

Isolated systems and their symmetries, part II: local and global symmetries of field theories

David Wallace

November 7, 2019

Abstract

Physical theories, for the most part, should be understood as modelling isolated subsystems of a larger Universe; doing so, among other benefits, greatly clarifies the interpretation of the dynamical symmetries of those theories. Building on a general framework for the symmetries of isolated systems developed in the prequel to this paper, I apply that framework to field theory. The resultant analysis provides a general basis for interpreting the physical significance of symmetries according to their topological and asymptotic features: global symmetries in general, and local symmetries insofar as they preserve boundary conditions and are asymptotically nonvanishing and/or topologically nontrivial, can be understood as physical transformations of an isolated system against the assumed background of other systems. Other symmetries in general must be understood as mere redescription, though in certain circumstances even non-boundary-preserving local symmetries can be afforded physical significance. The analysis largely reproduces — and so can be seen as a theoretical justification for — general practice in contemporary physics.

1 Introduction and summary of part I

When should a dynamical symmetry transformation be thought of as bringing about a genuinely new possibility, and when should it be seen as a mere redescription? The question dates back at least to the Leibniz-Clarke correspondence (Alexander 1956) but takes a new form in modern field theory due to the introduction of internal (non-spatiotemporal) symmetries and, more significantly, because of local symmetries — the gauge symmetries of modern particle physics and the diffeomorphism symmetries of general relativity. Philosophers by and large agree that local symmetries are mere redescription, but remain divided on global symmetries; physicists by and large regard both global symmetries and those local symmetries with certain asymptotic and/or topological features as genuine physical transformations, and only the remaining local symmetries as mere redescription. (For evidence of these claims see Belot (2018) and references therein.) A closely related controversy concerns when a symmetry transformation applied to a *subsystem* of a system gives rise to an empiri-

cally distinguishable state of affairs, the so-called ‘direct-empirical-significance’ question). The topic has been a locus of considerable philosophical debate in recent years (a partial list: Kosso (2000), Brading and Brown (2004), Greaves and Wallace (2014), Healey (2009), Belot (2018), Teh (2016), Friederich (2015), Gomes (2019)).

Most of this literature makes what in the prequel to this paper (Wallace 2019a) (henceforth Part I) I called the *Cosmological Assumption*: that theories should in the first instance be interpreted as descriptions of the whole Universe. I argued that this assumption, along with the related *Fundamentality Assumption* (that theories should in the first instance be interpreted as descriptions of the whole Universe), should be dropped in most interpretative contexts, and that doing so greatly clarifies our understanding of the interpretation of symmetries.

Also in Part I, I introduced the notion of a *subsystem-recursive* theory as an attractive model for how physical theories can be understood in the absence of the Cosmological Assumption. In such a theory, isolated subsystems of a model of the theory can be understood as modelled (in idealisation) by other sectors of the same theory, and conversely, any sector of the theory can be interpreted as modelling isolated subsystems of some larger model. I constructed a systematic framework based on subsystem-recursive theory to answer questions about the interpretation of symmetry, by (i) understanding the physical significance of a symmetry transformation applied to an in-practice dynamically isolated subsystem represented as part of a model of a larger system, and (ii) transferring the results to the interpretations of models of entire systems via subsystem-recursive theory.

Part I was concerned with the general presentation of these ideas and with their application to a variety of case studies in nonrelativistic particle mechanics in general, and in various formulations of Newtonian gravity in particular. In the present paper, my main concern is the interpretation of symmetries in field theories, partly as a further explication of the ideas of Part I but mostly because the various interpretative questions concerning field-theoretic symmetries are of considerable interest in their own right — especially in the context of the local symmetries of gauge theories, which have been much discussed in the recent literature. I repeat enough of the framework of Part I to make this paper formally self-contained, but I refer the reader back to Part I (and to Wallace 2019b) for detailed defenses of conceptual claims that are only briefly defended here.

The structure of the paper is as follows. In section 2 I state my abstract framework for discussing subsystems of a (sector of a) physical theory and the symmetries that preserve those subsystems, and explain in general how it applies to field theories. In section 3 I introduce three classical field theories, which serve as a source of examples in the rest of the paper. In section 4 I reprise in the field-theoretical context Part I’s account of how symmetry-related states of a subsystem might be indiscernible or might be extrinsically discernible by reference to some other subsystem of the same system, and how we can then use subsystem-recursive theory to draw conclusions about the interpretation of symmetries applied to the whole system. In section 5 I apply this account to the

internal symmetries of the three example theories discussed in section 3. In section 6 I show how my account generalises from internal to spacetime symmetries (via a reformulation of spacetime field theories originally presented in Wallace (2015)), and in section 7 I consider the complications introduced into my account if we add as a modelling constraint that the interactions between subsystems are (in an appropriate sense) weak. Section 8 is the conclusion.

Notation: Greek indices range over (0,1,2,3) and are raised and lowered with the Minkowski metric (signature $+- - -$) unless context implies otherwise. Arrow superscripts (e.g., \vec{x}) denote spatial vectors; given a point x in Minkowski space, \vec{x} denotes the spatial part of it expressed as a vector from the coordinate origin. In section 6, Roman indices from the early part of the alphabet also range over (0,1,2,3) and are raised and lowered with the Minkowski metric (written η_{ab}), while a general spacetime metric $g_{\mu\nu}$ is used to raise and lower Greek indices. The Einstein summation convention is assumed throughout.

2 Subsystem structure for field theory

Following Part I, I define a *subsystem structure* as a collection \mathcal{X} of objects (the subsystems) equipped with a relation \subset (where $X \subset Y$ should be read as ‘ X is a subsystem of Y ’), obeying these assumptions:

- Partial ordering: \subset is reflexive, transitive, and antisymmetric.
- Minimal element: there is some subsystem (written \emptyset) satisfying $\emptyset \subset X$ for all $X \in \mathcal{X}$.
- Maximal element: there is some subsystem (written \mathcal{X} in a mild abuse of notation) satisfying $X \subset \mathcal{X}$ for all $X \in \mathcal{X}$.

Two subsystems X, Y are *disjoint* if $Z \subset X, Z \subset Y$ implies $Z = \emptyset$.

In a field theory, the subsystems are identified with open subsets of whatever spacetime the theory is defined on (mostly Minkowski spacetime \mathcal{M} , in my examples) and \subset and \emptyset have their ordinary set-theoretic interpretations. A particularly significant class of subsystems are the *cylinders*: a cylinder is a smooth map of the cylinder $(a, b) \times B^3$ (B^3 is the unit open ball in \mathbb{R}^3 , and a, b are real numbers with $a < b$) into spacetime, with the sides of the cylinder mapped to a spacelike surface and the ends mapped to timelike surfaces.

Again following Part I, a *state structure* for a subsystem structure \mathcal{X} is then given by:

1. For each $X \in \mathcal{X}$, a set $S(X)$, the *state space* of X , satisfying $S(\emptyset) = \emptyset$.
2. For each $X, Y \in \mathcal{X}$ with $X \subset Y$, a map $r_{XY} : S(Y) \rightarrow S(X)$, the *restriction* of $S(Y)$ to $S(X)$, satisfying:
 - $r_{XX} = \text{id}_X$, where id_X is just the identity on $S(X)$;
 - $r_{XY} \cdot r_{YZ} = r_{XZ}$.

It is required to satisfy the additional condition of

Upwards consistency: Given $X \subset Y$ and $x \in \mathcal{S}(X)$, there is some $y \in \mathcal{S}(Y)$ with $y|_X = x$.

Given these conditions, the restriction from Y to X can be consistently written as $x \rightarrow x|_X$, since $(x|_Y)|_X = x|_X$.

In field theory, states of the whole system are maps from spacetime to some value space \mathcal{V} satisfying some (theory-dependent) continuity and differentiability conditions; the states of a subsystem X are just the restriction of those maps to X , and the restriction map is just ordinary functional restriction. As in the particle case, subsystems in this context are *four-dimensional* entities, so that a ‘state’ of such a subsystem is a specification of its physical properties over an extended period of time. This notion of ‘state’ should be contrasted with the dynamical notion that arises when we talk about ‘state space’ and the like, and in which the dynamics of a system evolves one state into another. I will distinguish between ‘4D state’ and ‘dynamical state’ where the context does not make it clear.

