Representational schemes for theories with symmetry

Henrique de A. Gomes

University of Oxford
Oriel College, Oxford, OX2 1TQ, UK

October 15, 2024

Abstract

In the philosophical literature, symmetries of physical theories are most often interpreted according to the general doctrine called 'traditional sophistication'. But even this doctrine leaves two important gaps in our understanding of such theories: (A) it allows the individuation of isomorphism-classes to remain intractable and thus of limited use, which is why practising physicists frequently invoke 'relational, symmetry-invariant observables'; and (B) it leaves us with no formal framework for expressing interesting counterfactual statements about different physical possibilities. I will call these *Limitations* of TS. Here I will show that a new Desideratum to be satisfied by theories with symmetries allows us to overcome these Limitations. The new Desideratum is that the theory admits what I will call *representational schemes* for its isomorphism-classes. Each such scheme gives an equally valid reduced formalism for a theory.

Contents

1 Introduction

Traditional Sophistication (henceforth TS) is a doctrine about symmetry-related models of theories of physics (cf. (Martens & Read, 2020, p. 323)). At a very broad level, before we get into the details, the doctrine says that any two models that are related by a dynamical symmetry can be taken to represent the same physical possibility.[1]

In the context of the relationist-substantivalist debate, and in particular in the narrower context of the hole argument, TS cannot be neatly located within either camp. It sides with the relationist in some issues and with the substantivalist in others. In the camp of the relationist, TS says that spacetimes related by a symmetry should represent at most one physical possibility; and, still against the traditional substantivalist, it *denies* that spacetime points enjoy a primitive identity, or *hacceity*. But, on the camp of the substantivalist, TS imposes no requirement to change or replace models of a theory so that each possibility is represented by only one model.[2]

Prima facie, there seems to be little reason to endorse this middle ground. Indeed, we can certainly conceive of dynamical symmetries as relating models that we would wish to interpret as representing very distinct physical possibilities (Belot (2013) gives many examples of this).

However, as widely recognised (cf. Jacobs (2022); Martens & Read (2020) for recent appraisals), TS *is* compelling for—and usually understood to apply only to—theories that exhibit a particular mathematical feature. To exhibit this feature (called Criterion (ii) in Gomes (2021c)),[3] a theory must first of all have its models be expressed by values of dynamical variables that are functions of a set of independent variables. This set constitutes a fixed, 'absolute', or non-dynamical background and this background is endowed with some structure of its own. For spacetime theories such as Newtonian mechanics, special or general relativity, it is a spacetime manifold endowed with a geometric structure, such as: Newtonian absolute space and time, a Minkowski metric, or just a smooth structure, respectively in the three cases (see (Earman, 1989, Ch. 2) for a longer and more detailed list). Models in Newtonian mechanics would be given by particle trajectories over the corresponding highly structured background; models in general relativity would be given by a metric and other tensors over the smooth manifold. But we could also have models of theories that incorporate other fields, such as those of the Standard Model of particle physics, express their dynamical content on a fixed background

[1]Following (Gomes, 2021d, p. 5), in this paper I will focus on theories whose equations of motion can be defined variationally from a Lagrangian, and for which we can define dynamical symmetries as 'a group of transformations that is independent of the models of the theory, but that has an action on these models—an action which may be model-dependent—which leaves the Lagrangian invariant up to boundary terms'.

[2]In the context of spacetime theories, the label 'sophisticated substantivalist' was used with a negative connotation in (Belot & Earman, 1999) (see also (Belot & Earman, 2001)), to characterize those who sought to simultaneously keep an ontological commitment to spacetime points while rejecting their context-independent, or primitive, identity across possibilities. Later on, the label was stripped of its pejorative connotation and accepted by many philosophers. Dewar (2017) extended the adjective "sophisticated" to other symmetries: where it is understood as allowing commitment to structure without reduction or elimination of the isomorphic copies.

[3]Criterion (i) in Gomes (2021c) requires dynamical symmetries to be in a 1-1 relation with the isomorphisms of some well-defined mathematical structure common to all of the models of the theory. This Criterion is very easily satisfied, due to the vast resources of mathematics and flexibility of what constitutes a well-defined mathematical structure. Criterion (ii) is a very straightforward generalisation of (Earman, 1989, p. 45)'s famous 'SP principles' for spacetime symmetries, as also described in Jacobs (2022); Martens & Read (2020).

structure that contains an 'internal' space over each spacetime point (this is called 'a vector bundle', cf. Gomes (2024)). The set of transformations that leave this background structure invariant are called the background's automorphisms.

Finally, Criterion (ii) then requires, for a given theory, that the dynamical symmetries be uniquely represented by automorphisms of the models' common background structure—e.g. the diffeomorphisms of a smooth manifold, or the fiber-preserving linear automorphisms of a vector bundle. For theories satisfying this Criterion, the difference between two symmetry-related models is invisible to the background structure: it is purely representational, more like a change of coordinates. For such theories, it is natural to interpret symmetry-related models as being physically identical, thus rendering TS compelling.[4]

Here, I am mostly concerned with two types of theory that exhibit the particular mathematical feature described above, and so for which TS is compelling: (a) general relativistic and (b) gauge, i.e. Yang-Mills, theories. In (a) dynamical symmetries of the Einstein-Hilbert Lagrangian are isometries of Lorentzian metrics, which are generated by smooth diffeomorphisms of the underlying smooth manifold where the metric is defined (and so satisfy Criterion (ii)). Similarly, in (b) dynamical symmetries of the Standard Model Lagrangian give a kind of isomorphism of models, called a *gauge transformation*, that can be construed as spacetime-dependent fiber-preserving linear isomorphisms for (internal) vector spaces attached to each spacetime point (and so similarly satisfy (ii)).

I take TS to significantly illuminate our understanding of such theories. But there are certain important questions about symmetry and physical equivalence that lie beyond even its scope. I call these questions 'limitations', and here I will argue that they are overcome by ideas that are broadly 'relationist'.

There are three main such Limitations, of which I here aim to overcome only two (which are aspects of what Gomes (2021c) called Worry (2)), which I will label Limitations (A) and (B):[5]

(A) TS gives us a sufficient condition of identification of what two models represent about the world (or contrapositively: a necessary condition of non-identification). That condition

[4]Suppose isomorphisms of the models were induced by a bijection of the background structure, but assume one such bijection is *not* an automorphism of the structure. Then it would still be true that two thus symmetry-related models would be distinct only insofar as they distributed properties differently over that background. But in this case, the background structure is *sensitive*, not indifferent, to the differences between the transformed models. This is what occurs, for example, for boosts in a system of Newtonian particles: boosts yield dynamical symmetries, but two boost-related models have different states of motion with respect to the background absolute frame.

