Observability, redundancy and modality for dynamical symmetry transformations

David Wallace

November 7, 2019

Abstract

I provide a fairly systematic analysis of when quantities that are variant under a dynamical symmetry transformation should be regarded as unobservable, or redundant, or unreal; of when models related by a dynamical symmetry transformation represent the same state of affairs; and of when mathematical structure that is variant under a dynamical symmetry transformation should be regarded as surplus. In most of these cases the answer is 'it depends': depends, that is, on the details of the symmetry in question. A central feature of the analysis is that in order to draw any of these conclusions for a dynamical symmetry it needs to be understood in terms of its possible extensions to other physical systems, in particular to measurement devices.

1 Introduction: the dynamical conception of symmetry

In physical practice, symmetry is for the most part a dynamical notion: a *dynamical symmetry* is a transformation that takes solutions of the equations of motion to other solutions. Such a notion has obvious practical applications in dynamical problems, such as the solution of the equations of motion and the identification of conserved quantities, but the significance of symmetry seems to transcend this: from symmetries, it is often said, we can infer consequences about what can be observed (symmetry-variant quantities are said to be unobservable), about what parts of theories do and do not do work (symmetry-variant quantities are sometimes said to be 'redundant'; symmetry-variant structures to be 'surplus'), about representation (symmetry-related models are often taken to play the same representational role) and about modality (symmetry-related situations are at least sometimes treated as the same situation differently described).

Just what these inferences really are supposed to be, and why they are justified — if indeed they are — is a vexed question. They cannot apply without some qualifications and restrictions: in an austere sense, a quite arbitrary permutation of the space of solutions is a symmetry, and even when we restrict attention to transformations which are closer to physical practice, still problematic examples abound. To borrow an example from Belot (2013): the orbit of a planet around the Sun is well modelled by the *central-force model*, in which we idealise the sun as at rest and exerting a centrally directed inverse-square force on the planet.[1] This theory has the familiar rotational and time-translation symmetries (boost and translational symmetry are broken by the fixed location of the Sun) but in addition, it has 'Lenz-Runge symmetries', somewhat complicated and unintuitive velocity-dependent transformations of the planet's position that do not leave its distance from the Sun invariant. Unless we make the desparate move of supposing that the distance of a planet from the sun is neither observable nor physically significant, none of our theses seem to apply to this symmetry.

Of course one can try to rule out such counter-examples by narrowing the scope of symmetries to which the theses of unobservability, representational equivalence and the like are supposed to obtain, but it is not easy to see how this could be done, much less what principled justification there could be for doing it. To quote from two influential recent criticisms of drawing metaphysical and epistemic conclusions from the dynamical conception of symmetry (whose starting points differ sharply):

> [T]he ways of encoding the content of laws that are most appealing to mathematicians and physicists appear to lead to notions of [dynamical] symmetry that are coolly indifferent to considerations of representational equivalence (Belot, *ibid*)

> [T]he notion [of symmetry] is often defined in purely formal, mathematical terms, so that whether a given transformation is a symmetry of a given set of laws depends just on the formal and mathematical features of those laws and their models. But why should those features of the laws have anything to do with metaphysics, with what's real? It's not obvious[.] (Dasgupta 2016)

A common response to this problem is to look for some substantive additional requirement for a transformation to count as a symmetry, along with the formal requirement that it takes solutions to solutions. There are two broad classes of strategy. The first (the 'epistemic strategy') effectively builds unobservability of symmetry transformations into the definition of a symmetry: symmetries are required to leave all observable properties of the system invariant. Ismael and van Fraassen (2003) and Dasgupta (2016) propose definitions explicitly along these lines; some advocates of the 'primitive ontology' approach to scientific metaphysics (e. g. Allori *et al* (2008) and Allori (2013)) propose something closely related (they require that symmetries leave relevant facts about the 'primitive ontology' invariant, and if one grants their account of scientific epistemology[2] then observations supervene only on the primitive ontology). The second (the

[1] It is even better modelled if we work in center-of-mass coordinates, but I avoid this complication for expository convenience.

[2] One should not; cf Wallace (2018).

'representational strategy') instead builds the representational equivalence of symmetry-related models into the definition, usually by requiring that symmetries are automorphisms of the appropriate mathematical space of models (hence preserve all structure, and thus all representation-apt features, of a model). This strategy has been widely explored in the philosophy-of-spacetime literature, where symmetries are often taken to be diffeomorphisms that preserve 'absolute structure'; see, e. g. , Earman (1989) and references therein. (The preservation of representational equivalence is also built directly into Healey's (2009) definition of a 'theoretical symmetry'.)

Both strategies would have the disappointing consequence of making at least some of the theses about symmetry trivial. if symmetry is *definitionally* observation-preserving, for instance, we can infer nothing *substantive* about what is observable by learning that a theory has a symmetry, since the only way to learn that would have been via deductions about observability; if symmetry is *definitionally* structure-preserving, we cannot learn about surplus structure via symmetry. But more importantly, both appear in conflict with physical practice. This is most sharply apparent for the epistemic strategy: as Dasgupta freely admits, "on our epistemic approach, symmetry-to-reality reasoning involves not just mathematical analysis, but also considerations that reach into the philosophy of perception and mind. A far cry from the purely mathematical gloss it is often given!" Since the definition of symmetry in physics pretty clearly does not reach at all into the philosophy of perception and mind, and indeed frequently involves identifying symmetries in contexts far from the directly-observable world, 'symmetry' in the epistemic-strategy sense is at any rate not what physicists mean when they talk about symmetry.

Matters are somewhat more subtle for the representational approach. It is, for example, perfectly possible to formulate Newtonian physics as a theory whose models are N-tuples of paths in $\mathbb{R}^4$, in which case the mathematical structure of a model includes many features (the x-component of the vector linking a particle to the spatial coordinate origin at time=0, for instance) that are not invariant under the Galilean symmetries. But it is possible to reformulate Newtonian physics on a more sparsely-structured space ('Galilean spacetime'; see, e. g. , Weatherall (2017)) for which the Galilean symmetries coincide with the automorphisms. And having made such a reformulation, we can maintain that the reformulated theory is the truer representation of the physics, and that in the original $\mathbb{R}^4$ formulation, models are representationally equivalent whenever they are symmetry-related even though not all their mathematical structure is preserved under the symmetry transformation.

But what makes such a reformulation appropriate? The answer, in the history of physics, is pretty much always based on identifying a *dynamical* symmetry under which some of the structure of the original models is variant: that is, it is based on the assumption that structure variant under a dynamical symmetry is surplus. The Galilean-spacetime reformulation is preferable to the original formulation precisely because the automorphisms match the dynamical symmetries. (Earman (1989) makes this argument explicit.)

(As another, more technical example, $U(1)$ gauge theory can be formulated, and historically was formulated, as a theory of a 4-vector field on spacetime. That theory has gauge symmetries as dynamical symmetries, and those symmetries do not preserve the mathematical structure of the theory: for instance, the property of having divergence zero is not gauge-invariant. The theory can be reformulated as a connection on a fibre bundle, and now the gauge symmetries are structure-preserving bundle automorphisms — but the rationale for the refomulation was precisely *because* the dynamical symmetries were not automorphisms in the original formulation.)

So the representational strategy, too, does not seem to do justice to the notion of symmetry actually used in physics: insofar as it is correct, it is only because we have carried out reformulations, and/or drawn conclusions about when models are representationally equivalent, based on a prior, dynamical notion of symmetry.

Only the dynamical approach, I think, can really do justice to the notion of symmetry *as we find it in physics*. And so my goal in this paper is to defend the claim that epistemic and (indirectly) metaphysical conclusions can after all be drawn from the presence of a dynamical symmetry, at least in some circumstances, and — more importantly — to clarify how that can even be possible. To preview the core claims:

(i) observation is itself a dynamical notion, so a dynamical symmetry can have consequences for what is observable just when that symmetry encompasses the physical processes of measurement;

(ii) models in physics are almost always used to describe subsytems of a larger universe, so the interpretational significance of a symmetry depends not just on the dynamics of the system itself but on (explicit or tacit) assumptions about what happens to the symmetry when the system interacts with other systems.

The structure of the paper is as follows. After outlining (section 2) the main mathematical concepts I will use, in section 3 I give sharp versions of four theses about symmetries in the literature: that they are unobservable, that they give rise to representationally-equivalent models; that they are a guide to surplus structure; that they do not represent genuinely different possibilities. In sections 4-5 I analyse the consequences of assuming that a dynamical symmetry of a system applies also to the mechanical processes used to observe that system, and establish a version of the thesis about unobservability. In sections 6-7 I use these results, and the broader framework developed in sections 4-5, to assess the remaining three theses. In sections 8–10 (which in large part can be read independently and can be skipped on a first reading) I consider three complications: time-dependent symmetries such as those found in gauge theories; quantum theory; cosmology. Section 11 is the conclusion. For the most part I restrict myself to the comparatively-elementary example of point-particle classical mechanics, though sections 8–10 consider more advanced examples from quantum mechanics and field theory.

This paper is one of a series of three papers considering symmetries of isolated systems. The second and third (Wallace 2019a, 2019b) presuppose the results of this paper and explore in more detail how dynamically-isolated subsystems of larger systems can be modelled, how in turn isolated systems can and should be interpreted as subsystems of unspecified larger systems, and the interplay of those ideas with the interpretation of a system's symmetries (the two papers focus, respectively, on examples from particle mechanics and from field theory).

2 Elements of classical mechanics and group theory

For the most part, in this paper I consider *classical* (i. e. , non-quantum) systems: this is mostly for grounds of exegetical convenience, though there are a few additional subtleties in the quantum case which I discuss in section 9. In this section I lay out some basic assumptions about classical mechanics which I draw upon throughout the paper (readers already familiar with the subject might want to skim this section to get familiar with my conventions, or else skip it entirely).

Classical mechanics, in the abstract, characterises a system's dynamical history as a time-indexed collection of *configurations.* In the classic example of point-particle mechanics (with, say, N particles), a configuration is a specification of the locations of each of the point particles, and is mathematically represented by an ordered N-tuple of points in 3-dimensional Euclidean space. The collection of all such N-tuples is *configuration space*, which I denote by $\mathcal{Q}$. More abstract (say, field-theoretic, or extended-body) systems have configuration spaces with different mathematical representations: the configuration space of a rigid-body system, for instance, is given by a point representing the center-of-mass position of the system and a triple of orthogonal vectors representing its orientation.