Given subsystems X, Y , and states $x \in \mathcal{S}(X), y \in \mathcal{S}(Y)$, there may or not be a state $z \in \mathcal{S}(X)$ with $Z|_X = x, Z|_Y = y$. If there is, I say that x and y are *compatible*. X and Y are *disconnected* if any $x \in \mathcal{S}(X)$ and $y \in \mathcal{S}(Y)$ are compatible. When systems are *contiguous* — that is, disjoint, but not disconnected — there are compatibility constraints on states in those systems. (Except in pathological cases, disconnectedness clearly implies disjointness, since x, y will not be compatible if their restrictions to a common subset of X and Y differ.)

In field theory, a necessary condition for two systems to be disconnected is normally that their topological closures are disjoint; if this is not the case, there will normally be continuity constraints on which states are compatible. Provided that the value space \mathcal{V} is simply connected (which is the case in all the examples I consider) then this is also a sufficient condition: if X, Y have disjoint closure, we can always find appropriately-continuous functions on the space that restrict to any given functions on X and Y separately.

Also in field theory, it is convenient to define a system $Y \subset X$ as *interior* to X if it is bounded (i. e., does not extend to spatial or temporal infinity) and its closure is contained in X . Provided that the value space \mathcal{V} is topologically contractible (as, again, is the case for all the examples I consider), any state of a subsystem interior to X is compatible with any subsystem disjoint from X ; hence, subsystems interior to X are disconnected from any subsystem disjoint from X .

A physical quantity for X is by definition just a map from $\mathcal{S}(X)$ to some other space T . For instance, a *property* is defined by $T = \{0, 1\}$, or $T = \{\text{True}, \text{False}\}$ if you prefer. A real-valued quantity requires $T = \mathbb{R}$, and so on. Physical quantities for X do *not* restrict to subsystems of X , because the value of the quantity might depend on details of the state that are erased by restriction to X . (They do have natural extensions from X to any Y with $X \subset Y$: the extension \tilde{f} of f from X to Y is defined by $\tilde{f}(y) = f(y|_X)$.)

In field theory, a quantity f for a system X is *local* if X has some interior subsystem Y such that if $x|_Y = x'|_Y$, $f(x') = f(x)$. Non-local quantities are called *asymptotic*: they depend on some features of the state at the boundary of the subsystem, or else on limiting features of the state at the boundary of the whole system at spatial or temporal infinity.

A *symmetry group* for \mathcal{X} is then an assignment, to each X , of a group $\mathcal{G}(X)$ in the algebraic sense, along with a homomorphism r_{XY} from $\mathcal{G}(Y)$ to $\mathcal{G}(X)$, satisfying the same composition, locality and upwards consistency axioms as for state structures. (Hence, we can consistently write $g|_X$ for the restriction of a group element to X). And an *action* of that symmetry group on a state structure is a group-theoretic action R_X of each $\mathcal{G}(X)$ on $S(X)$ such that

$$R_X(g|_X) = R_X(g)|_X. \quad (1)$$

In keeping with usual physics conventions, I will take the action of the group as tacit unless context demands otherwise, and, given $x \in S(X)$, will write gx for $R_X(g)x$. Given a group element $g \in \mathcal{G}(X)$, for any Y with $X \subset Y$ there will exist some (possibly non-unique) $\tilde{g} \in \mathcal{G}(Y)$ with $\tilde{g}|_X = g$; I call any such \tilde{g} a *lift* of g to Y .

In a field theory, the symmetries have the form of possibly-spacetime-dependent bijections of the value space \mathcal{V} : that is, a symmetry g has form

$$(g\varphi)(x) = g_x \cdot \varphi(x) \quad (2)$$

where $g_x : \mathcal{V} \rightarrow \mathcal{V}$ is a bijection. (This means that symmetries are formally always internal in the field-theoretic sense; it is, however, possible to incorporate spacetime symmetries into this framework with a little care, and I discuss this in section 6.)

For any given $x \in X$, the *isotropy group* of x (sometimes called the *stabilizer group*) is the subset of $\mathcal{G}(X)$ whose elements leave x invariant. Normally speaking, the isotropy group of a state is trivial except for extremely special choices of state: for that reason, I define state as *generic* if it has trivial isotropy group.

In field theory, symmetries of a proper subsystem X can usefully be classified as follows: $g \in \mathcal{S}(x)$ is:

- *interior* if for any Y disjoint from X , g has a lift \tilde{g} to Y with $\tilde{g}|_Y = \text{id}_Y$;
- *local* if for any Y disjoint from X , g has a lift \tilde{g} to Y with $\tilde{g}|_Y = \text{id}_Y$.

The collections $\mathcal{G}_{int}(X)$ of interior symmetries, and $\mathcal{G}_{loc}(X)$ of local symmetries, each form a normal subgroup of $\mathcal{G}(X)$, so that we can in each case construct the quotient group. I call $\mathcal{G}/\mathcal{G}_{int}$ the *physical symmetry group* of X , and $\mathcal{G}/\mathcal{G}_{loc}$ the *global symmetry group* of X . Neither group has a uniquely defined action on states of X , but the physical symmetry group defines an action that is unique up to a choice of interior symmetry, and the global symmetry group defines an action unique up to a choice of local symmetry.

Moving from a specific subsystem to the whole theory: a symmetry $g \in \mathcal{G}$ is *fully interior* (or just 'interior') if its restriction to each nonempty subsystem is

interior; likewise, it is fully local if its restriction to each non-empty subsystem is local. A symmetry group is *local-regular* if every local symmetry of X is the restriction to X of a fully local symmetry.

3 Examples of field theories

To illustrate these notions more concretely, I will consider three field theories. In the first two (T1, T2) the field is a complex-valued function φ . In the third (T3), it is a pair (φ, A_μ) , where φ is again a complex function and A_μ is a one-form field (i. e., a covariant vector field, or, in effect, a function taking values in \mathbb{R}^4 .) The dynamics of the theory are now specified by a Lagrangian in each case:

$$\text{T1: } \mathcal{L}_1 = \frac{1}{2} \partial^\mu \varphi^* \partial_\mu \varphi - V(|\varphi|) \quad (3)$$

$$\text{T2: } \mathcal{L}_2 = \frac{1}{2} \partial^\mu |\varphi| \partial_\mu |\varphi| - V(|\varphi|). \quad (4)$$

$$\text{T3: } \mathcal{L}_3 = \frac{1}{2} \mathcal{D}^\mu \varphi^* \mathcal{D}_\mu \varphi - V(|\varphi|) - \frac{1}{4q^2} F_{\mu\nu} F^{\mu\nu} \quad (5)$$

where

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu; \quad \mathcal{D}_\mu \varphi = \partial_\mu \varphi - i A_\mu \varphi.$$

(Physically speaking, T1 is complex Klein-Gordon theory, used to describe scalar fields like the pion; T3 is an example of $U(1)$ gauge theory, with a scalar matter field; T2 is a toy of no direct physical significance.)

In each case, V is some *fixed* function: in accordance with the discussion of section 2 in Part I, different choices of V determine different theories (as do different values of the parameter q in the case of T3). *Sectors* of the theory are individuated by boundary conditions at infinity, and each boundary condition is required to asymptotically minimise energy density. This amounts to requiring in each case that $|\varphi(x)|$ tends to a constant value that is a minimum of V , and in addition imposes conditions on the derivatives:

$$\text{T1: } \lim_{\vec{x} \rightarrow \infty} \partial_\mu \varphi(x) = 0 \quad (6)$$

$$\text{T2: } \lim_{\vec{x} \rightarrow \infty} \partial_\mu |\varphi(x)| = 0 \quad (7)$$

$$\text{T3: } \lim_{\vec{x} \rightarrow \infty} \mathcal{D}_\mu \varphi(x) = 0; \quad \lim_{\vec{x} \rightarrow \infty} F_{\mu\nu}(x) = 0. \quad (8)$$

$$(9)$$

The form of the Lagrangian also constrains the conditions we impose on continuity and differentiability of φ . In each case, in order to avoid divergent values of the energy density, we require that $|\varphi(x)|$ is a differentiable function of x . In T1, we need the stronger condition that $\varphi(x)$ is itself differentiable. In T3, we also need $A_\mu(x)$ to be differentiable. (Essentially for convenience, I normally require that functions that have to be differentiable are also smooth.)