[5]I will set aside what Gomes (2021c) called Worry (3): as argued persuasively by Belot (2018), there is a consensus among physicists that, for certain sectors of general relativity, TS is wrong, because some isomorphisms of models relate *different* physical possibilities. Namely, in the context of spatially asymptotically flat spacetimes, some diffeomorphisms—those that preserve the asymptotic conditions but don't asymptote to the identity map—are interpreted as relating different physical possibilities. The same occurs in Yang-Mills theory (Giulini, 1995). But the reasons given for TS (see Gomes (2021c)) were general; they applied at the level of the whole theory. So TS, as a doctrine about the whole theory, i.e. all its sectors, seems refuted.

is the existence of an isomorphism between the models. TS commits the theory to a symmetry-invariant ontology, but it does not require the theory to have explicit, symmetry-invariant descriptions of that ontology. Without such descriptions, at a practical level, TS's sufficient condition of identification falls short. For given two models, the sufficient condition is not, *prima facie*, tractable: one cannot try out every diffeomorphism—or, more generally, automorphism of the models' common background structure—to see if one will bring the two models into coincidence. This Limitation motivates many theoretical physicists to seek relational characterizations of the symmetry-invariant ontology; even while generally endorsing TS (as described in Belot & Earman (1999, 2001)).

(B) By denying that spacetime points have primitive identity across physical possibilities, TS relinquishes any notion of correspondence between spacetime points that belong to non-isomorphic spacetimes. Similarly, for gauge theory, it relinquishes any notion of correspondence between the values of gauge fields belonging to non-isomorphic models. But such correspondences are crucial for counterfactuals to be expressed and interpreted. And beyond counterfactuals, these correspondences are crucial for making rigorous sense of the idea of superpositions of classical spacetimes, a topic that is central to various efforts to combine quantum mechanics with gravity (for recent discussion of why such correspondences are needed in order to thus combine the fileds, cf. (Kabel et al., 2024) and references therein).

In this paper it will turn out that answering Limitation (A) in a certain way will also answer Limitation (B). So I begin in Section 2 by elaborating (A). I will there explain why these Limitations are not merely metaphysical: they are very relevant for current issues in theoretical physics. Then, in Section 3, I introduce the idea of *representational schemes*.[6] Conceptually, such schemes are choices of (idealised) physical relations with which we can completely describe each physical possibility without redundancy. So each scheme gives an equally valid reduced formalism for a given theory, and (a sub-class of) such schemes will suffice to answer both (A) and (B).

In Section 4 I provide several examples of representational schemes. Focusing on the case where these schemes are choices of gauge-fixing, I will illustrate some of their important properties, such as non-locality. In Section 5 I discuss how choices of representational schemes are possible in the absence of an empirical breaking of the symmetry of the theory. In Section 6 I will develop the theory of counterparthood based on representational schemes and show that overcoming Limitation (A) via (a sub-class of) representational schemes also overcomes Limitation (B). That is, the resolution of Limitation (A) via certain types of representational schemes offers not only tractable conditions of identity for the entire universe, they also provide local correspondence relations between spacetime points (and, respectively, for internal vector

[6]These were previously called 'representational conventions' (cf. (Gomes, 2022)). But a 'convention' over-emphasises arbitrariness, which is a misleading connotation; whereas 'schemes', while still signalling that a choice has to be made, connote that this choice must satisfy certain constraints.

spaces, in the case of gauge theory) in non-isomorphic models. And in Section 7 I conclude by describing a limitation of representational schemes and a possible route to overcome it.

2 The challenge of individuating structural content

Here is how I will organise this Section. In Section 2.1 I will expound Limitations (A) and (B). In Section 2.2 I will relate Limitation (A) to issues in the metaphysics of spacetime, but argue that the Limitation is not solely metaphysical: it has echoes on current theoretical physics. In Section 2.3 I will try to narrow down what could count as a resolution of Limitation (A); this is where I first introduce *representational schemes*, which will be developed in the rest of the paper.

2.1 Limitations A and B in more detail

2.1.a Limitation (A): the individuation of isomorphism-classes

TS provides a sufficient condition for two models to represent the same physical possibility. To apply this condition in any specific case, one needs to fix a context in which the two models are being applied; as argued in recent discussions (cf. (Fletcher, 2020; Pooley & Read, 2022)). But relative to such a fixed context of application, the sufficient condition is very simple: namely, that there is an isomorphism between the representing models.

But here is Limitation A: even once we have fixed a notion of isomorphism and a context of application for the theory, we may not have a tractable way to individuate the *isomorphism-classes*. That is, because we lack an explicit isomorphism-invariant description of the models, we may not have tractable conditions for assessing when two given models are isomorphic. For sometimes it is exceedingly hard, given any two models, to verify whether they are isomorphic or not, via direct inspection of all candidate isomorphisms. That is why mathematicians often seek a condition equivalent to the existence of isomorphisms that is easier to verify.

Indeed, the classification problem for a type of mathematical object is a very familiar topic within mathematics. It is sometimes trivial—e.g. finite-dimensional vector spaces are isomorphic iff they have the same dimensions—and it is sometimes hard, e.g. seeking all the topological invariants of topological spaces. (Note here that 'all' means that two spaces having the same set of invariants implies that the two spaces are homeomorphic.)[7]

An analogy to purely philosophical discussions may be helpful here. Thus we can take Leibniz's venerable Principle of the Identity of Indiscernibles (PII) to provide a sufficient condition of identity of objects: viz. if objects x and y have exactly the same qualitative properties, then $x = y$.[8] But this sufficient condition is almost always intractable, since we cannot check

[7]For instance, mathematicians place great importance in theorem's like Perelman's, which shows that all three-dimensional compact manifolds with a certain homotopy group are homeomorphic to the three-dimensional sphere.

[8]There are many ways to interpret every term in this abbreviated description: the scope of 'objects', the meaning of 'qualitative', etc. In my usage so far, qualitative properties correspond to those represented by

it for *all* the properties of an object (even just the qualitative ones). Hence the philosophical endeavour to formulate identity conditions for a specific sort of object, i.e. tractable sufficient conditions for identity (e.g. the claim that, as we have learned from crime dramas, for humans, the identity of fingerprints, or dental records, is sufficient for personal identity). So the analogy is that: (i) Leibniz's PII is like the way that the definition of a mathematical structure fixes, *ipso facto*, a sense of isomorphism, but gives no guidance about how to ascertain whether two structures are isomorphic; so that (ii) the search for tractable sufficient conditions for identity is like the mathematician's seach for a solution to the a classification problem.