A *history* of the system is then a smooth function $q : \mathbb{R} \to \mathcal{Q}$, assigning to each time t the configuration $q(t)$ of the system at that time. The dynamical equations of Newtonian mechanics distinguish dynamically possible from dynamically impossible histories; these equations are second-order in time, and so take the form[3]

$$F_J(\ddot{q}(t), \dot{q}(t), q(t)) = 0 \quad (1)$$

for some functions F_J. (Here as usual $\dot{q}(t)$ and $\ddot{q}(t)$ denote, respectively, the first and second derivative of q with time.) The particular form of these equations depends on the system in question: for Newtonian point-particles moving under gravity, for instance, they encode Newton's second law and the inverse-square law. In that case, and in most elementary cases, the dynamics are *deterministic*,

[3]For simplicity I ignore the possibility of explicit time dependence in the dynamics; it can be incorporated without substantive changes to my arguments, at the cost of a more cumbersome notation.

which means that (1) can be solved for $\ddot{q}$ in terms of q and $\dot{q}$:

$$\ddot{q}(t) = \mathcal{F}(q(t), \dot{q}(t)). \tag{2}$$

Given $q(t)$ and $\dot{q}(t)$, $\ddot{q}(t)$ is then fixed; normally[4] this allows us to integrate the equations forward and backwards to determine a unique history. Because these equations are second-order, however, even for a deterministic system the configuration at time t is insufficient to determine future or past configurations: additional information about the rate of change of the configuration is required. The *state* of the system at time t is the ordered pair $(q(t), \dot{q}(t))$ of the system's configuration and its velocity: the space of all such pairs is the *state space* of the system, which I denote by $\mathcal{S}$.[5] In a deterministic system, the state of the system at one time uniquely determines its state at past and future times.

Given a theory in this form, a *(dynamical) configuration symmetry* is a bijection of $\mathcal{Q}$ which maps solutions to solutions: that is, if $q \to gq$ is the map, then $gq(t)$ solves the equations of motion iff $q(t)$ does. The bijection induces a bijection on the space of *states* (called the *tangent lift*) in a natural way:

$$\dot{q}(t_0) \to \left. \frac{\mathrm{d}}{\mathrm{d}t}(gq(t)) \right|_{t=t_0} \tag{3}$$

(note that this depends only on $q(t_0)$ and $\dot{q}(t_0)$). For instance, spatial translation has no effect on velocity; spatial rotation rotates the velocity vectors as well as moving the configuration.

The requirement for g to be a symmetry can be characterised directly on state space too, it is a symmetry provided that $gx(t)$ solves the equations of motion on state space iff $x(t)$ does. The class of bijections of $\mathcal{S}$ satisfying this requirement contains some transformations that are not tangent lifts of configuration symmetries (the Lenz-Runge symmetries have this feature).

The class of dynamical symmetries (whether understood as configuration symmetries or directly as state-space symmetries) forms a group in the mathematical sense: the trivial map id, taking each point to itself, is obviously a symmetry; if g, h are symmetries then so is gh; the inverse of a symmetry is also a symmetry. It is often helpful to think of this group abstractly (i. e., just in terms of its algebraic structure) and to distinguish the group from its action via transformations. Some more terminology: given a group $\mathcal{G}$, an *action* R of the group on a space $\mathcal{S}$ is a collection[6] of bijections $R(g) : \mathcal{S} \to \mathcal{S}$ satisfying $R(g)R(h) = R(gh)$ for any $g, h \in \mathcal{G}$. I write e for the identity in a group (so

[4] There are subtleties in doing so in certain systems (see Earman (1986) for detailed discussion) but those subtleties are normally 'unphysical' in the sense that they mark a failure of the system to correctly represent real-world physics after a certain time; in any case, I set these complications aside.

[5] Technically-minded readers will recognise that this simplifies some of the geometry: the state space is better described as the tangent bundle $T\mathcal{Q}$ over the configuration space $\mathcal{Q}$. These details will not matter for what follows.

[6] Technical detail: $\mathcal{G}$ should be a Lie group, $\mathcal{S}$ should be a differentiable manifold, and the map $(g, x) \to R(g)x$ should be a diffeomorphism.

that if R is an action, $R(e) = \mathrm{id}$). For simplicity, I normally assume that the group action is *free*: that is, if $R(g)x = x$, then $g = e$.

This notation is unavoidably somewhat abstract, but (as I noted in the introduction) for most purposes it will suffice to consider one concrete example: the case of Newtonian point particles (I will occasionally refer to other, more complicated, examples below, but readers unfamiliar with the physics of these examples may wish to skim them and focus on the Newtonian case). For definiteness, I will consider N particles moving under the inverse-square law of Coulomb electromagnetism, where the nth particle has a charge q_n and a mass m_n, where the nth particle exerts a force on the mth particle proportional to $q_n \times q_m$ divided by the square of the distance between them, and where each particle has an acceleration directly proportional to the sum of the forces upon it and inversely proportional to its own mass. There are two classes of dynamical symmetry of this theory:

Galilean spacetime symmetries: the transformation induced by translating all the particles by a common vector, or rotating them by a common amount, or by increasing their velocities by a common amount, or by translating each of their trajectories by a common amount in time, is a symmetry. (Strictly, the last two require a slight generalisation of our definition, which I omit for simplicity.)

Permutation symmetries: If any two particles have the same charge and mass, the transformation that permutes them is a symmetry.

3 Four theses about symmetry

To ensure a clear target for later analysis, here I give fairly precise formulations of four common (though contested) claims about symmetry. The first makes sharp what I mean by symmetries being unobservable.

The Unobservability Thesis: Given a family of models of a system which are related by a symmetry transformation, it is impossible to determine empirically which model in fact represents the system.

The Unobservability Thesis is sometimes stated in more metaphysical terms, as the claim that quantities or properties that are 'symmetry-variant' — that is, that have values that are not invariant under a symmetry transformation of the system — are unobservable. I will argue that at least in some circumstances, and under some natural intepretations, that claim is false.

The next two theses both concern the way representation works in a theory with symmetry.

The Representational Equivalence Thesis: Given a family of models of a theory which are related by a symmetry transformation, insofar as one model successfully represents a system, so do all the others.

The Surplus Structure Thesis: Given a theory with a symmetry transformation, insofar as the symmetry falls short of being an automorphism of the mathematical structures used to define the theory's models, then this points to aspects of that structure which are redundant, do no representational work, and can be removed from the theory without loss.

These two are clearly closely related: insofar as the Representational Equivalence Thesis holds, then analytically any symmetry-variant structure does no representational work; conversely, once surplus structure is removed, symmetries preserve all mathematical structure, and plausibly mathematical models represent *only* via such structure. (Both should be distinguished from the much stronger claim that, given a symmetry, one should quotient out by that symmetry in order to understand the theory properly: that claim seems to me highly implausible in full generality in view of how ubiquitously modern mathematics deploys structures with nontrivial automorphism groups but in any case I shall have little to say about it.)

The final claim concerns the modal status of symmetries:

The Modal Equivalence Thesis: Two states of affairs related by the action of a symmetry transformation are really the same state of affairs, differently described.

Its relation to the Representational Equivalence thesis is subtle, as we will see in section 7.

4 Observations from within a system

Most physical systems we study in physics are simply not rich enough to model the process of measurement — indeed, hardly any are. The most straightforward model of the Earth-Sun system, for instance, has only six degrees of freedom (three coordinates of position for the centers of mass of each of the two bodies) and it makes little sense to imagine how any physical process in that system could deserve to be called a 'measurement'. But for the moment suppose that we *do* have a sufficiently rich system, and ask what we can infer about the measurability of a symmetry.

The standard model for measurement in physics is the measurement of one system's state by encoding it in another system's. It goes like this:

1. The system being measured (the 'target system') has some unknown initial state, represented by a point in its state space.
2. The system doing the measurement (the 'measurement device') is placed in some 'ready' state, independent of what state the target system starts off in. (If it were not independent, the observer would have to have information about the result of the measurement even before making it, defeating the purpose of the measurement.)

3. The target system and the measurement device are then allowed to interact for a while. As a consequence, the post-interaction state of the measurement device now depends functionally on the pre-measurement state of the target system.

4. If that functional dependence is non-trivial (that is, if the measurement device's post-interaction state sometimes varies according to what the target system's pre-interaction state was) then the measurement has succeeded to at least some degree.

Here we are considering measurements within a single complicated system, but any such system can be decomposed into subsystems — a measurement, in this context, is a correlation established between the pre-measurement value of some 'target' quantity and the post-measurement value of a different 'readout' quantity, given that the pre-measurement value of that second quantity was in the fixed 'ready' state.

To see how this plays out in a system with a dynamical symmetry, it will be helpful to get a little more formal. Given a group $\mathcal{G}$ and an action R of $\mathcal{G}$ on a space $\mathcal{S}$, we can define the *orbits* of the action as the equivalence classes of $\mathcal{S}$ under the relation: x is related to x' iff $x' = R(g)x$ for some $g \in \mathcal{G}$. Suppose we arbitrarily pick a single point ϕ_O inside each orbit O: then any point $x \in O$ can be expressed as $x = R(g)\phi_O$ for some group element g, and hence, that point can be represented by the ordered pair (O, g). (Since we are assuming that the group acts freely on $\mathcal{S}$ — which is to say: $R(g)x = R(h)x$ iff $g = h$ — then the representation will be unique; if we had not made this assumption, multiple pairs might correspond to the same point, but each pair would still pick out a unique point.) I write $x \sim (O, g)$ to denote $x = R(g)\phi_O$.

The value of this representation is that the action of $\mathcal{G}$ has a very simple form: if $x \sim (O, g)$, so that $x = R(g)\phi_O$, then $R(h)x = R(h)R(g)\phi_O = R(hg)\phi \sim (O, hg)$. Hence the action of the group is just:

$$R(h)(O, g) = (O, hg). \tag{4}$$

We have effectively decomposed the space into its $\mathcal{G}$-invariant part, represented by an orbit, and its $\mathcal{G}$-variant part, represented by an element of $\mathcal{G}$ itself.

Now let's take $\mathcal{S}$ to be the state space of our system. The Unobservability Thesis concerns whether it is possible to distinguish experimentally between situations that differ only by the action of the symmetry. In this notation, this is to say: can we measure the value of g from within the system? Since our 'readout' quantity must be independent of g, it follows that it must be a function of O alone, so that the question is: is there a dynamically possible measurement process that encodes the value of g in some function of O? Intuitively, if the action of $\mathcal{G}$ is a symmetry, the answer must be 'no', but we should establish this more carefully.