It is worth dwelling a little on this issue of continuity requirements, partly because it does some work in the sequel and partly because it illustrates some

broader themes. Mathematically rigorous presentations of field theory often separate kinematics and dynamics more sharply than I have done here, imposing continuity conditions of various sorts *a priori* and prior to the dynamics. But this makes sense only if we suppose that we are trying to give an exactly true description of Nature — in which case, presumably, the continuity assumption is one more law of nature. (This would be to make the Fundamentality Assumption discussed in Part I.) But in realistic applications, it would be absurd to suppose that a given field theory is remotely trustworthy in its claims about the structure of the fields on scales of 10^{-100}m , let alone $10^{-10^{100}}\text{m}$. Continuity assumptions, rather, are *modelling assumptions*: where discontinuities imply divergent energy density, they also imply a failure of the assumptions under which the theory is valid, and so we should prohibit them from initial data. (And when continuous initial data evolves discontinuities in finite time, as occurs in (e.g.) caustics and shock-wave phenomena, that in turn indicates a breakdown of the theory on those scales.)

(This approach to continuity conditions is mostly tacit in the physics literature and I am not aware of an explicit discussion; but, as an illustration that can be found in any undergraduate quantum mechanics textbook, consider tunnelling through a square-well potential. The potential function is discontinuous, but we still require that both the wavefunction and its derivative are continuous, because the kinetic energy depends on the second derivative and must remain finite. The discontinuity of the potential function translates only to a discontinuity of the second derivative of the wavefunction. In the more singular case of a delta-function potential, the first derivative is permitted to be discontinuous, because in this case the discontinuity implies a delta-function form for the kinetic energy density, leading to a finite kinetic energy in any finite region. In neither case, of course, is the discontinuous potential supposed to be taken literally: we are dealing, rather, with self-consistency requirements in our approach to modelling.)

Each of these theories is subsystem recursive, in the following sense: consider a subsystem that is a cylinder; and require that in some region around the spatial boundary of the cylinder, the fields satisfy the asymptotic boundary conditions of a sector of the whole theory: that is, the fields that region are constant (or covariant constant, $\mathcal{D}_\mu\varphi(x) = F_{\mu\nu}(x) = 0$, in the case of T3) and minimise the potential energy — or, more realistically, suppose that this condition is *approximately* met. And suppose further that there is no significant deviation from these conditions in the vicinity of the spatial boundary. Then, so long as no disturbance from outside propagates in, and so long as no disturbance from inside crosses the boundary, we can take the boundary to spatial infinity without meaningfully affecting the interior dynamics. In this sense, any such cylinder within a model in a sector can be idealised as another sector (or another interpretation of the same sector).

These boundary conditions in turn individuate the sectors of the theory. For T1, there is either a single sector (if the minimum of V occurs at zero) or a one-parameter family of sectors labelled by phase. (In quantum-mechanical and statistical-mechanical applications, the transition between these two states

is interpreted as a phase transition; see, e.g., Binney *et al* (1992).) In T2, there is only one sector, irrespective of the location of the minimum. In T3, there is a large family of sectors.

In each case, the dynamical symmetries of the theory are (by definition¹) those transformations of the field that leave the Lagrangian invariant. As I noted previously, I confine attention to the *internal* symmetries, which have the form $\varphi(x) \rightarrow F(\varphi(x))$ for some function F . It is easy to see that the internal symmetries of T1 are

$$\varphi(x) \rightarrow e^{i\theta} \varphi(x) \quad (10)$$

for arbitrary $\theta \in [0, 2\pi)$ while the symmetries of T2 are

$$\varphi(x) \rightarrow e^{i\theta(x)} \varphi(x) \quad (11)$$

where now θ is an arbitrary (not necessarily continuous) function taking values in $[0, 2\pi)$. The internal symmetries of T3 are

$$\varphi(x) \rightarrow u(x)\varphi(x); \quad A_\mu(x) \rightarrow A_\mu(x) - u^{-1}(x)\partial_\mu u(x) \quad (12)$$

where u is a smooth (or at least twice differentiable) complex function satisfying $|u(x)| = 1$. (We can express this as

$$\varphi(x) \rightarrow e^{i\theta(x)} \varphi(x); \quad A_\mu(x) \rightarrow A_\mu(x) - \partial_\mu \theta(x) \quad (13)$$

where θ is a function taking values in $[0, 2\pi)$, but the continuity requirements on θ are a little subtle: it is required to be continuous only modulo 2π .) Restricting each of these groups of transformations to the various subsystems defines a symmetry group in the abstract sense of section 2, and we can now apply the taxonomy from that section:

- The symmetry group of T1 is rigid, and its global symmetry group is just the full symmetry group.
- All the symmetries of T2 are fully interior, so that the physical symmetry group is trivial.
- All the symmetries of T3 are fully local, but none are fully interior. So the physical symmetry group is the full symmetry group, but the global symmetry group is trivial. (Note that this group is also local-regular: the X -local symmetries are the restriction to X of the fully local symmetries.)

We can also note that in each of T1–T3, a field φ on X is generic (that is, has trivial isotropy group) iff $\varphi(x) = 0$ for all $x \in X$ (supporting our use of the term ‘generic’). In T1 and T2 the zero field has the entire symmetry group of X as its isotropy group; for T3, for states where φ vanishes, the isotropy group is the subgroup of $\mathcal{G}(X)$ that is constant on connected components of X .

¹There are subtleties as to whether this Lagrangian conception of symmetry captures the full notion of dynamical symmetry; I leave these aside for simplicity, but see Brown and Holland (2004) and references therein for discussion.

4 Symmetry and discernibility for isolated systems

A major theme of Part I was that subsystem recursivity, combined with this framework of subsystems, could be used to interpret the symmetries of isolated systems as a whole; here I adapt that analysis to the field-theoretic case. As in Part I, the basic interpretative assumption I make about symmetries is that when two states of a *subsystem* are related by a symmetry, they are identical in their *intrinsic* properties, so that any change caused by the symmetry at most changes a system's extrinsic properties, i. e. its relation to other systems. (And I assume that our description does not contain redundancy beyond that encoded in symmetries, so that the converse is true: any two states *not* related by a symmetry differ in their intrinsic properties.)

Now suppose X and Y are disjoint subsystems, with states x and y that are compatible: that is, such that some state \tilde{x} of the total system satisfies $\tilde{x}|_X = x$, $\tilde{x}|_Y = y$. And suppose that a symmetry $g \in \mathcal{G}(X)$ takes x to some $x' = gx$ which is also compatible with y , with \tilde{x}' satisfying $\tilde{x}'|_X = x'$, $\tilde{x}'|_Y = y$. Do any such \tilde{x} , \tilde{x}' differ by some intrinsic property? Yes — provided that there is *no* symmetry $\tilde{g} \in \mathcal{G}$ with $\tilde{g}\tilde{x} = \tilde{x}'$.

This leads to the following definitions: for any two $x, x' \in \mathcal{S}(X)$:

- x' is *intrinsically discernible* from x if there is no symmetry $g \in \mathcal{G}(X)$ such that $x' = gx$.
- x' is *extrinsically discernible* from x if there is some Y , and some $y \in \mathcal{S}(Y)$, such that any two \tilde{x}, \tilde{x}' with $\tilde{x}|_X = x, \tilde{x}'|_X = x', \tilde{x}|_Y = \tilde{x}'|_Y = y$ are intrinsically discernible, i. e. are unrelated by any $g \in \mathcal{G}$. In this case, y is a *witness* to their extrinsic discernibility.
- x, x' are *indiscernible* if they are neither intrinsically nor extrinsically discernible

To apply this framework meaningfully, we need to require that the subsystem under consideration is small enough that we can regard the rest of the system as a reasonable proxy for all of its environment. For a field theory, a natural requirement is that the subsystem has compact closure: it does not extend off to spatial or temporal infinity.

We can also work at the level of quantities rather than states: a quantity f is:

- *intrinsic* if whenever $x' = gx$ for $g \in \mathcal{G}(X)$, $f(x') = f(x)$; or, equivalently, if $f \cdot g = f$ for all $g \in \mathcal{G}(X)$.
- *extrinsic* if it is not intrinsic, but whenever x' and x are indiscernible, $f(x') = f(x)$.
- *gauge* if $f(x') \neq f(x)$ for x, x' indiscernible.

Where a theory includes gauge quantities, it is descriptively redundant: some of its states do not differ in any physically-relevant way and yet are mathematically represented as distinct. (Again, it only makes sense to apply this analysis directly for small subsystems.)

Our distinction between local, global and interior symmetries fits naturally into this framework, especially if we assume (as is standard in field theory) that every subsystem contains at least one generic state, variant under all nontrivial symmetries. Suppose first that x, gx are states of X with $g \in \mathcal{G}_{int}(X)$. By definition, for any Y disjoint from X , there is a lift \tilde{g} of g such that $\tilde{g}|_Y = \text{id}_Y$ — and so x, gx are indiscernible. A similar analysis shows us that any quantity variant under an interior symmetry must be gauge.