As an example, in general relativity, we may postulate that models are Lorentzian manifolds whose geometric structure is invariant under isometry, and that a sufficient condition for two models representing the same physical possibility is 'being isometric'. But given two Lorentzian metrics, how are we to ascertain *when* they are isometric?

To show that two given models are *not* isomorphic, all that TS implies is that if we scan the entire infinite-dimensional space of gauge transformations or diffeomorphisms, and do not find a gauge transformation (respectively, diffeomorphism) between the models, then they are not isomorphic. To be more explicit about this problem: given two coordinate-based descriptions of metrics $g^1_{\mu\nu}, g^2_{\mu\nu}$, we would either have to stumble upon a coordinate change that relates them (and so infer that they represent the same physical possibility), or else stumble upon some coordinate-invariant function that can be defined on all spacetimes and that takes different values on the two metrics (and so infer that they represent different physical possibilities). And thus if we fail on both counts, we could not decisively conclude that the two metrics correspond to same geometry, nor that they correspond to different geometries.

In the following, I will focus on giving tractable and sufficient conditions of identity, or a tractable notion of individuation for isomorphism-classes, by providing invariant descriptions of the isomorphism-classes. But I do not aim to give unique such descriptions. Quite the opposite: I claim there are many equally valid choices, differing perhaps with respect to theoretical virtues and depending on the context of application. Each such choice provides an explicit, reduced formalism for the theory, that can be broadly classified as 'relational' in spirit.

In our search for such invariant descriptions, we face two main obstacles. The first is the obstacle of *completeness*: the description may only be partial; the conditions that it provides are insufficient to ascertain whether any two isomorphism-classes are identical. One could try to overcome this obstacle by taking the union of many such partial descriptions, but then one faces the second obstacle, of *consistency*. Partial invariant descriptions may be inconsistent, and assessing consistency may again be an intractable endeavour. I will now illustrate these obstacles in the case of Riemannian geometries.

As our first attempt to find tractable and sufficient conditions of individuation, take a familiar list of isomorphism-invariant (in this case, isometry-invariant) properties. For instance: (1) "the manifold contains a two-dimensional surface, of area X, bounded only by two geodesics,

symmetry-invariant properties of the models. Thus TS commits us to a qualitative ontology, but it gives us no explicit, qualitative description of that ontology. Having said that, for this short example, I need no such details.

of length, L"; or (2) "the manifold is geodesically convex"; or (3) "for all points x, there exists a unique point y whose geodesic distance πR from x is greater than all other points", etc. All of these properties are invariant under isometries. Property (3) is instantiated say, in the two-sphere, where it describes anti-podal points, and it is not instantiated on a plane. Indeed, it is only property (3) that comes close to fully characterising the geometry: it would characterise a metric sphere of some dimension. And that is only because the metric sphere is a highly unusual, highly symmetric model. In contrast, (1) and (2) are far from uniquely characterising a physical possibility, or even points or regions of the world in which these properties hold. To do so in general we would need an infinite list of properties like (1) and (2).

These examples illustrate the first obstacle that gets in the way of a description via a denumerable list of geometric invariants: *completeness*. A *complete* set of invariants (also called in the literature on gauge theory a complete set of 'observables', cf. (Henneaux & Teitelboim, 1992, Ch. 1.2), (Bergmann, 1961)) is a set whose elements, taken together, distinguish two isomorphism classes of models; i.e. any two models in two such classes disagree about the value of some or other observable in the set.

But clearly, any *finite*, consistent list of geometric invariants is, generically, *incomplete*. That is because a Riemannian geometry has infinitely many degrees of freedom. Indeed, the space of Lorentz metrics that satisfy the Einstein equations has continuously many degrees of freedom, i.e. two physical degrees of freedom per space point; and the space of Yang-Mills connections also has continuously many degrees of freedom (cf. (Henneaux & Teitelboim, 1992, p. 29)). This means that, generically, no finite or even denumerable list will completely fix a particular isomorphism-class of a field theory.[9]

This leads us to the second obstacle, of consistency: it is not clear from inspection when two items in a list such as the one above are inconsistent. For instance, in the simple case where (3) is part of the list, the L of item (1) would have to be smaller than $4\pi R^2$. We arrive at this explicit condition because in this simple case (3) fixes the geometry to a large degree, and we can then use familiar theorems of spherical geometry. But in general cases the consistency of a long list would pose an intractable problem: I call this the problem of *consistency* of generic lists of symmetry-invariants; or, more generally, the problem of *consistent descriptions*. We will encounter it again in Section 3.4, when we discuss non-locality. Thus, generically, any such description can be subject to the problems of incompleteness and inconsistency.

So much by way of stating Limitation (A). To sum up: though TS says only symmetry-invariant properties of the models can represent physical properties, it neither provides nor 'sees the need' to provide a symmetry-invariant description of those models. Thus, in practice, TS does not help us ascertain when two models of our theory are isomorphic; it gives no guidance about how to find a consistent and complete set of observables, i.e. a set of observables whose compossible values label, one-to-one, the isomorphism classes.

[9] And then there is the related worry, about a non-denumerable list of invariants being vastly overcomplete! In this case, the elements of the list could be invariant but, since vastly overcomplete, still have to satisfy constraints that are in practice intractable. So this problem is one of consistency, as above, rather than one of completeness.

2.1.b Limitation (B): the individuation of objects

I now turn to Limitation (B). Let us start by seeing it as an analogue, for objects or other features, such as regions, *within* a model, of Limitation (A) which applied to *models as a whole.* That is: if we are confronted with two models (whether isomorphic or not), and consider some intra-model object or feature in one model—such as a spacetime point or region in general relativity, or a frame for a vector space in gauge theory—then the general question arises: which (if any) object or feature of the other model corresponds to it?

Agreed: stated so generally, this question is vague: it of course depends on what the scientific or philosophical meaning or role of 'correspond' is meant to be. In narrowing down this meaning, we must be careful not to take the analogy to Limitation (A) too far, or 'correspondence' might be too constraining; it could require too much similarity to be of any use. For instance, in the case of general relativity, suppose we take the notion of 'correspondence' to require the matching of all observables attached to a region. In such a case, two regions would correspond iff all their physical properties matched. It goes without saying that such a notion of correspondence will not be of much use for counterfactuals!

In other words, we don't want a specific spacetime region to bear *all* of its symmetry-invariant (or qualitative) properties 'essentially'.[10] No, we want to uniquely characterise, but not individuate–which would be too strong—regions in each member of a generic family of non-isomorphic models. Again, TS gives no guidance about what properties or relations two spacetime points (or regions) in two models need to have in common in order to 'correspond' in this light-handed, flexible way. In particular, it gives no guidance about for which observables the points (regions) must match in their values, in order that they 'correspond'.