Now we assume that the system has some time-independent, deterministic dynamics, with equations of motion represented by some one-parameter family $U(t)$ of bjiections of $\mathcal{S}$, satisfying $U(t + s) = U(t)U(s)$. The solutions to the dynamics are then functions $x : \mathbb{R} \to \mathcal{S}$ satisfying $U(t)x(s) = x(s+t)$. The group action is a *dynamical symmetry* if, for any $x \in \mathcal{G}$ and any path $x(t)$ in $\mathcal{S}$, $x(t)$ is a model of the theory (i. e., satisfies its equations of motion) iff $gx(t)$ is.[7]

Let's see how the dynamics looks in the orbit/group representation of state space. An initial state (O, g) evolves over some time t to a subsequent state (O', g'), where O' and g' are determined by O, g and t. So there must be functions α and β such that the evolution has form

$$U(t)(O, g) = (\alpha(O, g, t), \beta(O, g, t)). \tag{5}$$

Suppose we evolve the system for time t and then apply a symmetry h: then the final state will be

$$hU(t)(O, g) = (\alpha(O, g, t), h\beta(O, g, t)). \tag{6}$$

If instead we apply the symmetry and then evolve the system, we get

$$U(t)(O, hg) = (\alpha(O, hg, t), \beta(O, hg, t)). \tag{7}$$

But if the action is a dynamical symmetry, these must be equal, so that we have

$$\alpha(O, hg, t) = \alpha(O, g, t) \quad \beta(O, hg, t) = h\beta(O, g, t). \tag{8}$$

Adjusting notation slightly, this is to say that the dynamics can be expressed as

$$U(t)(O, g) = (\alpha(O, t), g\beta(O, t)). \tag{9}$$

So the future evolution of O depends only on the present value of O, with no additional dependence on g. In other words: there is a self-contained dynamics for the invariant degrees of freedom of the system that is quite independent of the $\mathcal{G}$-variant features.

We can now see why no measurement internal to a system can distinguish whether a symmetry has been applied to the whole system. Such a measurement, we have argued, would have to have the form

$$(O_0, g) \to (O(g), g') \tag{10}$$

where O_0 is some g-independent state and $O(g)$ has some functional dependence on g. But $O(g)$ depends only on O_0 and, despite the notation, is not functionally dependent on g at all.

So the Unobservability thesis is a straightforward consequence of the dynamics — *if* we assume that we are considering a system rich enough to model its own dynamics, and that the system is measuring itself rather than being observed from outside.

[7] This is a slightly restrictive formulation, ruling out as it does time-dependent symmetries like Galilean boosts, and time translation itself. This is for technical convenience; both can be incorporated, but at a significant cost in complexity. Specifically, we could define a symmetry group by a group $\mathcal{G}$, a time-indexed family of actions R_t of $\mathcal{G}$ on configuration space, and an action r of $\mathcal{G}$ on the real line, such that $t \to \sigma(t)$ solves the equations of motion iff $t \to R_t(g)\sigma(r(g)t)$ does; and we could drop the automatic presupposition of time translation symmetry by replacing the one-parameter dynamical map $U(t)$ with a two-parameter map $U(t, t_0)$. I leave it to the reader to verify that my results go through *mutatis mutandis* under this approach.

5 Symmetries of systems and subsystems

In practice, it hardly ever makes sense for the models we use in physics to imagine that the model includes its own measurement processes. In virtually all cases the measurement process, if any, is external to the system, and it is not normally modelled explicitly unless we are actually in the business of designing lab apparatus. So the consequences of the previous section for symmetries of realistic systems are indirect. Nonetheless they are there, as we can see by considering some of the examples where dynamical symmetries clearly do *not* entail anything about observability.

Consider first the Lenz-Runge example I discussed in the Introduction, which is a symmetry of the central force problem — for concreteness, let's suppose we're using that problem to model the Sun-Earth system. A measurement of that system might, for instance, record the distance of the Earth from the Sun at some given time. Without any commitment at all to the details of the physics of that measurement process, we know it must have a schematic form *something* like[8]

$$\text{Earth at distance } r \text{ from Sun; device ready}$$

$$\longrightarrow \text{Earth at distance } r \text{ from Sun; device records } r.$$

Suppose that some instance of the Lenz-Runge symmetry, imagined to be applied before the measurement, changes r to $2r$. Then the measurement process will leave the device in some '$2r$' state. On the other hand, since the symmetry transformation acts only on the Earth and Sun, and is not even well-defined for whatever mixture of optics and mechanics implements the measurement process, applying the transformation *after* the measurement leaves the measurement device in the 'r' state. Which is to say, the Lenz-Runge transformation is *not* a symmetry of the combined system of measurement device plus Earth-Sun system.

A second example: (classical[9]) vacuum electrodynamics (that is, the theory of electromagnetic fields in the absence of charge) has a conformal symmetry in addition to the spacetime symmetries: it has no absolute notion of length or time, and the map $\mathbf{x}, t \to \lambda\mathbf{x}, \lambda t$, applied to spacetime points, takes solutions to solutions. The Unobservability Thesis would seem to imply that radiation frequency is undetectable, so that no observation can distinguish visible light from X-rays, and this is obviously absurd. But it is obvious *why* it is absurd: because the conformal symmetry is *only* a symmetry of radiation in the absence of matter, and ceases to be a symmetry of systems in which matter is present.

[8] Nothing about the concept of measurement actually requires the post-measurement Earth-Sun distance to equal the pre-measurement distance, but obviously any practically realistic measurement process will have this feature.

[9] Quantum electrodynamics is conformally invariant to leading order in the absence of charged particles, but that invariance is violated by quantum corrections, due to loop effects where particle-antiparticle pairs are created (and, indirectly, due to anomalous breaking of the conformal symmetry).

Here is the general point: we have defined symmetries as transformations on the state spaces of *isolated systems*: those systems have self-contained dynamics and symmetries preserve the solutions of those dynamics. But while there is a temptation in the foundational literature to regard such isolated systems as proposed models of the entire Universe,[10] in real physical practice the vast majority of 'isolated system' models are used to model certain subsystems of the Universe under the idealisation that (at least on certain timescales, at least to a certain degree of accuracy) they are dynamically independent of other systems. (The previous two examples illustrate this: the central force problem, and vacuum electrodynamics, are scientifically interesting because (respectively) one body's gravitational effect often dominates all others within a certain region, and because electromagnetic waves often exist for a time in regions with little matter present — not because of an abstract curiosity about distant possible worlds containing only two bodies, or containing no matter at all.)

So if an isolated system has a symmetry group $\mathcal{G}$, what happens to the symmetry when the assumption of isolation fails - for instance, when a measurement device interacts with the system? We can consider three substantively different possibilities. Firstly, $\mathcal{G}$ might be *subsystem-specific*: it has no extension to a symmetry of the combined system. This is the case for the Lenz-Runge symmetry of the central-force model: if we introduce another system (say, another planet passing close to the first, or a non-gravitational measurement device) then it doesn't really even make sense to ask how the Lenz-Runge symmetry affects this system, and even if we did come up with some arbitrary way of applying the transformation to the second planet, there is no reason at all to expect it to be a symmetry. The same applies to the conformal symmetries of vacuum electromagnetism.

Secondly, it might be that $\mathcal{G}$ is also a dynamical symmetry of the second system, with action R_2, and that the combined action $R_1 \times R_2$ makes $\mathcal{G}$ into a dynamical symmetry of the combined system: in this case, the symmetry group is *subsystem-global*. For instance, if we consider our Newtonian system of N point particles to be a subsystem of a larger system of particles, then the boost, translation and rotation symmetries are subsystem-global: by performing them on the whole of the combined system we still take solutions to solutions, but this is not the case if we, say, translate only the particles in the original subsystem but leave untranslated the particles in the new system with which they are interacting. (More generally, global symmetries in the usual physics sense — for instance, the global $U(1)$ symmetry of an ungauged complex scalar field — are subsystem-global.)

Thirdly, it might be that the action $R_1 \times \mathrm{id}$, where $\mathcal{G}$ acts trivially on the second system, makes $\mathcal{G}$ into a dynamical symmetry of the combined system. In this case, the symmetry group is *subsystem-local*. For instance, in the Newtonian-particle case, permutation symmetry is subsystem-local: a permutation of N identical particles is a symmetry whether or not those N particles are interact-

[10]In Wallace (2019a) I call this the 'Cosmological Assumption', and critique it more thoroughly.

ing with other particles. (For a more advanced example (cf Wallace (2019b)) in gauge theory the group of gauge transformations which vanish at the boundary of a given system is subsystem-local.)

In this third case, it will often be the case that $\mathcal{G}$ also acts on the second system as a subsystem-local symmetry. In this case the combined symmetry group of the total system will be $\mathcal{G} \times \mathcal{G}$: $\mathcal{G}$ can be applied, independently, to either system.

More complicated cases can occur, where different symmetries have different extendibility properties. Taking the second system as fixed, we can categorise individual symmetries as follows:

- g is *extendible* if there is a dynamical symmetry $\tilde{g}$ of the combined system which has the form $\tilde{g} = g \times g'$, i. e. which has a well-defined restriction to the original system that is equal to g.
- If extendible, g is *subsystem-local* if it has an extension $\tilde{g} = g \times \mathrm{id}$.

It is clear that the extendible symmetries form a group $\mathcal{G}_{ext} \subset \mathcal{G}$, and the subsystem-local symmetries form a subgroup $\mathcal{G}_{loc} \subset \mathcal{G}_{ext}$ of that group. We can then define the subsystem-global symmetry group as the quotient group $\mathcal{G}_{gl} = \mathcal{G}_{ext}/\mathcal{G}_{loc}$; elements of $\mathcal{G}_{gl}$ are equivalence classes of symmetries that differ by a subsystem-local symmetry, so that their action on the system is defined only up to a subsystem-local transformation. (In principle, this decomposition might be different when the system is coupled to a different system; in practice this does not normally occur, and for expository convenience I will usually ignore this subtlety.)

It should be clear that an inextendible symmetry has no implications at all for whether symmetry-variant quantities can be observed: by definition, as soon as an additional system is coupled to the original system, the symmetry is lost. By contrast, subsystem-local and subsystem-global symmetries have strong implications for what can and cannot be measured. For in both of these cases, by incorporating system and measurement device into a larger system, we can apply our previous results about measurements internal to a system.

Consider first the case of a subsystem-global symmetry. Using our orbit notation, we can write the combined state of target system and measurement device as $(O, g; O', g')$, with O and O' being orbits in the target system state space and the measurement-device state space respectively. The action of a symmetry transformation h is then

$$(O, g; O', g') \to (O, hg; O', hg'). \tag{11}$$

Importantly, since the symmetry acts simultaneously on the two systems, the symmetry-invariant information about the combined system is not exhausted by O and O' but also includes the relational quantity $g'^{-1}g$. If we now prepare the measurement device in some fixed 'ready' state (so that both O' and g' are held fixed, independent of O and g), then because this relational quantity covaries with the target system's variant quantity g, it is perfectly possible to measure g.