Using subsystem recursivity, we can pull these conclusions up to isolated systems: an interior symmetry of the whole system — that is, in field theory a symmetry that vanishes at spatial and temporal infinity — maps states to indiscernible states, and properties variant under it are gauge. So symmetries of this kind represent descriptive redundancy.

At the other extreme, suppose that x, gx are states of X and that g is not a fully local symmetry, so that for at least one Y disconnected from X , g has no lift \tilde{g} satisfying $\tilde{g}|_Y = \text{id}_Y$. Then if y is some generic state of Y , it is a witness to the extrinsic discernability of x and gx . Transferring this to symmetries of an isolated system via subsystem recursivity, we conclude that global symmetries represent physical transformations of the system, mapping states to intrinsically-identical but extrinsically-distinct states.

The case of local (but non-interior) symmetries is a bit subtler — but it is relatively straightforward at least for local *quantities* (which, recall, depend only on the field values in an interior subsystem) and as long as the symmetry group is local-regular. Let f be such a quantity, and let it depend only on the state in interior subsystem X_0 of a subsystem X ; and let Y be disjoint from X , and thus disconnected from X_0 . Then if f is variant under any local symmetry g , g has an extension to Y whose restriction to Y is the identity; hence, no element of Y can witness the discernability of any g -related states. And so if there are states x, gx of X_0 such that $f(x) \neq f(gx)$, f is gauge. Combining this with our previous results, we have established that a local quantity is intrinsic if it is invariant under all symmetries, extrinsic if it is invariant under local symmetries but variant under some global symmetry, and gauge if it is variant under some local symmetry.

For the more general case, let's consider some subsystem Y disjoint to, but contiguous with, X . By definition, no state $y \in \mathcal{S}(Y)$ can discern x, gx for any $x \in \mathcal{S}(X)$, because not all such x, gx will be compatible with y . The best we can do is use y to explore the discernability of states in the subset $\mathcal{S}_y(X) \subset \mathcal{S}(X)$ of states of x compatible with y . In the infinite limit, subsets like this will represent sectors of the full theory, so let's call $\mathcal{S}_y(X)$ the *y-sector* of X .

Again by definition, y can discern x, gx iff gx lies in the y -sector. So we confine our attention to the subgroup $\mathcal{G}_y(X) \subset \mathcal{G}(X)$ of y -sector-preserving symmetries: $g \in \mathcal{G}_y(X)$ iff $g \in \mathcal{G}(X)$ and gx is compatible with y whenever x is. A symmetry in this subgroup is *y-interior* iff it has a lift \tilde{g} whose restriction to

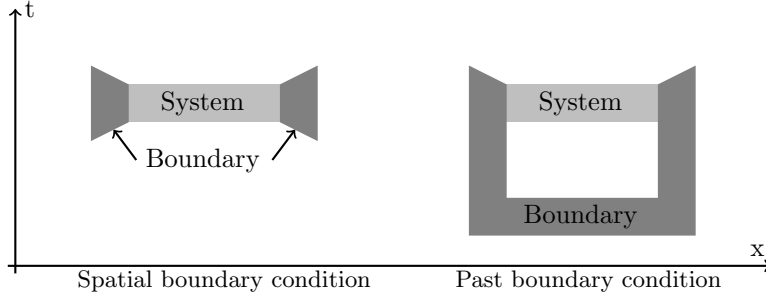


Figure 1: Spatial and past boundary conditions

Y is the identity; the y -global group is the quotient $\mathcal{G}_{y,G}(X) = \mathcal{G}_y(X)/\mathcal{G}_{y,I}(X)$ of the y -sector-preserving symmetries by the group of y -interior symmetries. The action of an element of the y -global group is defined up to an element of the y -interior group.

It is now easy to see that:

- Any two elements of the y -sector related by a y -interior symmetry are not discernable by y .
- Any two elements of the y -sector related by a non-trivial y -global symmetry are extrinsically discernable by y .

The main focus of this paper is interpreting the symmetries of entire sectors of a theory via subsystem recursivity. So the most useful choice of subsystem is a cylinder, and there are two physically useful choices of system contiguous to a cylinder. The first represents a spatial boundary condition for the subsystem: it is any other subsystem whose closure includes the spatial boundary of the cylinder. The second is more complex to state. Given a cylinder $c : (a, b) \times B^3 \rightarrow \mathcal{M}$, a *past extension* of c is another cylinder $\tilde{c} : (a', b) \times B^3 \rightarrow \mathcal{M}$, with $a' < a$, such that \tilde{c} coincides with c on the latter's domain. Our second subsystem type is any subsystem that includes both the spatial boundary of \tilde{c} and its (without loss of generality) its past temporal boundary; any such system can be understood as representing the dynamical state of the system at some time in the past given fixed spatial boundary conditions. Call these subsystems, respectively, *spatial boundaries* and *past boundaries*. (See figure 1 for an illustration, with two spatial dimensions suppressed.)

Assuming that our fields take values in a contractible space (as is the case for all theories I consider in this paper), the y -sector defined by a past boundary state y is no more restricted than the sector defined by the restriction of y to a spatial boundary, but topological reasons sometimes mean that the y -interior symmetry group is smaller than would be the case for a spatial boundary condition alone, as we will shortly see.

The real power of this somewhat abstract framework comes when we combine it with subsystem-recursivity to interpret the symmetries of isolated systems as

a whole by interpreting those systems as idealised descriptions of larger systems. Our previous observations about global and interior symmetries imply the following:

- Two states unrelated by any symmetry are intrinsically discernable.
- Two states related by a nontrivial global symmetry are extrinsically discernable: some distant system, idealised out to infinity, contains states that can witness their discernability.
- Two states related by an interior symmetry are indiscernable.
- A local quantity is intrinsic if it is invariant under all symmetries, extrinsic if it is variant under global symmetries but invariant under local symmetries; gauge if it is variant under local symmetries.

These results do not depend on any restriction to a particular sector, and indeed allow us to deduce something about the relation *between* sectors: if two sectors are related by a global symmetry, they can be interpreted as describing extrinsically different but intrinsically identical systems.

Further results can be obtained if we restrict attention to a given sector. Sectors are individuated by their spatial boundary conditions, which can be seen as an infinite idealisation of the spatial-boundary subsystems described above. For given boundary condition B , we can construct the subgroup \mathcal{G}_B of boundary-preserving symmetries, and within that group, identify the smaller subgroup \mathcal{G}_0 of *asymptotically vanishing* symmetries, which vanish at the boundary; these, respectively, are the idealisations of the y -sector-preserving symmetries and the y -interior symmetries discussed above. Two states differing by an asymptotically vanishing symmetry are not discernable by means of any state on the spatial boundary of the system; two states differing by a boundary-preserving symmetry which is not asymptotically vanishing are extrinsically discernable by a witness state on the spatial boundary. We can define the *asymptotic symmetry group* as the quotient $\mathcal{G}_\infty = \mathcal{G}_B/\mathcal{G}_0$.

It remains possible that two states indiscernable by means of the spatial boundary are nonetheless discernable by means of a past boundary. To explore this, we again use our results about local symmetries of sectors, idealised to the whole system: two states x, gx , where g is an asymptotically vanishing symmetry, will be indiscernable by a past boundary iff g is extendible to a symmetry that equals the identity on the past boundary. Under the idealisations in play, this amounts to the following condition on an asymptotically-vanishing symmetry g : there is a foliation $\mathcal{M} \simeq \mathbb{R} \times \Sigma$ of the spacetime, and a time t_0 (with respect to that foliation), such that for any $t < t_0$, we can find a symmetry \tilde{g} such that:

1. The restriction of \tilde{g} to times $> t$ coincides with g ;
2. For some $t' < t$, the restriction of \tilde{g} to times $< t'$ coincides with the identity.

Call any symmetry g satisfying this condition *past-eliminable*; we can conclude that states related by a past-eliminable symmetry are not discernable by means of any part of their boundary. States related by an asymptotically-vanishing, but non-past-eliminable, symmetry are discernable via their relations with a past state, but not via their relations with the spatial boundary. The past-eliminable symmetries form a subgroup \mathcal{G}_S of the asymptotically vanishing symmetries.