This is Limitation (B), and this paper will describe how we can overcome it. As with overcoming Limitation A, it will become apparent that there are many different, independent sets of relations that can 'uniquely characterise, or specify points as places in a structure' across possibilities; sets which differ only with respect to theoretical virtues specific to the context of application.[11]

2.2 An upshot for theoretical physics

In this Section I will present evidence that the implications of these Limitations are not strictly metaphysical: they are important for current issues in theoretical physics.

[10]In the context of the hole argument, to be briefly described below, in Section 2.2, this kind of rigidity is associated to (Maudlin, 1988, 1990)'s 'metric essentialist' resolution, which indeed has been criticised for being an argument valid only for the 'actual world' (Butterfield, 1989); and in (Gomes & Butterfield, 2023a) precisely for not allowing the expression of counterfactuals.

[11]Incidentally, this plurality of sets of relations that are sufficient for individuation is essentially absent from the allusions to individuation in the literature on symmetry and equivalence. On the contrary, this literature much more often (if not always) considers only the totality of properties and relations in which points stand with respect to each other. But, as in the analogy to Leibniz's PII, such totality gives us no tractable handle on these questions. Here I will propose a more tractable notion of relationism, that further picks out model-dependent subsets of relations.

I will start with Limitation (A), and with a connection to the hole argument, where the physical interpretation of symmetry-related models has high stakes (cf. Gomes & Butterfield (2023b); Pooley (2023) for recent appraisals).[12] The argument articulates the threat of a pernicious form of indeterminism that could arise in general relativity due to the the existence of isomorphic models. By stipulating that isomorphic models represent the same physical situation, TS disarms the threat. Of course, the stipulation is not unwarranted: as I argued in Section 2.1.b, it is compelling in the case of general relativity, and some judge that it is even guaranteed by a proper understanding of the mathematical formalism of Lorentzian manifolds (see e.g. (Fletcher, 2020; Weatherall, 2018)).

I only partly agree with this judgement. And I am not alone: the hole argument has been argued to be consequential for practising physicists, as most extensivelly catalogued by Belot & Earman (1999, 2001). They write:

> Far from dismissing the hole argument as a simple-minded mistake which is irrelevant to understanding general relativity, many physicists see it as providing crucial insight into the physical content of general relativity. (Belot & Earman, 1999, p. 169)

I take the spirit of the hole argument, at bottom, to be a gesture towards Limitation (A), about individuating isomorphism-classes, which is why the argument so often figures in the longstanding 'substantivalist-relational' debate, and why Einstein's 'point-coincidences' are so often cited in its resolution.[13]

Here is how Isham explicitly ties the hole argument to Limitation (A):

> The diffeomorphism group moves points around. Invariance under such an active group of transformations robs the individual points of M of any fundamental ontological significance. [...] This is one aspect of the Einstein 'hole' argument that has featured in several recent expositions. [...] It is closely related to the question of what constitutes an observable in general relativity—a surprisingly contentious issue that has generated much debate over the years. (Isham, 1992, p. 170)

As Isham remarks, it is a matter of fact that physicists also devote significant attention to characterising 'observables' (which, as I described in Section 2.1.a, are quantities that are invariant under the symmetries of the theory, that we could use to describe physical content without redundancy).[14] And this matter of fact is hard to reconcile with the idea that TS leaves

[12] In the case of theories with only time-independent symmetries, there is no threat of indeterminism; see e.g. (Wallace, 2002) for a philosophical appraisal.

[13] The main idea, to be discussed at length throughout the paper, is to specify points by their web of relations to other points: "All our spacetime verifications invariably amount to a determination of spacetime coincidences [...] physical experiences [are] always assessments of point coincidences." (Einstein, 1916). But, as I said, that general argument is vague, and does not help to explicitly and invariantly characterize isomorphism-classes.

[14] See e.g. (Donnelly & Giddings, 2016; Rovelli, 2007; Thiemann, 2003) for general arguments surrounding this difficult issue of 'gravitational observables'. But there are many examples from the string theory literature as well: e.g. see (Harlow & Wu, 2021) for an explicit, and rather complicated basis in the simplified context of a

open no further important questions about symmetry and equivalence once the mathematical notion of isomorphism between models is clarified.

As to Limitation (B), overcoming it may be necessary in order to discuss local spacetime counterfactuals in full generality, as I briefly described in Section 2.1.b and alluded to in Section 2.1.b. Modal intuitions— such a region of spacetime *could have been* more curved—invoke *some* relation between spacetime regions across physical possibilities. Since we took TS to deny that spacetime points have primitive identity across physical possibilities, we need an alternative way to establish *local* inter-world comparisons. This obstacle is widely recognised by physicists; as remarked by (Penrose, 1996, p. 591):

> The basic principles of general relativity—as encompassed in the term 'the principle of general covariance' (and also 'principle of equivalence') —tell us that there is no natural way to identify the points of one space-time with corresponding spacetime points of another.

In the context of spacetime, Limitation (B) has led to some puzzlement about how general relativity meshes with standard conceptions of scientific laws and scientific understanding. For instance, according to (Hall, 2015, p. 270), 'the ability to provide sharp and determinate truth conditions for a wide range of counterfactuals is precisely what lends a good physical theory its explanatory power'. And here is (Curiel, 2015, p.1):[15]

> General relativity poses serious problems for counterfactual propositions peculiar to it as a physical theory. [...] Given the role of counterfactuals in the characterization of, inter alia, many accounts of scientific laws, theory confirmation and causation, general relativity once again presents us with idiosyncratic puzzles any attempt to analyze and understand the nature of scientific knowledge must face.

Indeed, the topic of spacetime counterparts has become even more timely than that of finding invariant observables, since it is germane to current research on the superpositions of gravitational fields (cf. (Kabel et al., 2024) for a recent appraisal). Agreed, 'superspositions of spacetime' is a quantum theme, which I will steer away from here. But, as I will explain in Section 3, Limitations (A) and (B) only take center stage in theoretical physics in the context of quantization or in the treatment of subsystems. Strictly speaking, in the classical domain, and in the context of the whole Universe, there is no need for unique representation: a single representation is good enough, and TS suffices. This is why I partly agree with Fletcher (2020); Weatherall (2018)'s deflation of the hole argument.[16]

two-dimensional gravitational theory called Jackiw-Teitelboim gravity; and Witten (2023) proposes an algebra of operators along an observer's worldline as a background-independent algebra in quantum gravity.