In this sense, variant properties of a system can after all be measured. However, because the measurement proceeds *only* via the relational quantity $g'^{-1}g$, any such measurement is always reinterpretable as a measurement of that relational quantity. If the measurement device had instead been prepared in the alternative ready state (O', hg'), then the measurement process would be a measurement of $h^{-1}g$ rather than of g — and there would be no further measurement to be done that could determine whether it had been so prepared, because that pair of measurements would jointly consist of a measurement of a $\mathcal{G}$-variant quantity of the combined system from within that system, which we have seen to be impossible.

Being able to measure $\mathcal{G}$-variant quantities in even this thin sense might seem counterintuitive, but a little reflection shows that it fits our everyday and scientific practice. The speedometer of my car (which operates via interactions between my car and its environment) does indeed measure my car's velocity, in the sense that different velocities give rise to different measurements; the GPS on my phone measures my location in just the same sense. But in both cases, the measurement can equally be described as a measurement of a relative quantity (my car's velocity relative to the road; my location relative to selected coordinates on the Earth), and in neither case can the measurement determine the quantity independently of the equivalent quantity for the Earth.

(Note also that nothing in this analysis turns on whether the information about the system is encoded in the variant or the invariant degrees of freedom of the measurement device; in either case, we can establish a functional dependence on a g-variant quantity of the target system; in both cases, that functional dependence is reinterpretable as a functional dependence on an invariant relation between system and measurement device.)

What if the symmetry is subsystem-local? Then things are simpler: now the symmetry-invariant information about the combined system is exhausted by (O, O'), and (by our previous results) no measurement can determine either g or g'. So quantities variant under a subsystem-local symmetry are unobservable.

We can extend this to more complicated symmetry groups straightforwardly. Taking the symmetry group $\mathcal{G}$ as fixed, a given physical quantity (formally: a given map from $\mathcal{S}$ to some other space X, such as $\{0, 1\}$ for properties or $\mathbb{R}$ for real-valued quantities) can be classified as

- *locally variant* if it is variant under the action of a subsystem-local element of $\mathcal{G}$;
- *globally variant* if it is invariant under the action of a subsystem-local subgroup of $\mathcal{G}$, but variant under the action of a subsystem-global element of $\mathcal{G}$;
- *invariant* if it is invariant under all extendible elements of $\mathcal{G}$.

We have now established a fairly precise version of the Unobservability Thesis:

- Invariant quantities are observable (or, at least, there is no symmetry-induced bar to observing them);
- Globally variant quantities are observable from a measurement device outside the system, but those observations are always reinterpretable as observations of an invariant relation between system and measurement device;
- Locally variant quantities are unobservable.

Applied to our core examples, we can see that variant spatial quantities like center-of-mass position, orientation, and center-of-mass velocity are observable, but only via processes reinterpretable as measurements of invariant spatial quantities; on the other hand, no measurement can distinguish two states that differ by a particle permutation. This is satisfactorily in accord with our pre-theoretic understanding in both cases.

I pause to make two observations. Firstly, our analysis does not assume that we know how to model a given measurement device. All it needs is the assumption that the measurement device, too, is bound by the symmetry. That assumption is fallible, of course — if we were to discover a set of physical processes that allowed us to probe a system but without conforming to some symmetry of the system, we could indeed measure symmetry-variant properties of that system (and not just in the thin sense discussed above). But this is an advantage of the analysis, not a weakness: it is indeed impossible to know infallibly whether a symmetry will turn out to be exact and universal, or just obeyed by some systems in some contexts, and the history of physics repeatedly displays examples of this. Consider parity, for instance: it is possible to measure the orientation of one system relative to another, but only a device governed by parity-violating interactions allows one to determine parity in a more absolute sense.[11]

Secondly, several recent authors (notably Roberts (2008), but see also Dasgupta (2016)) have noted that dynamical reasons prohibit us from encoding variant features of a system in its invariant features. But this generally has not been taken to resolve the question of unobservability, unless we make a supposedly question-begging assumption — after all, variant features can be encoded in other variant features. (Roberts, for instance, then appeals to anthropic features of the sort of beings we are to complete the story). My dynamical analysis shows that something more systematic can be said: for an isolated system there is in an important sense only one variant quantity (represented by g in my orbit-group notation) and it cannot be encoded in any way in any other independent quantity. It is true that if we have multiple systems, there are multiple independent quantities that are separately variant under the application of the symmetry to subsystems, but these quantities *can* be measured, it is just that the measurement is always reinterpretable as a measurement of an invariant relation between systems. For a collection of particles moving under Newtonian forces, for instance, the issue is not whether we can record the position of one

[11] There remain no subtleties in this case: see, e. g. , Huggett (2000), Pooley (2003), and Saunders (2007) for more discussion.

particle in the position of another: indeed we can, and it is useful to do so. The issue is rather whether the center of mass position of the whole system can be recorded in any independent degree of freedom; as we have seen, no such recording process is physically possible.

6 Surplus structure and representational equivalence

We can now give a fairly strong argument in favor of the Surplus Structure thesis — which, recall, is that symmetry-variant structure (e.g., an absolute rest frame) is surplus. For we have seen that no measurement process can determine whether or not a symmetry transformation has been performed on the combined system of target plus measurement device. Since the definition of symmetry-variant structure is that it is variant under such a transformation, it follows that there is no process by which it can be detected: no process, for instance, that can determine whether or not a system is moving *relative to absolute space* as opposed to moving relative to another system.

As a corollary, symmetry-variant structure can have no dynamical consequences for structure which can be detected: if surplus structure influenced the dynamical evolution of detectable quantities, it would itself be detectable indirectly, and our highly abstract analysis of measurement would pick this up.

So: consider two formulations of a theory — one containing symmetry-variant structure, one with that structure excised. (Here as usual, by 'symmetry' I mean 'subsystem-extendible dynamical symmetry'.) The second has all the resources of the first as regards describing and predicting the phenomena, which is to say that the symmetry-variant structures have no scientific work to do. There are now well-known arguments for taking the excised theory as the better representation of the physics: Ockhamist reasons based on the elimination of useless structure; semantic reasons based on the lack of any mechanism by which we could describe what the variant structure represents; pragmatic reasons based on the folly of including something demonstrably scientifically redundant in a scientific theory.

(Møller-Nielsen (2017) resists these arguments on the grounds is that there might in fact not exist a reformulation of the theory which eliminates the structure, so that we are stuck with the 'surplus' structure as the cost of using the theory at all. (One might think, for instance, that to move from a Newtonian to a Galilean spacetime requires mathematical creativity, and is not a guaranteed possibility merely because Newtonian spacetime has symmetry-variant structure.) But I think this underestimates the powerful, general resources by which structure can be substracted from mathematical theories. For instance, in an extremely wide class of spacetime theories, methods dating back to Klein allow the automorphism group of the theory to be widened to include whatever spacetime symmetry one wishes (see Wallace (2019d) and references therein; see also Dewar (2019) for a general defense of the legitimacy of this approach.). In the admittedly very abstracted setting of category theory, a class of mathematical models can be understood as *defined* by the transformations between models, and removing symmetry-variant structure is as simple as enriching the class of transformations. (See, e. g. , Weatherall (2016); note that this can be seen as a generalisation of the Kleinian method.) In any case, the concern is fairly theoretical: I am not aware of *any* theory in extant physics (even construing 'extant' fairly broadly) which does not have a well-understood reformulation in which dynamical symmetries and automorphisms coincide, even if one eschews Kleinian and categorical tricks and insists on a more purist conception of reformulation.[12])

If the surplus-structure thesis is accepted on these grounds, the Representational Equivalence thesis follows fairly directly. Suppose M and M' are symmetry-related models. Insofar as they are not isomorphic (with respect to the notion of methematical isomorphism appropriate to the class of models from which they are drawn), this can only be because there is surplus structure. Once that structure is excised, symmetry-related models are guaranteed to be mathematically isomorphic. And since the means by which mathematical models represent the physical is entirely structural (we have no other way to make reference to mathematical objects, even assuming a background philosophy of mathematics in which that would make sense) one model is representationally adequate exactly insofar as the other is.

It is tempting to forge on to the Modal Equivalence thesis: if symmetry-related models are representationally equivalent, how could a symmetry transformation represent anything other than a mere redescription? We will see this is far too quick, again for reasons that follow from the fact that physical theories are virtually always used to describe isolated systems in a larger universe, not the Universe as a whole.

7 Modal equivalence and the intrinsic/extrinsic distinction

Does a symmetry transformation, applied to a system, bring about a different possibility? The most common way to understand that question (see, e. g. , Belot (2018)) is as a thesis about the entire Universe: if we apply a symmetry transformation to a model of the Universe as a whole, does that bring about a new possibility? Understood that way, it seems a rather abstract question, turning as much on one's general theory of modality as anything else: conventionalists about modality, for instance (like Belot) could argue that this is partly

[12] The nearest I know to a counter-example (which lies well outside 'extant physics') occurs in supersymmetry, a theorised extension to current quantum field theory: standard '$\mathcal{N}$=1' supersymmetry has a reformulation on 'superspace' which is analogous to the reformulation of special relativity on Minkowski spacetime, but the superspace move does not appear to be available for so-called 'extended' supersymmetry. The issue has not been extensively explored; see Wess and Bagger (1983) and references therein for general features of supersymmetry, and Menon (2018) for some preliminary philosophical considerations.

a sociological point, to be answered by looking at the way physicists actually in fact use the concept. But in accordance with the themes of this paper, I wish to answer it more locally, and more naturalistically, by considering symmetry transformations on isolated subsystems of a larger universe. In this case (which is, in any case, the *actual* context in which physicists use the concept), if the symmetry transformation is subsystem-global then the answer is quite straightforwardly yes. In the notation of the previous sections, if the first system has state (O, g) and the symmetry is h, the transformed system has state (O, hg). Given any second system with state (O', g'), the result of the symmetry is to transform the invariant quantity $g^{-1}g'$ to $hg^{-1}g'$ — and this transformation is directly observable, and so clearly indicates that we are dealing with a new possibility.