We thus have a chain of subgroups $\mathcal{G}_S \subset \mathcal{G}_0 \subset \mathcal{G}_B$, and can construct various useful quotients from them:

- The *asymptotic symmetry group* is $\mathcal{G}_\infty = \mathcal{G}_B/\mathcal{G}_0$, and its action on a state is defined up to an asymptotically vanishing symmetry. States related by a nontrivial element of the asymptotic symmetry group are extrinsically discernable via the spatial boundary.
- The *topological symmetry group* is $\mathcal{G}_{top} = \mathcal{G}_0/\mathcal{G}_S$, and its action on a state is defined up to a past-eliminable symmetry. States related by a nontrivial element of the topological symmetry group are not discernable via the spatial boundary, but are discernable by reference to an earlier state of the system.
- The *sector-preserving physical symmetry group* is $\mathcal{G}_{SPP} = \mathcal{G}_B/\mathcal{G}_S$, and its action on a state is defined up to a past-eliminable symmetry. States related by a nontrivial element of this group are extrinsically discernable at the spatial boundary and/or by reference to an earlier state.

The local symmetry groups in physics are normally defined by some *local group* G , such that elements of \mathcal{G} can be identified with maps $g : \mathcal{M} \rightarrow G$, with $g\varphi(x)$ depending only on φ and g in a neighborhood of x . (The local group in T3, in particular, has this form.) When this is the case, we can simplify the analysis as follows. Firstly, pick some foliation of the spacetime as above, and write $\bar{\mathcal{G}}$ for the group of functions from an arbitrary (but fixed) hypersurface in that foliation to G ; the symmetries of the theory can then be identified with maps $g : \mathbb{R} \rightarrow \bar{\mathcal{G}}$. A given boundary condition B can normally be translated into a condition on group elements $g(x)$ as x approaches spatial infinity; this restricts $\bar{\mathcal{G}}$ to $\bar{\mathcal{G}}_B$. The elements of $\bar{\mathcal{G}}_B$ which vanish at spatial infinity form a subgroup $\bar{\mathcal{G}}_0 \subset \bar{\mathcal{G}}_B$, and we can identify the boundary-preserving symmetries, asymptotically vanishing symmetries, and elements of the asymptotic symmetry group as, respectively, maps from \mathbb{R} to $\bar{\mathcal{G}}_B$, $\bar{\mathcal{G}}_0$, and $\bar{\mathcal{G}}_B/\bar{\mathcal{G}}_0$.

Now: in accordance with usual usage, two symmetries in $\bar{\mathcal{G}}_0$ are *homotopic* to one another if one can be continuously deformed to the other (without leaving $\bar{\mathcal{G}}_0$, and in particular, a symmetry is *small* if it is homotopic to the identity. Given an asymptotically vanishing symmetry $g : \mathbb{R} \rightarrow \bar{\mathcal{G}}_0$, the smoothness of g means that $g(t)$ and $g(t')$ must be homotopic to one another for any t, t' , so we can meaningfully describe any such symmetry as small or non-small. It is now easy to see that an asymptotically-vanishing symmetry is past-eliminable iff it is small. The question of whether there exist past-ineliminable asymptotically-vanishing symmetries then comes down to the topology of the

space of asymptotically-vanishing maps from spacelike hypersurfaces to the local group \mathcal{G} .

Assuming we are working in Minkowski spacetime, those hypersurfaces have the topology of \mathbb{R}^3 , but the requirement that the symmetry vanishes asymptotically means that we can consistently compactify \mathbb{R}^3 to the three-sphere S^3 and extend each symmetry appropriately. This means that there are past-ineliminable asymptotically-vanishing symmetries iff the space of maps from S^3 to G is topologically non-trivial; indeed, if $H_3(G)$ is the third homotopy group of G , we have

$$\mathcal{G}_{top} \simeq H_3(G). \quad (14)$$

5 Examples of field-theoretic symmetries

Let's apply this rather abstract analysis to the theories T1-T3 discussed in section 3. Beginning with T1: the symmetries here are all global. So two states related by a phase transformation are intrinsically identical, but extrinsically discernable by means of distant systems. Effectively, it makes sense to regard φ and $\varphi \exp(i\theta)$, for nonzero constant θ , as different fields because their phase relations to some (indeed, any) distant system are different. If the potential function V has a nonzero minimum, so that there is a family of sectors labeled by phase angle, we can likewise regard any two *sectors* as extrinsically different (and not just as redescrptions of the same physics) because those sectors, too, are discernable by means of their phase relation to distant systems. At the level of quantities, by writing an arbitrary field as

$$\varphi(x) = \rho(x) \exp(i\alpha(x)) \quad (15)$$

we can usefully define the local quantities $\rho(x)$ and $\alpha(x)$ for each x . In T1, $\rho(x)$ is intrinsic, $\alpha(x)$ is extrinsic. We can define intrinsic functions of the phase by considering derivatives: the local quantity $\partial_\mu \alpha(x)$, for instance, is intrinsic.

Turning to T2, in which all the symmetries are interior: any two fields $\varphi(x)$, $\varphi(x) \exp(i\theta(x))$ are indiscernable. As such, the phase information in T2 is simply redundant: here, we might as well eliminate the symmetry entirely and reformulate the theory as a theory of the real field $\rho(x) = |\varphi(x)|$, with no internal symmetries. (This is actually pretty obvious by inspection, but it is gratifying to find that our results reproduce this intuitive fact about T2). Local quantities $\rho(x)$ and $\alpha(x)$ can be defined, just as in T1; $\rho(x)$ remains intrinsic, but — again, as expected — $\alpha(x)$ is gauge, as are derivatives like $\partial_\mu \alpha(x)$. Similarly, even if the potential has a nonzero minimum so that formally there are many sectors, those sectors are related by interior symmetries and so describe indiscernable states of affairs.

T3 is the most interesting case. There are no global symmetries, and so any two sectors describe the same physical goings-on and cannot in full generality be interpreted as extrinsically different versions of a system. (We will see in section 7 that such an interpretation can be recovered if certain assumptions are made about the background conditions). Even if the potential has a nonzero

minimum, so that there exist different phase-related sectors, those sectors can be extrinsically discerned neither by a distant system (since there are no global symmetries) nor by the spatial boundary (since no such boundary is compatible with different sectors).

Identifying the boundary-preserving symmetries requires us to specify the minima of V . If it has its minimum at zero (the ‘unbroken phase’, in particle-physics parlance), so that the boundary condition is $\lim_{\vec{x} \rightarrow \infty} \varphi(x) = 0$, then the condition for a symmetry $u(x)$ to be boundary-preserving is

$$\lim_{\vec{x} \rightarrow \infty} u(x) = e^{i\Theta} \quad (16)$$

for some fixed Θ . If instead the minimum is at some nonzero value, so that there is a family of minima related by angle (the ‘Higgs phase’), the condition is the more stringent

$$\lim_{\vec{x} \rightarrow \infty} u(x) = 1. \quad (17)$$

This latter condition is also the condition for a symmetry to be asymptotically vanishing. So in the unbroken phase, the asymptotic symmetry group (the quotient of the boundary-preserving symmetries by the asymptotically vanishing ones) is $U(1)$; in the Higgs phase, it is trivial. As a consequence, in the unbroken phase, fields φ and $\varphi \exp(i\theta)$, for nonzero constant θ , represent extrinsically different fields in the same sector; in the Higgs phase, no such interpretation is available, and the fields must be understood just as different ways to represent the same physics.

At the level of quantities, we can again define local quantities $\rho(x)$ and $\alpha(x)$; the former is intrinsic, while the latter is gauge. Derivatives like $\partial_\mu \alpha(x)$ are also gauge quantities, though the covariant derivative $\mathcal{D}_\mu \alpha(x) \equiv \partial_\mu \alpha(x) - A_\mu(x)$ is intrinsic. But we can also define asymptotic quantities like the total phase:

$$A_\Sigma = \int_\Sigma dx \alpha(x) \quad (18)$$

where the integral is over any spacelike hypersurface Σ . This will be infinite in the Higgs phase, but is well-defined in the unbroken phase, and it is an extrinsic quantity: states of different total phase are related by an element of the asymptotic symmetry group.

Whatever the boundary condition, $H_3(U(1))$ is trivial, and hence so is the topological symmetry group. To find an example of a theory with nontrivial topological symmetry group, we can generalise T3 to a non-Abelian gauge theory: the fields φ and A_μ take values, respectively, in \mathbb{C}^2 and in the Lie algebra $\mathfrak{su}(2)$, and the dynamical symmetries are

$$\varphi(x) \rightarrow U(x)\varphi(x); \quad A_\mu(x) \rightarrow U(x)A_\mu(x)U^{-1}(x) - U^{-1}(x)\partial_\mu U(x) \quad (19)$$

for arbitrary smooth functions $U : \mathcal{M} \rightarrow SU(2)$. (The actual dynamical equations will not be needed; see Peskin and Schroeder (1995, pp.486–491), or any

modern quantum field theory textbook.) $SU(2)$ has the same topology as the three-sphere S^3 , and so the topological symmetry group is

$$\mathcal{G}_{top} = H_3(SU(2)) = H_3(S^3) = \mathbb{Z}, \quad (20)$$

the group of integers. For detailed consideration of this example, in the context of the physical interpretation of symmetries, see Teh (2016).