[15]See Jaramillo & Lam (2021), for a more recent echo of this puzzlement, in the context of the initial value formulation of general relativity.

[16]It is important to note that, though these arguments and counterarguments have been given within the context of general relativity, there are closely analogous ones for Yang-Mills theory.

2.3 What are we looking for?

In Section 2.1.b, I reformulated Limitation (2) more precisely as stating two shortcomings of TS. To recap: Limitation (A) is that we have not specified how individual physical possibilities are to be tractably individuated. Limitation (B) is that we have not provided invariant descriptions of objects within each physical possibility, in a way that allows some notion of correspondence—or rather, 'counterpart relations'; cf. Gomes & Butterfield (2023a)—between objects across possibilities.

Focussing on Limitation (A) in Section 2.1, I argued that denumerable lists of geometric invariants at best incompletely characterise an isomorphism-class, or at worst characterise nothing at all. So what *would* succeed in overcoming Limitation (A), by providing consistent representations of isomorphism-classes with tractable conditions of identity? I propose that it is a:

Definition 1 (Representational scheme) *A representational scheme for a given type of (isomorphism-invariant) structure is a complete, isomorphism invariant description of the isomorphism-classes of the theory, such that two such descriptions are identical iff the two isomorphism-classes are identical. Thus, mathematically, such a description is an explicit map* σ *from the space of models of the theory, call it* Φ*, to some fixed value space* V *(the space of tuples, maybe infinite tuples, of values of a complete set of observables) such that, for two models* $\phi, \phi' \in \Phi$*,*

$$\sigma(\phi) = \sigma(\phi'), \qquad \textit{iff} \qquad \phi \sim \phi', \tag{2.1}$$

where the equivalence relation is given by isomorphism.

Two remarks are in order about this definition. First, I said the map is 'explicit': I did this to avoid implicit definitions that are in practice often intractable.[17] Second, it should be clear that I intend the definition to encompass many different choices of equally valid reduced formalisms. Although there is an element of choice, they are mathematically and physically constrained: this is why I called them *schemes* (cf. footnote 6).

The paradigmatically successful example of a characterization that would overcome Limitation (A) and satisfy Definition 1 can be found, as I briefly mentioned in Section 2.1, in the case of vector spaces, e.g. when Φ consists of all finite-dimensional, real vector spaces, whose isomorphisms are invertible linear transformations. Here specifying a single number—the dimension of the space—fully characterizes each isomorphism class. But it is not the finiteness of the characterization that makes it successful: the space of scalar fields such as temperature or density in a fixed spacetime background also admits an equally transparent characterization.

In this case of the temperature field suppose there are no isomorphisms apart from the identity. The descriptions are complete and *local*, in the sense that two temperature distributions are physically identical iff they numerically match at each point of spacetime. So we

[17] The condition excludes, for example, very abstract maps such as the projection onto the equivalence classes: $\pi : \Phi \to \Phi/\sim$. It is not always clear that the definition makes mathematical sense, in the first place; but, even if it did, it would not be an explicit map and thus would fail to satisfy Definition 1.

can tractably assess the consistency of partial descriptions (i.e. descriptions of temperature for regions of spacetime). But what are good examples of representational schemes for field theories with non-trivial isomorphisms, the cases that this paper focuses on?

In order to hone Definitions and find such examples, let us briefly recap both the merits and shortcomings of a description of the isomorphism-classes via fields that admit non-trivial isomorphisms, for instance, the metric tensors g_{ab}. One merit is that they provide clearly consistent and complete descriptions of the invariant structure—e.g. the geometry—of a region where they are defined. But of course, a metric is not isomorphism-invariant, so it does not explicitly provide tractable conditions for the identity of isomorphism-classes (as I argued in Section 2.1, given two metric tensors, it is hard to tell whether they are in the same isomorphism-class, even if it is obvious that each *completely* represents *some* isomorphism-class).

And as to Limitation (B), from a generic metric tensor or connection, one could in principle derive sufficient individuating relations for regions or frames of a vector bundle. For instance, generically we could specify each spacetime point by its web of metric relations to other points. But, as with applying Leibniz's PII in Section 2.1, it is not clear which tractable subset of relations we should take for this purpose. Moreover, in order to be able to formulate counterfactuals about the contents of spacetime regions in different possibilities, such subsets of relations should not fully determine the content of those regions: we don't want spacetime regions to bear their content *essentially*; that would leave no room for interesting counterfactual statements. So we want representational schemes to fix representational redundancy, but not limit the physical content that they are able to represent.

So we are looking for a consistent description of the invariant structure that has the nice properties of the metric and the connection but that can also tractably individuate possibilities and uniquely characterise regions and frames *across possibilities*.

In field theories, representational schemes that satisfy these desiderata are most conveniently provided through choices of gauge fixing. In that case, we require the output of a representational scheme to be itself a model of the original theory. This is why, in Definition 1, I denoted representational schemes with the familiar notation for sections of a fibre bundle, viz σ; in later Sections this will be a central case. So we define:

Definition 2 (Representational scheme via gauge-fixing) *A representational scheme via gauge-fixing is a 1-1 map between the equivalence classes given by isomorphisms of the theory and a subset $\mathcal{F}_\sigma$ of models of the theory, $\sigma : \Phi \to \mathcal{F}_\sigma \subset \Phi$, such that the two models in the subset are identical iff their arguments are isomorphic; i.e. they obey* (2.1).

Although it can be applied more widely, this type of representational scheme has many advantages, particularly in the case of field theories.

The first is that, for descriptions given through representational schemes satisfying Definition 2 the problem of *consistency* of descriptions, as described in Section 2.1, is also tractable. This advantage of Definition 2 over Definition 1 will be explained in Section 3.4.b. The second advantage is tightly related to the first: it is only for representational schemes satisfying Definition 2 that we can straightforwardly overcome Limitation (B).

Lastly, I should mention that it is not only gauge-fixings that bear a relationship to my notion of representational schemes as originally stated in Definition 1. In the current literature, they also encompass 'relational' (or 'material', or even 'quantum') reference frames, as well as the notion of *dressing*. And while a construal via gauge-fixing or dressed quantities is most apt to overcome Limitation (A), a construal via relational reference frames is most apt to overcome Limitation (B). We will visit the relationship between these three construals in what follows (see especially Section 3.3).

In sum: I will show Limitation (A) is overcome when we have provided a representational scheme for our models that satisfies Definition 1. As to Limitation (B), I will show it it is partially overcome for representational schemes via gauge-fixing, i.e. satisfying Definition 2.