How is this to be reconciled with the Representational Equivalence Thesis? As an illustration, consider our standard particle-mechanics example, and consider two systems (say, with N and M particles). Mathematically speaking, an instantaneous configuration of the first system is an ordered set of N points in Euclidean space E; an instantaneous configuration of the second is an ordered set of M points in another copy E' of Euclidean space. To compare the two systems (to check, for instance, whether they are sufficiently spatially separated that we can treat them as isolated) we need to identify E and E' — that is, specify an isomorphism between them — and there is no preferred or canonical way to do so. So given configurations q, q' of the systems separately, we have not been given enough information to describe their joint configuration: that requires, in addition, a *representational convention* as to how points in the two configuration spaces are to be compared. Such a convention is inevitably required whenever we combine subsystems into a joint system. (In practice, the convention is often given by a choice of coordinate systems, and/or of reference frames, in the two subsystems.)

Prior to stipulating any such convention, there is no sense in which (q, q') specifies a different joint configuration from $(R(g)q, q')$, since q and $R(g)q$ are representationally equivalent. Given a choice of representational convention, though, it is clear that applying the symmetry transformation to one system gives rise to a different total configuration (and that this is true independent of what the actual representational convention is). So: symmetry-related configurations can be understood as representing different possible configurations *if we hold fixed the choice of representational convention.*[13]

Introducing some metaphysical terminology: we have established, in effect, that the *intrinsic* properties of a system are invariant under symmetry transformations of that system: these properties do not depend on relations with any other systems and so representational equivalence entails equivalence of intrinsic properties. A subsystem-global transformation brings about a change in a system's *extrinsic* properties. And this in turn gives some insight into what we are actually saying, physically, about a system when we establish that a certain transformation is a symmetry: we are identifying the fact that some properties

[13]Thanks to Oliver Pooley for this observation.

of the system are extrinsic properties. The relativity principle, for instance, can be understood physically as telling us that velocity is an extrinsic property of any system. (The idea that symmetries leave intrinsic properties unchanged is far from new, and is often taken as *definitional* of a symmetry: see, e. g., (Healey 2009).)

Although I have carried out this analysis at the level of configurations, it applies *mutatis mutandis* at the level of states or entire histories of the system (with the caveat in the latter case that dynamically possible histories of subsystems can only be combined into dynamically possible histories of the combined system insofar as the subsystems are non-interacting).

What if the symmetry-transformation is subsystem-local? In this case, the transformation can be extended to a symmetry transformation which acts trivially on any other systems — which is to say that the symmetry transformation leaves invariant the intrinsic features not just of the system being transformed but of any larger system within which that system is embedded, up to the entire Universe, in principle. This seems to suggest that subsystem-local transformations bring about no physical change in a subsystem, and that features of the system that are variant under a subsystem-local symmetry do not represent physical features of that system.

As long as we are working at the level of histories, this seems correct: at any rate, it is difficult to see what naturalistic work is being done by a notion of modality that treats histories related by a subsystem-local symmetry as physically distinct, and it is unclear how to understand the possibility of a representational convention for the Universe as a whole which could allow us to reconcile modal inequivalence with representational equivalence. (In metaphysical terms, any new possibility would be purely haeccistic.) But matters are much subtler if we consider applying the transformation at the level of configurations (or, mutatis mutandis, states): at that level, there is a different kind of modal question we can ask for which subsystem-local transformations do real physical work. Namely: given that at time t the system has configuration q_0, and given two configurations $q, R(g)q$ related by a symmetry transformation, do these represent different configurations that the system can evolve into, or redescriptions of the same configuration?

Here too the intrinsic/extrinsic distinction is useful. The intrinsic features of a configuration are the symmetry-invariant features; symmetry-variant features might be understandable as extrinsic features dependent on the relation with configurations of *other systems*, or as dependent on the relation with configurations of the *same* system at different times. If two particles are in circular orbits around one another, for instance, the intrinsic features of the system are time-invariant, but we can understand the system as changing *either* in the sense that its orientation relative to other systems is changing, *or* in the sense that its orientation at one time is different from its orientation at another.

We can get further insight into this by returning to the orbit/group analysis, this time applied to configurations. Given a group $\mathcal{G}$ of dynamical configuration symmetries, we can represent a configuration as a pair (o, g) where now o is an orbit in configuration space (here again we suppose that we pick a representive point $q(o)$ in each orbit o, so that any point can be written as $gq(o)$ for some (o, g)). A point in *state* space can then be represented as $(o, g; \dot{o}, \dot{g})$, but this representation does not fully separate variant from invariant features of the state, because the quantity $g^{-1}\dot{g}$ is invariant under the group action.[14] The quantity $\xi = g^{-1}\dot{g}$ is called the *body velocity* in the physics literature, and can be understood physically as representing the rate of change of the system's variant features not as measured by an external observer, but as measured by a reference frame embedded in the system itself.

An orbit in the state space is represented by a triple $(o, \dot{o}, \xi)$, comprising the intrinsic features of the system's configuration, the rate of change of those intrinsic features, and the body velocity. In this notation (and continuing to assume deterministic dynamics), section 4's dynamical analysis can be rewritten as

$$\ddot{o}(t) = \mathcal{F}_O(o(t), \dot{o}(t), \xi(t); g(t); t) \tag{12}$$

$$\dot{\xi}(t) = \mathcal{F}_G(o(t), \dot{o}(t), \xi(t); g(t); t) \tag{13}$$

$$\dot{g}(t) = g(t)\xi(t) \tag{14}$$

for some functions $\mathcal{F}_O, \mathcal{F}_G$. (In the physics literature this is known as *dynamical reduction*, or sometimes as *Marsden-Weinstein reduction*; see Marsden and Ratiu (1999) and references therein for details.)

Returning to the case of relations between configurations at different times, notice that once we have solved the equations of motion for the intrinsic degrees of freedom, we can go on to solve (14) to determine $g(t)$: that solution will have the form $g(t) = h(t)g(0)$ where $h(t)$ depends only on the intrinsic features between times 0 and t. $h(t)$ gives a dynamically determined relation between $g(0)$ and $g(t)$; given $g(0)$ and this relation, $g(t)$ is fixed for any other time. So the extrinsic features of the system at any time can be read off from the intrinsic features of the system's history and the extrinsic features at any one other time — just as the extrinsic features of a subsystem can be read off from the intrinsic features of the system and the extrinsic features of any other subsystem. Let's call extrinsic features of this kind *temporally extrinsic*, as distinct from the *subsystem-extrinsic* features that arise from considering a system's relation to other systems.

(If the symmetry is discrete rather than continuous – as in the case of permutation symmetry — the analysis goes slightly differently, because now $\dot{g}$ is not meaningful. The configuration space can still be decomposed into orbits, but the pairs $(o, \dot{o})$ fully describe orbits in state space and there is no body velocity. But it remains the case that given a solution to the intrinsic dynamics, we can determine some group element $h(t)$ such that if $g(0)$ represents the system's variant features at time 0, $g(t) = h(t)g(0)$ represents them at time t.)

Should we regard temporally extrinsic features of a system's state as physically meaningful? — or, put another way, are two states that differ only with

[14]Technical note: strictly $\dot{g}$ is a point in the tangent space over g, and $g^{-1}\dot{g}$ is the left translation of that point back to the tangent space over e.

respect to their temporally extrinsic features different representations of the same state, or representations of different possible states? The question is moot as long as we are considering subsystem-global symmetries, because in that case, any temporally extrinsic feature is also subsystem-extrinsic. But it is significant when we consider subsystem-local symmetries. For instance, given a particle permutation π, consider configuration-space trajectories $q(t)$ and $q'(t)$ where for some t_0 we have $q(t_0) = q'(t_0)$, and for some t_1, $q'(t_1) = \pi q(t_1)$. In this case, there will be no symmetry relating the two trajectories, and we can meaningfully regard $q'(t_1)$ and $q(t_1)$ as extrinsically different even though intrinsically identical.

Should we then regard πq and q as *different* possible configurations? The question seems largely conventional: it is, after all, up to us how we define configuration space (or phase space). If we want to quotient out configuration space (which amounts to working with *un*ordered N-tuples of points in Euclidean space) we can do that, albeit at the cost of making the topology a bit awkward; if we want to continue working with the unquotiented configuration space, so that points in the space represent temporally extrinsic as well as intrinsic features of configurations, we can do that too. The modal facts are in any case unambiguous at the level of entire histories, so I find it difficult to see any way of making this a matter of convention-independent fact. (We will see in section 9 that the case for this apparent conventionality is sharpened in quantum theory.)

To summarise our results so far:

If $\mathcal{G}$ is a group of non-extendible dynamical symmetries then no conclusions about observational, representational, or modal equivalence follow from the symmetry.

If $\mathcal{G}$ is a group of extendible dynamical symmetries then

- $\mathcal{G}$-variant features of a system are unobservable from within that system.
- As a corollary, $\mathcal{G}$-variant mathematical structure in a theory is surplus and symmetry-related structures are representationally equivalent.

If in addition $\mathcal{G}$ is a group of subsystem-global dynamical symmetries then

- $\mathcal{G}$-variant features of a system can be measured from outside that system, but any such measurement can be reinterpreted as a measurement of a $\mathcal{G}$-invariant relation between system and measurement device.
- A symmetry transformation leaves the intrinsic features of a system invariant but changes its system-extrinsic features.
- Symmetry transformations bring about new possibilities.

If instead $\mathcal{G}$ is a group of subsystem-local dynamical symmetries then

- $\mathcal{G}$-variant features of a system are unobservable

- A symmetry transformation of an entire system (at all times) does not bring about a new possibility, and does not change any physical features of the system.
- A symmetry transformation of a system's state at a given time changes the temporally extrinsic features of the system (while leaving its intrinsic features unchanged); it appears to be conventional whether symmetry-related states represent different possibilities.

In the case where $\mathcal{G}$ contains a system-local subgroup $\mathcal{G}_{loc}$, so that we can define the system-global symmetry group as $\mathcal{G}_{gl} = \mathcal{G}/\mathcal{G}_{loc}$, and define its action on the system uniquely up to a system-local symmetry:

- $\mathcal{G}_{loc}$-variant features of a system are unobservable; features that are $\mathcal{G}_{loc}$-invariant but variant under a subsystem-global symmetry can be measured from outside that system, but any such measurement can be reinterpreted as a measurement of a $\mathcal{G}$-invariant relation between system and measurement device.
- A system-local symmetry transformation of an entire system (at all times) does not bring about a new possibility, and does not change any physical features of the system; a system-global symmetry transformation (defined only up to a system-local symmetry) changes its system-extrinsic features but not its intrinsic features.

In the remainder of the paper I consider complications to this picture which arise from more advanced examples of symmetries, due (respectively) to time-dependent symmetries, to quantum theory, and to cosmology. The technical level of the next three sections is significantly higher than for the rest of the paper; they can be skipped on first reading without loss of continuity.