Gratifyingly, these results seem to directly match standard usage in physics, as several examples will show:

1. In the constrained Hamiltonian formulation of gauge theory, two states are regarded as physically equivalent if they are related by a symmetry transformation that (a) vanishes at the spatial boundary, (b) is homotopic to the identity; states related by symmetries which are either topologically nontrivial or asymptotically nonvanishing are treated as physically distinct. See again Teh (2016) for a detailed examination of this case, including a construction (under slightly different names) of exactly the asymptotic, topological, and sector-preserving physical symmetry groups.
2. The interpretation of which states are and are not physically equivalent is not just a matter of physics *terminology*: it is directly demonstrated by statistical mechanics and by the path-integral formulation of quantum field theory, in which sums and integrals count states related by large transformations or by asymptotically non-vanishing, boundary-preserving transformations as distinct, but avoid summing over multiple states related by small asymptotically-vanishing symmetries. (The need to avoid double-counting in this way was a major issue in the quantization of non-Abelian gauge theories — see Peskin and Schroeder (1995, pp.512–515), or any modern QFT textbook, for details.) Similarly, in quantum field theory the physical symmetries are represented by unitary operators, but the small asymptotically-vanishing symmetries have no physical representation.
3. It is standard in field theory — cf Witten (2000, pp.1151,1251) — to regard constant phase rotations of a field as generating physically different states *either* where the phase symmetry is global, *or* in the unbroken phase where the phase symmetry is local but the fields are asymptotically vanishing; by contrast, constant-phase rotations of a field with local phase symmetry in the Higgs phase are not taken to generate physically distinct states. This apparently idiosyncratic behavior finds a natural interpretation here: in the first case, phase-related states can be extrinsically discerned by means of their relation to some distant system; in the second, they can be discerned by means of their relation to the spatial boundary; in the third, neither option is available. (Belot (2018, section 3.4) raises this puzzle²

²To be precise, Belot’s challenge is ‘[to] make sense of the fact that Witten, within a single set of lectures, takes global phase to be unphysical in the vacuum sector of a Yang-Mills-Higgs theory, but physical in the monopole sector of the same theory’; following the references, by

as a challenge for the interpreter of symmetries; this section, and indeed this paper, may be taken as a response to that challenge!)

6 Spacetime symmetries

So far we have been concerned entirely with *internal* symmetries of field theories: spacetime symmetries, such as the translation symmetries of special-relativistic field theories and the diffeomorphism symmetries of general relativity, are not caught by my analysis. The prequel to this paper was concerned with spacetime symmetries, but of particles rather than fields.

It is not entirely trivial to extend the analysis to the spacetime symmetries of fields, and the reason is interesting. Central to my approach has been the existence of a subsystem structure, such that symmetries of systems restrict to symmetries of subsystems. In the particle case, the subsystems are collections of particles, whose spacetime locations can vary from state to state but play no role in individuating the systems themselves; in the field case, they are regions of spacetime, so that spacetime properties actually individuate the regions. This makes it awkward to transfer my analysis to the spacetime symmetries of fields.

There is, however, a trick to get around this: *parametrisation*. I discuss it at some length in (Wallace 2015) (see also references therein for related approaches): here I give only a brief account. The idea is simplest to explain for field theories where the spacetime is a fixed Minkowskian background (as in gauge field theory, for instance). The idea is that instead of defining a field as a map $\varphi : \mathcal{M} \rightarrow V$, we define it as a pair of maps $(X, \varphi) : \mathcal{P} \rightarrow \mathcal{M} \times V$ where \mathcal{P} is a differential manifold diffeomorphic to \mathcal{M} . Points on the manifold can be thought of as parts of the field: the maps X and φ assign to each infinitesimal part of the field, respectively, its spacetime location and its field value. Subsystems are identified with open subsets of \mathcal{P} , and so different fields assign not just different field strengths, but different locations to each subsystem.

Given any dynamics for a field theory on \mathcal{M} , we can construct a dynamics for the parametrised field theory. If we use Roman letters to denote Minkowski coordinates on \mathcal{M} , and Greek letters to denote coordinates on the manifold \mathcal{P} , we can define

$$e_{\mu}^{\alpha} = \partial_{\mu} X^{\alpha}. \quad (21)$$

If the dynamics of the original theory is written in a generally covariant form, and then the metric tensor $g_{\mu\nu}$ is redefined as

$$g_{\mu\nu} = e_{\mu}^a e_{\nu}^b \eta_{ab} \quad (22)$$

we obtain an equivalent dynamics for the parameterised theory equivalent to the original theory. (All this can be done in a coordinate-free fashion if desired,

the 'vacuum sector' Belot means a sector where the internal symmetry is completely broken (so that, on my analysis, there is no boundary-preserving asymptotic or topological symmetry, and so no physical symmetry transformation); by the 'monopole sector' he means a sector where a $SU(2) \times U(1)$ local symmetry is broken down to to $U(1)$ (so that, on my analysis, the asymptotic $U(1)$ symmetry is boundary-preserving and thus represents a physical symmetry.

essentially just by interpreting the indices as abstract rather than coordinate indices.)

The resultant dynamics has two interesting properties. Firstly, it has a diffeomorphism symmetry, which can be interpreted as reflecting the fact that points on the manifold are mere labels and the physical content is given by the coincidences of field value and spacetime location. (Note that this symmetry does not preserve the subsystem structure and so is not amenable to this paper’s analysis: its natural interpretation is as pure descriptive redundancy, without physical content.) Secondly, both the spacetime and internal symmetries of the original theory appear in the parametrised theory as internal symmetries. For T1, for instance, the full group of symmetries can be written as

$$(X^a(x), \varphi(x)) \rightarrow (R_b^a X^b(x) + \Lambda^a, \varphi(x)e^{i\theta}) \quad (23)$$

where R_b^a is a Minkowski-metric-preserving matrix (i. e., element of $SO(3,1)$), Λ^a is a vector, and θ is, as usual, a phase. All of these symmetries are global; our previous analysis returns the result that symmetry-related states are extrinsically discerned by their relation to distant systems — phase relation for the ‘true’ internal symmetries, relations of spatiotemporal distance, relative velocity, and orientation for the spatiotemporal symmetries.

The parametrisation trick can also be carried out for general relativity (GR), though the details are a little complicated and I give them only in outline, referring interested readers to Wallace (2015) for the details. Essentially, we start with a parametrised theory and then promote the global Poincaré symmetry to a local symmetry in just the same way as is done in gauge theory. The resultant theory is characterised by a connection that takes values in the Lie algebra of the Poincaré group, together with a collection of parametrised fields. In the simplest version of this process, we start with the trivial theory in which the field is a ‘location field’, map from \mathcal{P} to \mathcal{M} , with no additional field values: this might be called ‘pure location theory’. Gauging the theory’s global symmetry group gives a theory characterised by a Poincaré connection together with the location field. As demonstrated in Wallace (2015), various choices of gauge fixing can reduce this theory to more familiar forms of general relativity, such as the tetrad formulation with torsion. (As in the case of gauge theory, this localisation process generates only a kinematics: the dynamics needs to be specified manually.)

For our purposes, the important thing is that the resultant theory has as its symmetries (i) the diffeomorphism symmetries of \mathcal{P} ; (ii) local Poincaré symmetries, specified by a smooth assignment to every point in \mathcal{P} of an element of the Poincaré group; (iii) any ‘true’ internal symmetries of the unparametrised theory.

To interpret these symmetries for GR, we need to think about how subsystem recursivity works in that theory. It needs to be conceded right away that there are *some* models of GR for which the Cosmological Assumption is clearly correct, and for which no interpretation as a subsystem of a larger Universe is available. Most notably, spatially-closed models (such as the constant-positive-

curvature versions of Friedman-Robertson-Walker spacetime) have no asymptotic features to interpret as idealised boundaries with a larger system: they are only interpretable as whole Universes. Plausibly, other exotic spacetimes also lack a natural interpretation as subsystems.