3 Representational schemes

Following up on Section 2.2, in Section 3.1 I will briefly describe two contexts in which representational schemes take center stage in theoretical physics; in which, that is, physicists cannot avoid facing Limitations (A) and (B): quantization and the treatment of subsystems. In Section 3.2, I will introduce the mathematics of representational schemes via gauge-fixing. In Section 3.3 I will make the relationship between representational schemes via gauge-fixings, dressings, and relational reference frames explicit. In Section 3.4 I will discuss the nonlocality of representational schemes.

3.1 Representational schemes in physics: a brief history

Section 3.1.a will explain why representaional schemes are important in quantization, and introduce the idea that gauge-fixings, dressings, and relational reference frames may be different sides of the same coin. Section 3.1.b describes the second context in which representational schemes are important: the treatment of subsystems. Here we will see how a representational scheme construed in terms of a reference frame allows us to invariantly describe the physics of subsystems in gauge theory.

3.1.a Gauge-fixing and relational observables in quantization

In Section 2.2, I argued that theoretical physicists are also preocuppied with matters of symmetry and equivalence; more specifically, with overcoming Limitation (A) by finding complete sets of 'observables' that fully describe the theory and tractably individuate isomorphism-classes. But in the classical domain, there is less need for unique representation, or for conditions of individuation in general: any model representing a given physical situation is as good as any other. What does it matter if there are many ways to represent the same physical situation? One is enough.

But passing over to the quantum domain, it matters: one is enough, more than one is too many. In the context of quantization, we need to eliminate redundancy of representation, even for technical reasons. So physicists mostly talk about complete observables in the context of quantization, where, at least in certain approaches it is ineluctable. For instance, in the path integral, or sum-over-histories approach, the tools of perturbative field theory fail if we do not treat the symmetry-related histories as being one and the same, or physically identical.[18]

There are in general two preferred ways to treat symmetry-related histories as one and the same. One approach is to reconstrue the theory, e.g. the path integral, in terms of symmetry-invariant, relational variables: this is the 'purist's' approach; common among experts in general relativity and those working on quantum gravity. In contrast, particle physicists working on gauge theory will not spend much time discussing relational invariant observables. Within that community the usual approach to redundancy is to just "fix the gauge". To fix the gauge is to satisfy Definition 2 by implementing auxiliary conditions that are to be satisfied by a single element in each class of symmetry-related models. Gauge fixing is very useful in practice and widely employed in the actual calculations of gauge theory and (classical and perturbative) gravity. But it is also often (erronously) assumed to do violence to the original symmetry of the theory.

The contrast is clear: in the less applied, more conceptually-driven domains, gauge-fixing is often overlooked in lieu of a choice of 'relational observables' and, more recently, in lieu of 'dressed fields', which are a particular type of relational observables.[19]

But I maintain that this assumption, this contrast sketched above, between gauge-fixing, relational observables (and reference frames) is in one way misleading. For as we will see in Section 3.3.b, gauge fixings can be understood in terms of symmetry-invariant composites of the original fields, called *dressed fields*, and therefore do no violence to the original symmetry of the theory. Dressed fileds give rise to a complete set of observables that can be understood relationally, i.e. as providing a description of the isomorphism class with respect to a physical or *relational reference frame*. On the other hand, some complete sets of 'observables' admit no construal in terms of gauge-fixing; which is why the contrast is not *completely* misleading. In Section 3.3, I will give a conceptually unified account of dressed quantities, material reference frames, gauge-fixings, and relational observables.

3.1.b A brief history of relational reference frames and subsystems

Apart from the more familiar context of quantisation, discussed in the previous Section, representational schemes in the guise of dressed quantities—which will be discussed at length in Section 3.3.b—sometimes understood as descriptions of a physical state relative to a relational reference frame, have recently been invoked in the treatment of subsystems in gauge theory and general relativity. Indeed, our treatment of representational schemes here is inspired by an

[18]In slightly more detail: the propagator involves the inverse of the Hessian around a classical solution; since the Hessian of the action vanishes along the generators of the symmetry, it diverges. Thus one would obtain divergences at every order of the perturbation series: these divergences cannot be neatly separated into energy scales and so they are immune to the machinery of effective field theories.

[19]Bagan et al. (2000); Lavelle & McMullan (1995, 1997) were, to my knowledge, the first to explore the relationship between dressings and gauge-fixings, in the context of QCD.

application to subsystems, which came chronologically first, in (Gomes & Riello, 2017) (where they were described as 'abstract material reference frames', or 'relational connection-forms'). This initial general treatment was illustrated with explicit examples in (Gomes et al., 2019; Gomes & Riello, 2018, 2021), where a concept of 'abstract reference frames' was compared to dressings in gauge theory and gravity; in (Gomes, 2019, 2021a) I gave a conceptual treatment of these ideas (I'll add some detail in Section 7.2).

The reason these notions of dressing or relational reference frames were first introduced in the context of subsystems was that there are subtleties in construing the isomorphisms of subsystem states as dynamical symmetries, subtleties which can be treated by appealing to properties of representational schemes. In particular, there are subtleties about the symmetry-invariance of a bounded subsystem's dynamical structures, such as its intrinsic Hamiltonian, symplectic structure, and variational principles in general.[20]

But by describing a model relative to some subset of degrees of freedom, or from the point of view of 'relational reference frames', our description is rendered explicitly gauge invariant for all gauge transformations, even in the presence of arbitrary boundaries.

Summing up Section 3.1: in Section 2.1 I justified Limitations (A) and (B) from a conceptual and mathematical perspectice. Here I argued that it is clear that the Limitations pertain also to physics; and indeed they must be resolved when dealing with quantum systems, subsystems, and counterfactuals.

3.2 Representational schemes via gauge-fixing: general properties

At the abstract level explored in this and the next Sections, there are no substantive differences between diffeomorphisms and the symmetries of gauge theories. Indeed, even those of non-relativistic particle mechanics, which mostly lie outside of the scope of this paper, can still be

[20]The point being that dynamical structures on bounded subsystems are *not* invariant under arbitrary gauge transformations of the boundary state, so, in order to preserve gauge-invariance at the level of the subsystems, the boundary state must be understood as 'already gauge-fixed', even without an explicit gauge-fixing condition. This would in practice fix gauge transformations to the identity at the boundary of a subsystem. In effect, fixing the gauge transformations in this way, for finite boundaries, requires the stipulation of a single value for the fields at the boundary. Of course, this may be very well for a particular subsystem, but it is not a condition we would wish to enforce for *any* subsystem, especially if we are to take subsystems of subsystems. In contrast, a representational scheme does not impose any physical limitation on the boundary states, and so restrictions to subsystems is explicitly recursive here.