8 Time-dependent symmetries

Recall our original definition of a dynamical symmetry, expressed with respect to the configuration-space formulation of classical mechanics: it is specified by a group $\mathcal{G}$ and an action R of $\mathcal{G}$ on the configuration space $\mathcal{Q}$, such that $R(g)q(t)$ is a dynamically possible history iff $q(t)$ is. A *time-dependent* symmetry is again specified by a group and action, but now we require that for any smooth path $g(t)$ in $\mathcal{G}$ itself, $R(g(t))q(t)$ is dynamically possible iff $q(t)$ is. Or put more colloquially : if the symmetry is time-dependent then we can apply it independently at different times, subject only to the requirement that the symmetry group applied changes smoothly with time. We will assume for the moment that all elements of $\mathcal{G}$ are *small*: which is to say, each element is connected by a smooth path to the identity e. (The significance of this requirement will emerge shortly).

If a symmetry is time-dependent, it follows that the dynamics for the theory is (formally) indeterministic. For suppose that $g(t) = e$ for $t < t_0$ but $g(t) \neq e$ thereafter. Then if $q(t)$ is a dynamically permissible trajectory, so is $R(g(t))q(t)$

— but these two histories are identical up to time t_0 but diverge afterwards. As such, insofar as the theory's dynamics are specified by a differential equation, that equation must admit multiple solutions.

(For an elementary discussion of time-dependent symmetries, with examples and references, see Wallace (2003).)

The notions of extendibility of symmetries, and of subsystem-local and subsystem-global symmetries, generalise straightforwardly to the time-dependent case. Generalising our analysis of measurement to indeterministic symmetries is a little delicate since there is no unique equation of motion, but a natural extension is to say that one system measures another if the *set* of possible post-measurement values of the measuring-device state is functionally dependent on the pre-measurement value of the target-system state. Then our previous results again generalise: subsystem-global symmetries are measurable, subsystem-local symmetries are not.

In practice, subsystem-locality and time dependence are closely related: to the best of my knowledge, there are no physically relevant examples of a time-dependent, subsystem-global symmetry group.[15] (The reason, in a nutshell, is relativity, which forces transformations of the same system at different times to be closely related to transformations of different systems at the same time.) The interesting case is time-dependent, subsystem-local symmetry.

Our previous analysis showed that state quantities variant under a time-independent symmetry could always be understood as extrinsic features of the state, representing symmetry-invariant relations either with other systems, or with states of the same system at different times (with the latter being the only available possibility if the symmetry was subsystem-local). In general this understanding is not available for quantities variant under a time-dependent symmetry: given an arbitrary $g_0 \in \mathcal{G}$, and an arbitrary time t, we can find $g(t)$ such that $g(t_0) = g_0$, and $g(t) = e$ except for values of t within some arbitrarily short distance of t_0. So just as a subsystem-local transformation of a system can be extended to a transformation that equals the identity on any other given system, so a (small) time-dependent symmetry of a *state* can be extended to a transformation that equals the identity on any other given state. For this reason, we can usefully call elements of a small time-dependent symmetry-group *time-local*.

I have already argued that when a subsystem-local transformation is applied to an entire history, there is no good reason to regard that transformation as bringing about a new possibility. Since any time-local transformation of a state can be extended to a transformation of the entire history that equals the identity except in an arbitrarily small neighbhorhood of the state, it follows directly that time-local transformations of *states* also fail to give rise to new histories (and hence the formal indeterminism associated with these symmetries does not entail any physical indeterminism).

[15] The closest I know to examples are the variant theories of Newtonian mechanics developed by Barbour and Bertotti (1982) and Saunders (2013), but (for reasons I develop in more detail in Wallace (2019a)) neither can be interpreted as representing a subsystem of a larger system.

We can also understand this through dynamical reduction. Given a system with configuration-space history $q(t) \simeq (o(t), g(t))$, a time-dependent symmetry $h(t)$ transforms it to $(o(t), h(t)g(t))$ — which is to say that the symmetry-variant part of the system's history is completely unconstrained by the dynamics. As a consequence, we can consistently reduce the dynamics to a dynamics on the space of orbits, with some differential equation

$$\ddot{o}(t) = \mathcal{F}(o(t), \dot{o}(t); t). \tag{15}$$

(If symmetry is the only cause of indeterminism, this equation will be deterministic.) So the symmetry-variant properties of a system have no dynamical effect either on the symmetry-invariant properties, or even on the symmetry-variant properties at later times. They are, in the fullest sense, dynamically redundant.

The requirement for symmetries to be small is doing essential work here, as the above arguments have required us to construct smooth paths from e to whatever group element we were considering. They fail for large symmetries (those which cannot be so connected to the identity); two states variant under a large element of a time-dependent symmetry group are not time-local, and quantities variant under such symmetries (but invariant under the small part of the symmetry group) can be understood as time-extrinsic just as for time-independent symmetries. Important examples arise in non-Abelian gauge theories (I discuss them further in Wallace (2019b); see also Teh (2016)).

All of this has a gratifyingly direct fit in the constrained Hamiltonian formalism (cf Earman (2003), Teh (2016), and references therein). There, the phase-space dynamics is deterministic only up to *small* time-dependent symmetries, which are generated by the constraints; there, 'observables' are required to be invariant under such symmetries (but not necessarily under large elements of the symmetry group); there, phase-space points related by a small symmetry are normally interpreted as physically equivalent.

9 Symmetries in quantum theory

In the Hilbert-space formulation of quantum mechanics, a dynamical symmetry is a 1:1 transformation of Hilbert space which leaves all transition amplitudes invariant.[16] By Wigner's theorem, any such transformation must be a unitary or anti-unitary operator; writing it as $\widehat{R}$, the condition is that if $\widehat{U}(t;t_0)$ is the time evolution operator for an isolated system which has $\widehat{R}$ as a dynamical symmetry, then $\widehat{R}$ commutes with $\widehat{U}(t;t_0)$. (Any such transformation has the property that it maps solutions to solutions, but the converse is not true. In quantum theory, preserving solutions is too coarse a notion: we want to preserve the full structure of transition probabilities.)

A dynamical symmetry *group* is then a group $\mathcal{G}$ together with an action $g \to \widehat{R}(g)$ of $\mathcal{G}$ by means of (anti-)unitary operators, such that each $\widehat{R}(g)$ is a

[16]In fact, it suffices for the transition amplitudes to be invariant up to phase. Also, and as with the classical case, I simplify slightly by leaving out velocity boosts and time translation.

dynamical symmetry. The notions of an extendible symmetry, and of subsystem-global versus subsystem-local symmetries, carry over *mutatis mutandis*.

The analysis of sections 4-5 was extremely abstract — so abstract that it can be carried directly over to quantum theory. So given an isolated quantum system with a symmetry, that symmetry is unmeasurable from within the system; there is no in-principle difficulty in measuring it from a different system, even if the symmetry is extendible, but that measurement will always be reinterpretable as a measurement of some relational quantity. There are, however, interesting subtleties that arise from the richer structure of quantum theory.

To explore further, note that if a measurement is of some quantity invariant under a symmetry, quantum-mechanically that means that the operator representing the measurement must be invariant under the adjoint action of the operator representing the symmetry. So given a symmetry group $\mathcal{G}$ with action $g \to \widehat{R}(g)$ as above, any internally-measurable observable $\widehat{X}$ must satisfy

$$\widehat{R}(g)\widehat{X}\widehat{R}^{-1}(g) = \widehat{X} \tag{16}$$

for all $g \in \mathcal{G}$, and so given a quantum state ρ, the expectation value of that measurement satisfies

$$\langle\widehat{X}\rangle_\rho \equiv \mathsf{Tr}(\widehat{X}\rho) = \mathsf{Tr}(\widehat{R}(g)\widehat{X}\widehat{R}^{-1}(g)\rho) = \mathsf{Tr}(\widehat{X}\widehat{R}^{-1}(g)\rho\widehat{R}(g)) \tag{17}$$

which is to say that ρ can be replaced by $\widehat{R}^{-1}(g)\rho\widehat{R}(g)$ without affecting the result of any (internally) measurable quantity. If we now integrate over all such group actions, using the invariant Haar measure $\mu_\mathcal{G}$[17] over $\mathcal{G}$, we can define

$$\rho_\mathcal{G} = \int_\mathcal{G} \mathrm{d}\mu_\mathcal{G}\, \widehat{R}^{-1}(g)\rho\widehat{R}(g) \tag{18}$$

which, following Bartlett, Rudolph, and Spekkens (2007), we can call the $\mathcal{G}$-twirl of ρ. Then any measurement of ρ, provided it is physically possible and performed within the system, is equivalent to the same measurement of $\rho_\mathcal{G}$.

Now, recall that the action of a group on a Hilbert space can always be decomposed into *irreducible representations* (irreps), so that if $\widehat{R}_I$ is the Ith irrep and acts on Hilbert space $\mathcal{H}_I$, the total Hilbert space can be decomposed as

$$\mathcal{H} = \bigoplus_I \mathcal{K}_I \otimes \mathcal{H}_I \tag{19}$$

with $\mathcal{G}$ acting trivially on each $\mathcal{K}_I$: that is, the overall group action is

$$\widehat{R}(g) = \sum_I \mathrm{id}_I^K \otimes \widehat{R}_I(g) \tag{20}$$

where id_I^K denotes the identity on $\mathcal{K}$. $\rho_\mathcal{G}$ can be written in this notation as

$$\rho_\mathcal{G} = \sum_I \rho_I \otimes \mathrm{id}_I^\mathcal{H} \tag{21}$$

[17]Strictly this is only defined for compact $\mathcal{G}$, so these results need to be treated with some care if applied to non-compact groups such as translation.

where ρ_I is a density operator on $\mathcal{K}_I$ and now $\mathrm{id}_I^{\mathcal{H}}$ is the identity on $\mathcal{H}_I$. As long as we confine ourselves to internally-performable measurements, then, the quantum state is empirically equivalent to a probabilistic mixture of states defined on the $\mathcal{G}$-invariant spaces $\mathcal{K}_I$.

In other language, all of this is to say (i) that as far as internal measurements are concerned, symmetries effectively define superselection rules, and (ii) that within each superselection sector the quantum state is represented for all empirical purposes by one of the invariant states ρ_I.