But in realistic physical applications, these cases are exceptional. By far the most common applications of GR are to subsystems: pulses of gravitational radiation, coalescing black holes, precision models of the solar system, neutron-star spindowns, and the like. Even in cosmological applications, this is the case: empirical evidence now supports a positive cosmological constant, and so cosmological models asymptotically match anti-de Sitter space, but there seems very little appetite among cosmologists³ to treat these models as genuinely infinite, rather than as sufficient to describe the finitely-large *observable* Universe. (The nearest there is to a consensus about the Universe far beyond our horizon is the ‘eternal-inflation’ model, in which different patches of the Universe have very different values of the cosmological constant and in which quantum fluctuations threaten to invalidate the entire classical-spacetime picture — but even eternal inflation is not *that* close to a consensus, and the default assumption in cosmology is really just that we *don’t know*.)

In any case, for the moment let us set the cosmological constant to zero and restrict attention to the sector of GR with asymptotically flat boundary conditions. Thus restricted, the theory *can* be thought of as subsystem-recursive: subsystems are isolated if their boundaries are approximately Minkowskian, and such subsystems can be modelled, approximately, by models of the theory as a whole, with other matter removed to infinity. (The technical details could be filled in in a number of ways; at this level of analysis they are not material.) The local Poincaré symmetries are local in the sense of section 4, and so symmetry-related states are discernable if at all only by reference to their spatial boundaries; applying the analysis of that section, we deduce that:

1. The boundary-preserving symmetries are those which are asymptotically constant;
2. The asymptotically-vanishing symmetries are (of course) those which are asymptotically trivial;
3. The asymptotic symmetry group is just the (global) Poincaré group;
4. The topological symmetry group is trivial (at least where \mathcal{P} itself is topologically trivial).

So asymptotic symmetries can be interpreted as translations, boosts and rotations of the system relative to reference systems at its spatial boundary, while the asymptotically-vanishing symmetries can be understood as pure gauge transformations, relating different representations of the same physics. If we map these conclusions back to the conventional (i. e., non-parametrised) formulation of

³I concede that this observation is mostly based on conversations and spoken comments.

GR, in which all symmetries are diffeomorphisms, we get the result that diffeomorphisms that either (a) fail to be asymptotically constant in local Minkowski coordinates, or (b) vanish asymptotically, fail to represent physical symmetries, but diffeomorphisms that are asymptotically constant Poincaré transformations represent physical transformations of the system: boosts, rotations, translations.

All of this is again absolutely in keeping with physics orthodoxy: for adducement of evidence, see Belot (2018); for a very clear example in the physics literature, see Guica *et al* (2009, section 4); for further discussion see Wallace (2017b) (the latter paper avoids the complications of the parametrised approach at the cost of considerable technical awkwardness).

Essentially the same analysis could be applied to the asymptotically de Sitter or anti-de Sitter sectors of GR. I am inclined to think it should be somewhat broader still: introducing a sufficiently small (positive or negative) cosmological constant to a system makes no real difference to the physics of local systems and is really just a way of enforcing an asymptotic boundary condition, so in non-cosmological applications of GR we ought to be able to think of sectors with *different* values of the cosmological constant just as different ways of idealising the boundary conditions of systems, irrespective of the actual value of the constant. (This is standard practice in large parts of quantum gravity, in particular in the AdS/CFT correspondence: for discussion, see Wallace (2017a) and references therein.) But further development of these thoughts lies beyond the scope of this paper.

7 The flat-background assumption and approximate global symmetries

Consider a field theory, like T3, whose symmetry group is local. We have seen that if X and Y are two disconnected, spatially distant systems, then symmetry-related states of X cannot be extrinsically discerned using Y as a witness: it will always be possible to extend the symmetry that relates them into a symmetry of the whole system that leaves Y invariant. Only a system contiguous with X can act as a witness in this case. We can understand this in directly physical terms: given a local symmetry then there is no fact of the matter about the relative phase of X and Y , because different ways of filling in the intervening space can alter that relative phase without affecting the states of X or Y .

However, this result relies on the assumption that this ‘intervening space’ might have a quite arbitrary state. In many physically relevant contexts, we are concerned not with this very general context, but with a more restricted context in which we can confine our attention to a subset of models in which substantive assumptions can be made about the background space surrounding the systems of interest. Specifically, given some collection X_1, \dots, X_n of mutually disconnected, simply connected subsystems, define the *flat-background regime* of (a sector of) T3 as the subclass of models in which the curvature $F_{\mu\nu}$ vanishes outside X_1, \dots, X_n . In the flat-background regime, there will be a well-defined

phase angle between the boundaries of systems X_1, X_2 respectively:

$$\Delta u(X_1, X_2) = \exp \left(i \int_{X_1}^{X_2} (\partial_\mu \varphi - A_\mu) dx^\mu \right) \quad (24)$$

where the integral is along any path from a point on the boundary of X_1 to a point on the boundary of X_2 that does not pass through one of the subsystems. The vanishing of the curvature outside the subsystems (and the topological triviality of Minkowski spacetime and the subsystems) means that this integral is independent of the specific choice of path taken. Furthermore, these phase angles satisfy

$$u(X_3, X_2)u(X_1, X_2) = u(X_3, X_1). \quad (25)$$

And in the more physically realistic case where the curvature only approximately vanishes, these results hold to that degree of approximation.

This suggests a qualification to the results of section 4. There, I argued that local symmetries can be understood as physical only when they are boundary-preserving, so that systems related by a boundary-non-preserving symmetry must be regarded as redescrptions of the system, rather than as physically different systems. That result applies if (as I have assumed previously) we are using the theory to model subsystems of a larger system without any assumptions about the rest of the system; if, however, we are using it to model one of the subsystems of a flat-background-regime system, we can after all regard the symmetry as making a *physical* difference, just as in the case of a global symmetry (where we saw that there was no requirement that boundaries were preserved). In effect, assuming a flat-background regime promotes the asymptotic local symmetries of a system to global symmetries, whose effects on one subsystem can be directly witnessed by another.

A variant of this scenario (the ‘spatially-flat-background regime’) occurs if, with respect to a fixed spacetime foliation, the spatial part F_{ij} of the curvature were to vanish. In that circumstance, the *same-time phase angle* between two systems is a well-defined function of time: it is the exponentiated integral of (24) along a path of constant time. In this scenario, asymptotic symmetries are promoted to symmetries that are local with respect to time but global with respect to subsystems.

There are physically significant examples of both scenarios. The first concerns spontaneous symmetry breaking in the Higgs sector of T3. In T1 (that is: in complex scalar field theory with its global $U(1)$ symmetry), in physically relevant contexts symmetry can break in different directions in different regions of space. This can lead to topological defects akin to the domain walls that occur in ferromagnets when symmetry breaks in different directions for different spatial parts of the magnet. It is generally assumed in physics that similar effects can occur in theories like T3 where the broken symmetry is local. Indeed, vortices in superconductors are normally understood in this way, as are cosmic strings in certain cosmological models. But it is not immediately obvious that topological defects make sense for a local symmetry: in the global-symmetry

case we have seen that Higgs-phase systems with differing boundary phases can be understood as physically different when the phase symmetry is global, but not when it is local.

There was a spirited discussion on exactly this point in the physics literature in the 1990s (see Kibble and Vilenkin (1995) and references therein), but the consensus — in accordance with my discussion in this section — is that it does make sense to apply the domain-phenomena results for global symmetry breaking to the local case, precisely because the flat-background-regime assumption holds: curvatures are sufficiently low in the relevant models that it is *de facto* possible to compare the phases of spatially distant subsystems in a meaningful way.

The second example is the ‘Faraday’s Cage’ experiment (in somewhat idealised form). In that experiment, the eponymous cage is charged to high voltage, without any discernable effect on the physical phenomena occurring within the cage. Within Coulombic electrostatics, this can be interpreted straightforwardly as an empirical demonstration of the symmetry $V \rightarrow V + \Delta V$: this symmetry is subsystem-global, and so our previous results (in the particle version of the analysis discussed in this paper’s prequel) establish that models related by a potential difference can be interpreted as describing intrinsically the same system, with the potential difference being an extrinsic difference witnessed by any distant electrostatic system. If we try applying this to T3 (used as a model of a scalar field coupled to electromagnetic fields) then the appropriate symmetry is

$$\begin{aligned}\varphi(x, t) &\rightarrow \varphi(x, t) \int_{t_0}^t d\xi \exp(i\Delta V(\xi)) \\ A_0(t) &\rightarrow A_0(t) + \Delta V(t) \\ A_i(t) &\rightarrow A_i(t).\end{aligned}\tag{26}$$

Healey (2009) has argued that in a theory like this we can no longer interpret the Faraday-cage experiment as empirically demonstrating a symmetry, essentially because on Healey’s analysis local symmetries always count simply as redescrptions of a system. The present framework gives a more nuanced result: the symmetry in question is local but not boundary-preserving, and so cannot *in full generality* be understood as a physical transformation of the system, but it can be so understood if the spatially-flat-background regime applies. Interpreting this physically: if there are strong magnetic fields filling some region that surrounds the cage, then its potential difference compared to a distant system is not well defined and so the charging up of the cage cannot be understood as an intrinsically undetectable change of that potential difference. (Indeed, in this general case it will be difficult at best even to define what we mean by “charging up the cage” in this situation. But if there are no such magnetic fields, or if they are confined to isolated regions, then we can meaningfully define the cage’s potential as an extrinsic property of the cage, and interpret the symmetry transformation (26) as altering that property.

