction between the (non-interior) symmetries that are 'boundary-preserving' on states satisfying the condition for dynamical isolation, and the (non-interior) symmetries that are *not* boundary-preserving even on such special states. For this reason, it is t'Hooft's beam-splitter that is *precisely* analogous to Galileo's ship, while Faraday's cage is precisely analogous to empirical realisations of the Equivalence Principle.

Acknowledgements

References

Brading, K. and H. Brown (2004). Are gauge symmetry transformations observable? *British Journal for the Philosophy of Science 55*, 645–665.

Brown, H. and R. Sypel (1995). On the meaning of the relativity principle and other symmetries. *International Studies in the Philosophy of Science 9*, 235–253.

Deutsch, D. and P. Hayden (2000). Information flow in entangled quantum systems. *Proceedings of the Royal Society of London A456*, 1759–1774.

Galileo (1632/1967). *Dialogue Concerning Two Chief World Systems* (2nd revised ed.). Berkeley, California: University of California Press. Translated by S. Drake.

Healey, R. (2007). *Gauging What's Real*. Oxford: Oxford University Press.

Healey, R. (2009). Perfect symmetries. *British Journal for the Philosophy of Science 60*, 697–720.

Kosso, P. (2000). The empirical status of symmetries in physics. *British Journal for the Philosophy of Science 51*, 81–98.

Maxwell, J. C. (1881). *An Elementary Treatise on Electricity*. Oxford: Clarendon Press.

Norton, J. D. (2008). The hole argument. In *The Stanford Encyclopedia of Philosophy* (Winter 2008 ed.). http://plato.stanford.edu/archives/win2008/entries/spacetime-holearg/.

Saunders, S. (2003). Physics and Leibniz's principles. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, pp. 289–308. Cambridge: Cambridge University Press.

t'Hooft, G. (1980). Gauge theories and the forces between elementary particles. *Scientific American 242*, 90–166.

Wallace, D. (2009a). QFT, antimatter, and symmetry. *Studies in the History and Philosophy of Modern Physics 40*, 209–222.

Wallace, D. (2009b). The relativity and equivalence principles for self-gravitating systems. Forthcoming in *Einstein Studies*; available online at http://philsci-archive.pitt.edu/4533/.

Wallace, D. and C. Timpson (2007). Non-locality and gauge freedom in Deutsch and Hayden's formulation of quantum mechanics. *Foundations of Physics 37(6)*, 951–955.