As long as we are considering subsystem-global symmetries, this analysis has largely practical consequences: it tells us that when analysing a measurement, we can treat the symmetry as superselected with respect to the combined system of target system plus measurement device, which interestingly constrains what measurements are possible. A very early result in this field, the WAY theorem (named for Wigner (1952) and Araki and Yanase (1960); see Busch and Loveridge (2013) and references therein for modern presentations) established that no completely sharp measurement of a quantity was possible unless it commuted with all additive conserved quantities, and showed that the degree of accuracy that could be obtained by an unsharp measurement depended on the initial degree of asymmetry of the measurement device. More recent work (see Bartlett, Rudolph, and Spekkens (2007) and references therein) has quantified this result and developed it into a general resource theory of measurements, where the resource is the initial asymmetry. (Note the crucial role of the assumption that the conserved quantity is *additive*, which in our terminology is equivalent to it being extendible.)

The consequences are more dramatic for a subsystem-local symmetry, for then our analysis entails that a superselection rule with respect to that symmetry effectively applies to any system even as regards measurements made outside the system. Consider permutation symmetry, for instance: no measurement, including from outside the system, can distinguish the quantum state of N identical particles from the state obtained by $\mathcal{G}$-twirling that state. So the system's state is empirically equivalent to a probabilistic mixture of states each of which transforms under some irrep of the permutation group — and which, furthermore, is invariant under the action of that irrep. At this point, normal physics practice would be to say that the actual system is in one or other irrep, and the probability distribution over them represents our ignorance of the actual irrep. In the case of permutation, the completely symmetric and completely asymmetric irreps represent, respectively, bosonic and fermionic statistics; the higher-dimensional, mixed-symmetry irreps represent parastatistics.

In light of these results, let us return to the discussion of section 5 on the redundancy of subsystem-local symmetries. There I argued that it seems to be largely conventional whether we quotient the state space of a system by a subsystem-local symmetry: it depends whether we want the state space to represent the extrinsic relation between states at different times which differ by the action of the symmetry, or just to represent the intrinsic properties at each time. In quantum theory, quotienting amounts to replacing the Hilbert space of the system with the direct sum of the invariant $\mathcal{K}_I$, with the additional stipulation that coherent superpositions of states in different subsystems have no physical meaning: put another way, the actual Hilbert space of the quotiented system is *one* of the $\mathcal{K}_I$. But given our superselection results, this is really just the question of whether to take the state to be ρ_I or $\rho_I \otimes \mathrm{id}_I^{\mathcal{H}}$. Even a fairly modest structuralism in our metaphysics will make this a distinction without a difference. (In the specific case of permutation symmetry, this underdetermination has been extensively discussed in the literature: see French (2015) and references therein.)

What is *not* conventional is the fact that two quantum states at time t which are intrinsically identical but which differ extrinsically via a symmetry transformation can both be reached from the same quantum state at some earlier (or later) time t_0, and that both possibilities contribute coherently to the overall transition amplitude. The dynamical significance of the choice of irreducible representations is that it determines any relative phase factors between the contributions. This is particularly clear if we restrict attention to Abelian symmetries (or to Abelian representations of symmetries, such as the symmetric or antisymmetric representations of the permutation group). Let x be some initial configuration, and let y and y' be configurations at some later time, related by the action of the symmetry. If we quotient out by the symmetry, y and y' are mapped to the same point, but two trajectories $x \to y$ and $x \to y'$ will not be mapped to the same trajectory, so the action for each must be included in the path integral — and in formulating the dynamics, we always have the possibility of adding some topological phase factor to their relative contributions. If we do not quotient out, then the two trajectories correspond to different transition amplitudes $\langle x| \widehat{U} |y\rangle$ and $\langle x| \widehat{U} |y'\rangle$ — but because the final state transforms under an irrep of the symmetry, it will be a coherent superposition of $|y\rangle$ and $|y'\rangle$ with equal weight, differing only by a phase factor. So the total transition amplitude between physical states will again involve a coherent sum of contributions from the two paths — only this time, the relative phase is determined by the particular representation (-1 for fermions, +1 for bosons, for instance).[18]

I end this discussion with two brief remarks in more technical areas of physics, which may be skipped by readers unfamiliar with the relevant material.

1. The physical significance of a subsystem-local symmetry is contained (I have argued) in the existence of inequivalent trajectories between intrinsically identical initial and final states, which can be understood as an extrinsic property of those states. This takes a subtle form in the case of identical particles in two dimensions (or in systems like thin films, which can be approximated as two-dimensional), because two trajectories might be topologically inequivalent not just because they differ by a particle permutation, but because they differ in the way in which the permutation was implemented over time: in two dimensions, trajectories often cannot

[18] The path-integral analysis for non-Abelian groups — for parastatistics, for instance — is decidedly subtler, and indeed at one time (Laidlaw and Morette DeWitt 1971) it was argued that parastatistics could not be incorporated in the path-integral formalism. See Greenberg and Mishra (2004) and references therein for further details.

be deformed into one another without crossing. The concrete implication is that it is the *braid group*, not the permutation group, which determines the statistics of identical particles in two dimensions: this leads to the phenomenon of *anyons*. See Stern (2008) for a review.

2. In non-Abelian gauge theory (specifically, in QCD) there is a subsystem-local (but non-small) symmetry generated by topologically non-trivial, but boundary-vanishing, gauge transformations. These transformations form a group isomorphic to the integers, so that we can write an element of the group action as $\widehat{R}(n)$, with $\widehat{R}(n+m) = \widehat{R}(n)\widehat{R}(m)$; a state $|\varphi\rangle$ in an irrep satisfies
$$\widehat{U}(n)\,|\varphi\rangle = \mathrm{e}^{-in\theta}\,|\varphi\rangle \tag{22}$$
for some angle θ, so that the irreps are labelled by this θ, and in the path integral, trajectories that terminate on states differing by the action of $\widehat{R}(n)$ pick up a relative phase factor $\exp in\theta$, which has empirical consequences. Experiment suggests that it is equal to zero to very high accuracy, and the so-called *strong CP problem* is the problem of identifying why this is.

   It has been argued (Fort and Gambini 1991; Healey 2010) that if we worked in the quotiented configuration space right from the start, the strong CP problem would dissolve; but this seems overstated. The topologically distinct paths would still contribute to the path integral and we would still have the option of adding a θ phase term. Of course, we could decide not to, but equally, we could decide to set $\theta = 0$ in the non-quotiented formalism. The real issue with $\theta = 0$, in either formulation, is that in quantum field theory, quantum corrections would be expected to introduce a large non-zero θ term even if it is omitted initially: obtaining $\theta = 0$ at the level of phenomenology requires a very large, and very precisely chosen, value of θ in the bare parameters of the theory. (It is a highly contested issue whether *naturalness* arguments like this should really be seen as persuasive: I make the case that they should in Wallace (2019c).)

10 Symmetries of the whole Universe

I have argued in this paper that interpretatively there is a fundamental distinction to be drawn between symmetries based on if, and if so how, they can be extended to other systems. But what if — as is extremely common in philosophical writings — we are concerned with symmetries of the entire Universe, in which case the distinction seems to evaporate?

I am highly tempted to dismiss this question as premature. Even in cosmology, we are hardly ever working with a model intended to describe the *entire Universe* (as opposed to, say, some large patch of it); even in quantum gravity, very large additional problems occur when we move from considering some isolated system (like a black hole) to considering the entire universe. Indeed, at present many high-energy physicists would claim that we have a quantum satisfactory theory of gravity, via the AdS/CFT correspondence (and its interpretation as describing a system in a box; cf discussion and references in Wallace (2017), but there is no comparable case within that research program that we have a quantum theory of *cosmology*. Modelling the whole Universe raises conceptual and technical problems that do not occur in normal uses of physics and we have very little empirical evidence to help us resolve them.

But suppose we throw caution to the winds and assume (following philosophy-of-physics orthodoxy) that at least some of our current theories have physically-significant interpretations in which their models are models of the whole Universe. At this point, the key result of section 4 — that no measurement internal to a system can distinguish between symmetry-related states of that system — applies to any dynamical symmetry. (This assumes, of course, that we are studying a theory rich enough to represent scientists, measurements, and observations inside its own structure. Given a possible world consisting only of two point-particles interacting via a central force, for instance, there is nothing like the structure to permit us to model measurement, and so no conclusion that can be drawn about what is observable. If the reader nonetheless wants to know how to draw epistemic and/or metaphysical conclusions about that (supposedly) possible world from its dynamical symmetries, they will have to do it on their own time: I have no intuitions about the situation, and no inclination to trust any such intuitions anyway.)

Assuming our symmetries are time-independent, the discussion in section 5 on subsystem-local symmetries now applies with equal force to subsystem-global symmetries: in an N-particle Newtonian theory interpreted as a theory of the whole Universe, for instance (and ignoring severe concerns about the modellability of measurement in any such theory) then configurations at different times will have a symmetry-invariant and empirically determinable relative orientation even while the overall orientation of the system remains undetectable. But by the same token, the considerations of that section suggest that possible histories differing by a symmetry transformation ought to be regarded as redescriptions of the same history; or, at any rate, if they are not to be so regarded it must be for metaphysical, not naturalistically-motivated, reasons. (This is of course the original context of the 'Leibniz shift".)

In the quantum context, this would imply that any symmetry of the whole Universe — even a global one — ought to be regarded as defining a superselection rule. Applying this to the Hamiltonian itself (i. e., to time translation symmetry) has radical implications for dynamics in any such universe, a problem explored extensively in the quantum gravity literature (the classic discussion in Page and Wootters (1983)) and related to (but distinct from) the famous 'problem of time'.

But we may be getting ahead of ourselves, for it is not at all clear that we should expect a truly cosmological theory to have any such symmetries. Modern physics provides (admittedly speculative) grounds to think that global symmetries in this sense are not exact and/or do not apply to the Universe as a whole:

1. The apparent exact global internal symmetries of particle physics (e.g., the symmetry underlying baryon-number conservation) are so-called 'accidental' symmetries (see, e. g. , Duncan (2012, pp.490-1)) which occur as a consequence of renormalisability but we have no reason to expect to hold exactly when non-renormalisable corrections to the dynamics are added.

2. Quantum gravity offers some further reasons to expect global symmetries not to be exact: the only quantities conserved in black hole evaporation are those associated with long-range forces, which in high-energy physics corresponds to those associated with asymptotic local symmetries. The lack of exact global symmetries is also provable perturbatively in string theory and can be argued for via the AdS/CFT correspondence: see Banks, Johnson, and Shomer (2006) and references therein for details.