Faraday’s cage offers a nice illustration of the interpretative value of dropping both the Cosmological and the Fundamentalist assumptions. According to

the Cosmological assumption, we should be interpreting an isolated electromagnetic system (at least in the first instance) as alone in the universe, in which case its interpretation should be univocal; but in reality, whether we regard a transformation like (26) as a physical transformation or a mere redescription depends on whether we take the system to be modelling a subsystem of a fully general electromagnetic system, or a subsystem of an electromagnetic system in which the flat-background assumption holds. According to the Fundamental assumption, we should be interpreting both Coulombic electrostatics and gauge-theoretic electromagnetism as if they were exactly true theories, so that the replacement of the former by the latter can cause us to 'discover' that something which appeared to be an empirical symmetry was after all no such thing (this is Healey's view). But if we regard both theories as valid descriptions of the system in different circumstances (with the Coulombic description applicable to a proper subset of situations to which the gauge-theoretic description is applicable) then we ought to be able to hold on to the interpretation of Faraday's cage as a demonstration of symmetry when we go from the one theory to the other, at least in the situations where both are applicable — and our analysis obtains exactly that result.

The flat-background assumption, and the physical significance of a system's boundary-non-preserving symmetries when that assumption holds, generalise beyond T3: they apply, *mutatis mutandis*, to non-Abelian gauge field theory, and indeed to the parameterised form of general relativity I discussed in section 6. In the latter context, we can interpret the non-boundary-preserving symmetries as representing time-dependent accelerations of the system relative to other, distant systems (something that only makes sense when curvature is sufficiently low that the acceleration of a system is well-defined). In this regime, the equivalence principle can be understood as a physical symmetry principle, just as in non-relativistic gravity (see the discussion in section 6 of Part 1).

I note in passing that this section's analysis completes and corrects Greaves and Wallace (2014). In the latter paper, we argued that symmetries that did not preserve boundary conditions could still have physical significance in some circumstances, and advanced Faraday's Cage and Einstein's Elevator as examples, but we gave no detailed prescription for those circumstances. The discussion above provides such a prescription, and has the corollary that the invariance of physics inside a conductor, and inside a sealed box undergoing free-fall, can be understood as a demonstration of a symmetry principle only in the flat-background regime.⁴

⁴In a late stage of preparation of this paper, I became aware of Murgueitio Ramírez and Teh (2019), who also consider the Einstein's-elevator scenarios in Greaves and Wallace (2014). Their analysis takes the concept in a different direction from that which I defend here; I postpone the relation between the two approaches to future work.

8 Conclusion

Field theories in physics, with the exception of a very few genuinely cosmological models, are most naturally interpreted as descriptions of effectively-isolated subsystems of a larger Universe, with the boundary conditions idealized out to spatial and temporal infinity. These field theories also have a rich subsystem structure of their own which permits them to be understood as subsystem-recursive and so allows us to draw conclusions about the interpretation of models of isolated systems from studying the dynamically-isolated subsystems of larger models.

When this approach is applied to the symmetries of field theories, we find that the physical symmetry group of a sector of a field theory is the group of boundary-preserving symmetries, quotiented by the topologically-small boundary-vanishing symmetries. Asymptotically nonvanishing symmetries represent transformations of a system relative to its spatial boundary; asymptotically-vanishing but topologically non-trivial symmetries represent transformations of a system relative to its own previous state; asymptotically-vanishing, topologically-small symmetries are pure gauge and simply represent a mathematical redescription.

The requirement that symmetries in the ‘physical symmetry group’ leave the boundary conditions invariant arises because these symmetries represent transformations which can be performed upon a given system, while leaving its spatial boundary unaffected (at least in idealisation). Transformations which violate the boundary condition cannot be so interpreted, and so cannot be understood as describing a physical change relative to the system’s boundary. If the symmetry is global, they can still be understood as transformations of the system relative to other, distant, systems, and so need not be understood as mere redescription. This strategy is not available in full generality for local symmetries which do not preserve the boundary conditions; however, under the assumption that the system is to be interpreted as a subsystem of a larger system which is for the most part approximately flat in an appropriate sense, we can recover an interpretation of boundary-non-preserving transformations (modulo boundary-vanishing transformations) as describing changes of the system relative to distant systems.

These results all fit closely to contemporary physical practice. They can be seen as providing a theoretical underpinning for that practice; in addition, the closeness of the fit can be seen as further evidence — if evidence is needed — that the natural way to interpret systems in field theory is as subsystems of the Universe and not as fictional universes in their own right.

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

- Alexander, H. (Ed.) (1738/1956). *Th Leibniz-Clarke Correspondence, together with extracts from Newton's Principia and Optics*. Manchester: Manchester 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>.
- Binney, J. J., N. J. Dowrick, A. J. Fisher, and M. E. J. Newman (1992). *The Theory of Critical Phenomena : an introduction to the renormalisation group*. Oxford: Oxford University Press.
- Brading, K. and H. Brown (2004). Are gauge symmetry transformations observable? *British Journal for the Philosophy of Science* 55, 645–665.
- Brown, H. R. and P. Holland (2004). Dynamical versus variational symmetries: Understanding Noether's first theorem. *Molecular Physics* 102, 1133–1139.
- Friederich, S. (2015). Symmetry, empirical equivalence, and identity. *British Journal for the Philosophy of Science* 66, 537–559.
- Gomes, H. (2019). Gauging the boundary in field-space. *Studies in the History and Philosophy of Modern Physics* 67, 89–110.
- Greaves, H. and D. Wallace (2014). Empirical consequences of symmetries: a new framework. *British Journal for the Philosophy of Science* 65, 59–89.
- Guica, M., T. Hartman, W. Song, and A. Strominger (2009). The Kerr/CFT correspondence. *Physical Review D* 80, 124008.
- Healey, R. (2009). Perfect symmetries. *British Journal for the Philosophy of Science* 60, 697–720.
- Kibble, T. and A. Vilenkin (1995). Phase equilibration in bubble collisions. *Physical Review D* 52, 679–688.
- Kosso, P. (2000). The empirical status of symmetries in physics. *British Journal for the Philosophy of Science* 51, 81–98.
- Murgueitio Ramírez, S. and N. J. Teh (2019). Abandoning Galileo's Ship: The quest for non-relational empirical significance. Forthcoming.
- Peskin, M. E. and D. V. Schroeder (1995). *An introduction to Quantum Field Theory*. Reading, Massachusetts: Addison-Wesley.
- Teh, N. J. (2016). Galileo's gauge: Understanding the empirical significance of gauge theory. *Philosophy of Science* 83, 93–118.
- Wallace, D. (2015). Fields as bodies: a unified presentation of spacetime and internal gauge symmetry. <https://arxiv.org/abs/1502.06539>.

- Wallace, D. (2017a). The case for black hole thermodynamics, part II: statistical mechanics. <https://arxiv.org/abs/1710.02725>.
- Wallace, D. (2017b). The relativity and equivalence principles for self-gravitating systems. In D. Lehmkuhl, G. Schliemann, and E. Scholz (Eds.), *Towards a theory of spacetime theories*, pp. 257–266. New York: Birkhäuser.
- Wallace, D. (2019a). Isolated systems and their symmetries, Part I: General framework and particle-mechanics examples. Forthcoming.
- Wallace, D. (2019b). Observability, redundancy and modality for dynamical symmetry transformations. Forthcoming.
- Witten, E. (2000). Dynamics of quantum field theory. In P. Deligne, P. Etingof, D. S. Freed, L. C. Jeffrey, D. Kazhdan, J. W. Morgan, D. R. Morrison, and E. Witten (Eds.), *Quantum Fields and Strings: a Course for Mathematicians*, Volume 2, pp. 1119–1424. Providence, RI: American Mathematical Society.