The modern guise of the obstruction to gauge-invariance posed by boundaries was first highlighted in (Donnelly & Freidel, 2016) (which also introduced 'edge-modes' as new degrees of freedom that fixed this obstruction), which caused a flurry of papers on subsystems in gauge theory (cf. e.g. (Carrozza & Höhn, 2022; Donnelly & Freidel, 2016; Geiller, 2017; Geiller & Jai-akson, 2020; Gomes, 2019; Gomes et al., 2019; Gomes & Riello, 2021; Ramirez & Teh, 2019; Riello, 2021) and references therein). (Gomes & Riello, 2017) introduced relational reference frames as a tool to restore gauge-invariance in a subsystem-recursive manner. It was there fist suggested that the resolution of the gauge-invariance problem did not require new degrees of freedom at the boundary, but physical reference fields with respect to which one should describe variations *throughout* the subsystems, and not just at the boundary. See (Carrozza & Höhn, 2022) for a more recent and comprehensive review of the relationship between dressings, relational reference frames, and gauge-fixings.

encompassed by the formalism. And so we treat all of these symmetries uniformly, labelling them with the group $\mathcal{G}$, which could be infinite-dimensional.

Given the space of models of a theory, Φ, we assume it has a smooth manifold structure (infinite-dimensional, in the case of field theories) and it admits an action of $\mathcal{G}$, i.e. a map $\Phi : \mathcal{G} \times \Phi \to \Phi$, that preserves the global structure on Φ (e.g. is smooth, in the topology of Φ), and preserves dynamics, e.g. the Hamiltonian, the action functional, or the equations of motion (in the language of Gomes (2021c), they are S-symmetries). More formally: there is a structure-preserving map, μ, on Φ that can be characterized element-wise, for $g \in \mathcal{G}$ and $\varphi \in \Phi$, as follows:

$$\begin{aligned} \mu : \mathcal{G} \times \Phi &\to \Phi \\ (g, \varphi) &\mapsto \mu(g, \varphi) =: \varphi^g. \end{aligned} \tag{3.1}$$

The symmetry group partitions the space of models into equivalence classes in accordance with an equivalence relation, $\sim$, where $\varphi \sim \varphi'$ iff for some g, $\varphi' = \varphi^g$. We denote the orbit of φ under $\mathcal{G}$ by $\mathcal{O}_\varphi := \{\varphi^g, g \in \mathcal{G}\}$, which as a set is isomorphic to the equivalence class of φ under the equivalence relation, usually denoted by square brackets, $[\varphi]$.[21] Writing the canonical

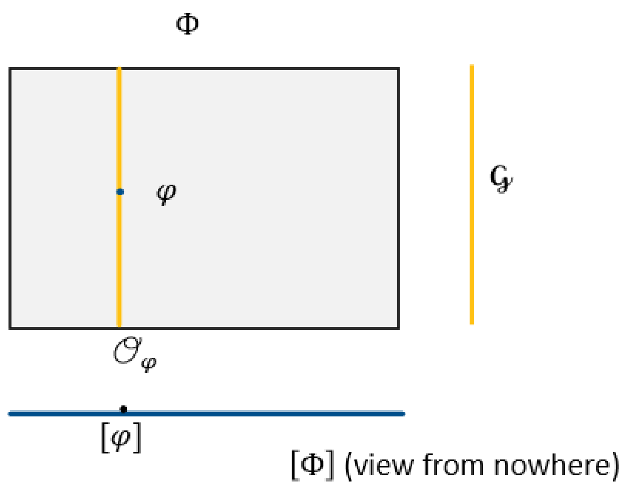

Figure 1: The space of states, 'foliated' by the action of some group $\mathcal{G}$ that preserves the value of some relevant quantity, S, and the space of equivalence classes. In field theory, S is the value of the action functional, and the spaces $\mathcal{G}$, Φ, and $[\Phi]$ are infinite-dimensional (Frechét) manifolds.

projection operator onto the equivalence classes, $\mathsf{pr} : \Phi \to \Phi/\mathcal{G} =: [\Phi]$, taking $\varphi \mapsto [\varphi]$, then the orbit $\mathcal{O}_\varphi$ is the pre-image of this projection, i.e. $\mathcal{O}_\varphi := \mathsf{pr}^{-1}([\varphi])$.

We can now improve on Definition 2 with the more precise:

[21] I here see no important difference between $[\varphi]$ and $\mathcal{O}_\varphi$; the only difference is that the latter is usually seen as an embedded manifold of Φ, whereas the former exists only implicitly, or abstractly, outside of Φ (which is why I called it 'the view from nowhere'; see Figure 1).

Definition 3 (Representational scheme via gauge-fixing) *A representational scheme via gauge-fixing is an injective map*

$$\begin{aligned}\sigma : [\Phi] &\to \Phi \\ [\varphi] &\mapsto \sigma([\varphi])\end{aligned} \tag{3.2}$$

that respects the required mathematical structures of Φ*, e.g. smoothness or differentiability and is such that* $\mathsf{pr}(\sigma([\varphi])) = [\varphi]$*, where* pr *is the canonical projection map onto the equivalence classes.*

Armed with such a choice of representative for each orbit, a generic model φ could be written uniquely as some doublet $\varphi = ([\varphi], g)_\sigma := \sigma([\varphi])^g$, which of course satisfies: $\varphi^{g'} = (\sigma([\varphi]))^{gg'} = ([\varphi], gg')_\sigma$.[22] Thus we identify $\Phi \simeq [\Phi] \times \mathcal{G}$ via the diffeomorphism:

$$\begin{aligned}\overline{\sigma} : [\Phi] \times \mathcal{G} &\to \Phi \\ ([\varphi], g) &\mapsto \sigma([\varphi])^g\end{aligned} \tag{3.3}$$

Now, as I mentioned, the space $[\Phi]$ is abstract, or only defined implicitly; this is why I called it "the view from nowhere". Since we cannot usually represent elements $[\varphi]$ of $[\Phi]$ intrinsically, we in practice replace σ by an equivalent projection operator that takes any element of a given orbit to the image of σ:

Definition 4 (Projection operator for σ) *A map*

$$\begin{aligned}h_\sigma : \Phi &\to \Phi \\ \varphi &\mapsto h_\sigma(\varphi) = \sigma([\varphi]),\end{aligned} \tag{3.4}$$

is called the projection operator for σ *of Definition 3. It maps a model to an isomorphic copy chosen as a representation of its orbit by* σ*. The end result,* $h_\sigma(\varphi)$*, is called a* dressed φ*.*

Since $[\varphi^g] = [\varphi]$, we must have

$$h_\sigma(\varphi^g) = h_\sigma(\varphi). \tag{3.5}$$

Now, $h_\sigma(\varphi)$ is an explicit map on Φ, i.e. it is a function of φ. So h_σ uniquely and concretely represents the invariant, or structural, content. In other words, two given models, φ, φ', give the same value for *all* symmetry-invariant quantities iff $h_\sigma(\varphi) = h_\sigma(\varphi')$. Thus the invariant structure of each isomorphism-class $[\varphi]$ is represented uniquely, according to σ, by the corresponding value of the map $h : \Phi \to \Phi$.