3. Although local gauge symmetries (and the Poincaré gauge symmetry underlying general relativity, are believed to be exact, and give rise to subsystem-global symmetries through asymptotic effects, those asymptotic effects rely on boundary conditions 'at infinity', which are normally interpreted to describe the way in which a system embeds into a larger system and which are not applicable in cosmology. The only consistent way we currently know how to avoid such boundary conditions is to assume that space is compact, in whch case there are no asymptotic symmetries.

In the end, I return to the caution with which I began this section: we just do not know, reliably, how to think about theories — and, *a fortiori*, about symmetries of those theories — when the theories are genuinely being interpreted cosmologically. Nor do we need to in order to understand the role of symmetries in non-cosmological (i. e. , almost all) applications of our theories.

11 Conclusion

Dynamical symmetries place constraints on observability because, and to the extent that, they apply to the physical processes by which those observations are made. Although this implies fairly straightforwardly that symmetry-variant properties of a system cannot be observed by measurement processes confined to the system itself, in practice we very rarely model a measurement process explicitly in physics, and so the implications of a symmetry of a system for observations made on that system depends on whether, and how, the symmetry can be extended to other systems, including the measurement device. If the symmetry cannot be so extended, it has no particular consequences for observation; if it is subsystem-local, so that it can be extended to a symmetry which acts trivially on any other system, it is unobservable; if it is subsystem-global, so that it can be extended to a symmetry that includes the measurement device but which does not act trivially on it, then the action of the symmetry on the system alone has measurable consequences, but these can be reinterpreted as measurements of the symmetry-invariant relations between system and measurement device.

These results strongly suggest that symmetry-invariant features of a system should be interpreted as intrinsic features, that features that are variant under subsystem-local symmetries should be interpreted as descriptively redundant, and that features that are invariant under subsystem-local symmetries but variant under subsystem-global symmetries should be interpreted as extrinsic features. It follows that subsystem-local symmetry transformations should not be regarded as bringing about different states of affairs, but that subsystem-global symmetry transformations should be so regarded.

These various results extend, albeit with some subtleties, to the physically important cases of time-dependent and quantum-mechanical symmetries. They do not extend to theories which are intended to be understood as cosmological, in which the local/global distinction collapses, but cases of this sort are extremely unusual in physical practice and there are strong, if speculative, reasons to think that a genuinely cosmological physical theory should have no global symmetries. In any case, the cosmological case only strengthens the general point that normal modelling practice in physics is to treat an isolated system as modelling a small part of a larger universe, and that the epistemology and metaphysics of symmetry in physics is much clearer once this is properly allowed for.

The results of this paper go some way towards seeing how genuinely substantive intepretative results can follow from a purely formal concept of symmetry (and do not require us to abandon that formal concept, as is the case for some non-dynamical definitions of symmetry). But a puzzle remains: my analysis turns on how a symmetry can be extended beyond a system to other systems in dynamical contact with it, and definitionally (it would seem) no formal feature of a system's symmetries can fix this question of extension to other systems. If so, then once again we seem to need to make substantial interpretative assumptions in order to get interpretative results from a theory's symmetries, and to abandon the hope of extracting such results from a formal conception of symmetry.

This problem cannot be completely solved: nothing about a theory, in isolation, can rule out the possibility that new physics, applying to new systems, will break a symmetry of the old system once the two are dynamically coupled. Nonetheless it is still possible to draw tentative conclusions about the extendibility of a theory's symmetries from the theory itself, once it is recognised that (in many cases) physical theories themselves contain rich information about their subsystems and the restriction of symmetries to those systems. I develop this observation in depth in the companions to the present papers (Wallace 2019a; Wallace 2019b).

Acknowledgements

I am grateful to Gordon Belot, Harvey Brown, Neil Dewar, Hilary Greaves, Eleanor Knox, Oliver Pooley, Laura Ruetsche, Simon Saunders, Nic Teh, and Jim Weatherall for the many conversations on symmetry which led to this paper. It also benefitted significantly from feedback at meetings in Leuven and

Hannover in 2017, and in Salzburg in 2019.

References

Allori, V. (2013). Primitive ontology and the structure of fundamental physical theories. In *The Wave Function: Essays on the Metaphysics of Quantum Mechanics*, pp. 58–75. New York: Oxford University Press.

Allori, V., S. Goldstein, R. Tumulka, and N. Zanghí (2008). On the common structure of Bohmian mechanics and the Ghirardi-Rimini-Weber theory. *British Journal for the Philosophy of Science 59*, 353–389.

Araki, H. and M. M. Yanase (1960). Measurement of quantum mechanical operators. *Physical Review 120*, 622–626.

Banks, T., M. Johnson, and A. Shomer (2006). A note on gauge theories coupled to gravity. *Journal of High Energy Physics 09*, 049.

Barbour, J. and B. Bertotti (1982). Mach's principle and the structure of dynamical theories. *Proceedings of the Royal Society of London 382*, 295.

Bartlett, S. D., T. Rudolph, and R. W. Spekkens (2007). Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics 79*, 555–609.

Belot, G. (2013). Symmetry and equivalence. In R. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics*, pp. pp. 318–339. Oxford: Oxford University Press.

Belot, G. (2018). Fifty million Elvis fans can't be wrong. Nous, forthcoming; available at https://onlinelibrary.wiley.com/doi/abs/10.1111/nous.12200.

Busch, P. and L. D. Loveridge (2013). Quantum measurements constrained by symmetries. In C. Bai, J.-P. Gazeau, and M.-L. Ge (Eds.), *Symmetries and Groups in Contemporary Physics: Proceedings of the XXIX International Colloquium on Group-Theoretical Methods in Physics, Tianjin, China, 20-26 August 2012*, pp. 587–592. Singapore: World Scientific.

Dasgupta, S. (2016). Symmetry as an epistemic notion (twice over). *British Journal for the Philosophy of Science 67*, 837–878.

Dewar, N. (2019). Sophistication about symmetries. *British Journal for the Philosophy of Science 70*, 485–521.

Duncan, T. (2012). *The Conceptual Framework of Quantum Field Theory*. Oxford: Oxford University Press.

Earman, J. (1986). *A Primer on Determinism*. Dordrecht: Reidel.

Earman, J. (1989). *World Enough and Space-Time*. Cambridge, Massachusetts: MIT Press.

Earman, J. (2003). Tracking down gauge: an ode to the constrained Hamiltonian formalism. In K. Brading and E. Castellani (Eds.), *Symmetry in*

Physics: Philosophical Reflections, pp. 140–162. Cambridge: Cambridge University Press.

Fort, H. and R. Gambini (1991). Lattice QED with light fermions in the P representation. *Physical Review D 44*, 1257–1262.

French, S. (2015). Identity and individuality in quantum theory. In E. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition). Available at https://plato.stanford.edu/archives/fall2015/entries/qt-idind/.

Greenberg, O. and A. K. Mishra (2004). Path integrals for parastatistics. *Physical Review D 70*, 125013.

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

Healey, R. (2010). Gauge symmetry and the theta-vacuum. In M. Suaurez, M. Dorato, and M. Redei (Eds.), *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association*, Volume 2, pp. 105–116. Springer.

Huggett, N. (2000). Reflections on parity nonconservation. *Philosophy of Science 67*, 219–241.

Ismael, J. and B. van Fraassen (2003). Symmetry as a guide to superfluous theoretical structure. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, pp. 371–392. Cambridge University Press.

Laidlaw, M. G. G. and C. Morette DeWitt (1971). Feynman functional integrals for systems of indistinguishable particles. *Physical Review D 3*, 1375–1378.

Marsden, J. E. and T. S. Ratiu (1999). *Introduction to Mechanics and Symmetry*. Springer-Verlag.

Menon, T. (2018). Taking up superspace: the spacetime structure of supersymmetric field theory. Forthcomnig in N. Huggett, B. Le Bihan, and C. Wuthrich (eds.), *Philosophy beyond Spacetime* (OUP, forthcoming). Available online at http://philsci-archive.pitt.edu/14682/ .

Møller-Nielsen, T. (2017). Invariance, interpretation, and motivation. *Philosophy of Science 84*, 1253–1264.

Page, D. N. and W. K. Wootters (1983). Evolution without evolution: Dynamics described by stationary observables. *Physical Review D 27*, 2885–2892.

Pooley, O. (2003). Handedness, parity violation, and the reality of space. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, pp. 250–281. Cambridge: Cambridge University Press.

Roberts, J. T. (2008). A puzzle about laws, symmetries and measurability. *British Journal for the Philosophy of Science 59*, 143–168.

Saunders, S. (2007). Mirroring as an a priori symmetry. *Philosophy of Science 74*, 452–480.

Saunders, S. (2013). Rethinking Newton's *Principia*. *Philosophy of Science 80*, 22–48.

Stern, A. (2008). Anyons and the quantum Hall effect - a pedagogical review. *Annals of Physics 323*, 204–249.

Teh, N. J. (2016). Galileo's gauge: Understanding the empirical significance of gauge theory. *Philosophy of Science 83*, 93–118.

Wallace, D. (2003). Time-dependent symmetries: the link between gauge symmetries and indeterminism. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, pp. 163–173. Cambridge: Cambridge University Press.

Wallace, D. (2017). The case for black hole thermodynamics, part II: statistical mechanics. https://arxiv.org/abs/1710.02725.

Wallace, D. (2018). On the plurality of quantum theories: Quantum theory as a framework, and its implications for the quantum measurement problem. Forthcoming in S. French and J. Saatsi (eds.), *Scientific Metaphysics and the Quantum* (Oxford: OUP, forthcoming).

Wallace, D. (2019a). Isolated systems and their symmetries, part I: General framework and particle-mechanics examples. Forthcoming.

Wallace, D. (2019b). Isolated systems and their symmetries, part II: Field theory. Forthcoming.

Wallace, D. (2019c). Naturalness and emergence. Forthcoming in *Monist*; available at http://philsci-archive.pitt.edu/15757/.

Wallace, D. (2019d). Who's afraid of coordinate systems? An essay on representation of spacetime structure. *Studies in the History and Philosophy of modern Physics 67*, 125–136.

Weatherall, J. (2016). Understanding gauge. *Philosophy of Science 83*, 1039–1049.

Weatherall, J. (2017). Classical spacetime structure. Forthcoming in E. Knox and A. Wilson (ed.), *The Routledge Companion to Philosophy of Physics* (Routledge, forthcoming). Available online at https://arxiv.org/abs/1707.05887 .

Wess, J. and J. Bagger (1983). *Supersymmetry and Supergravity.* Princeton: Princeton University Press.

Wigner, E. (1952). Die messung quantenmechanischer operatoren. *Zeitschrift fur Physik 133*, 101. An English translation by Paul Busch is available at https://arxiv.org/abs/1012.4372 .