3.3 Gauge-fixing, dressing, and relational frames in field theories

The three Sections below will respectively focus on gauge-fixing, dressings, and relational frames. The common thread in this Section, that I will argue runs through each of these three topics, is the notion of representational scheme.

[22] Our notation is slightly different than Wallace (2019, p. 9)'s, who denotes these doublets as (O, g) (in our notation $([\varphi], g)$), and labels the choice of representative (or gauge-fixing) as φ_O (our φ_σ). We prefer the latter notation, since it makes it clear that there is a choice to be made.

3.3.a Gauge-fixing

In equation (3.3), describing the product structure of the space of models, I oversimplified: finite-dimensional principal bundles need not be trivial, and, accordingly, the space of models need not be globally isomorphic to the product between the space of physical states and the group of gauge transformations, $[\Phi] \times \mathcal{G}$. Indeed, even a local product form for the space of models is only guaranteed to exist in the finite-dimensional case.

Nonetheless, it is in fact true that the space of models Φ—for both Riemannian metrics and gauge connection-forms—is mathematically very similar to a principal fiber bundle, with $\mathcal{G}$ as its structure group (resp. diffeomorphisms and vertical automorphisms). But there are important differences between the infinite-dimensional and the finite-dimensional case. In the finite-dimensional case, a free (and proper) action of the group on the manifold suffices for that manifold to have a principal G-bundle structure, usually written as $G \hookrightarrow P \to P/G$. In the infinite-dimensional case, these properties of the group action are not enough to guarantee the necessary fibered, or local product structure: one has to construct that structure by first defining a *section*. I summarise the construction of that structure and its main obstructions in Appendix A.

A choice of section is essentially a choice of embedded submanifold on the model space Φ that intersects each orbit exactly once. In the 2-dimensional figure 1 above, this would be represented by a 'nowhere-vertical' curve: i.e. a submanifold that is transversal to all the $\mathcal{G}$-orbits and intersects each orbit once.

A convenient way to define such sections is to impose further functional equations that the model in the aimed-for representation must satisfy; this is like defining a submanifold through the regular value theorem: in finite dimensions, defining a co-dimension k surface $\Sigma \subset N$ for some $n > k$-dimensional manifold N, as $\mathcal{F}_\sigma^{-1}(c)$, for $c \in \mathbb{R}^k$, and $\mathcal{F}_\sigma$ a smooth and regular function, i.e. $\mathcal{F}_\sigma : N \to \mathbb{R}^k$ such that $\mathsf{ker}(T\mathcal{F}_\sigma) = 0$.

Once the surface is defined, σ can be seen as the embedding map with range $\sigma([\Phi]) = \mathcal{F}_\sigma^{-1}(0) \subset \Phi$. The next step is to find a gauge-invariant projection map, h_σ, that projects any configuration to this surface.

3.3.b Dressing

I have called $h_\sigma(\varphi)$ the dressed variables. What are its clothes? In the case of dressings associated to gauge-fixings, they are the gauge transformation $g_\sigma(\varphi)$ required to transform φ to a configuration $\varphi^{g_\sigma(\varphi)}$ which belongs to the gauge-fixing section σ.[23]

Here we encounter the dual interpretations of gauge-fixings and dressed variables. In the

[23] This is not the most general type of dressing function. Firstly, there is the infinitesimal version, given by the relational connection-form ϖ (cf. Gomes & Riello (2017, 2018, 2021) and Gomes et al. (2019); and Gomes (2019, 2021a) and (Gomes & Butterfield, 2023a, Appendix A) for philosophical introductions). I will briefly summarise this construction in Section 7.2. Second, there is a more complete theory of dressings and dressed fields, which only requires the right covariance properties and can restrict to subgroups of the gauge group: see François (2021); François et al. (2021) for a review.

case of fields, the projection g_σ is a function of φ that is usually non-local and 'relational', in the sense that it stands for a comparison of the values of different fields at different points. Conceptually, the image of this projection—the dressed field—is a symmetry-invariant composite, or function, of the original degrees of freedom, which uniquely describes each isomorphism-class relationally. The local interpretation as a gauge-fixed field has the field satisfying an auxiliary set of differential equations besides their equations of motion, by belonging to $\mathcal{F}_\sigma$. We can thus see such conditions as implicitly defining the invariant relational composites that we call 'dressed variables'.

In more detail: a function $\mathcal{F}_\sigma : \Phi \to W$, valued on some general vector space W, defining the section (surface) $\sigma([\Phi])$ as its level surfaces $\mathcal{F}_\sigma^{-1}(0)$ is suitable as a gauge-fixing iff it gives rise to a unique dressing by satisfying two conditions:

• *Universality* (or existence): For all $\varphi \in \Phi$, the equation $\mathcal{F}_\sigma(\varphi^g) = 0$ must be solvable by a dressing $g_\sigma : \Phi \to \mathcal{G}$.

That is, there is a Φ-structure-preserving map,

$$g_\sigma : \Phi \to \mathcal{G}, \text{ such that } \mathcal{F}_\sigma(\varphi^{g_\sigma(\varphi)}) = 0, \text{ for all } \varphi \in \Phi. \tag{3.6}$$

This condition ensures that $\mathcal{F}_\sigma$ doesn't impose *physical* constraints on what can be represented, i.e. that each orbit possesses at least one intersection with the gauge-fixing section.

• *Uniqueness*: g_σ is the unique functional solution of $\mathcal{F}_\sigma(\varphi^{g_\sigma(\varphi)}) = 0$. This condition ensures that each orbit is represented uniquely.

From these two conditions, it follows that $\varphi^{g_\sigma(\varphi)} = \varphi'^{g_\sigma(\varphi')}$ if and only if $\varphi \sim \varphi'$, meaning that models have the same projection onto the section iff they are isomorphic. In one direction, the claim is immediate: $\varphi^{g_\sigma(\varphi)} = \varphi'^{g_\sigma(\varphi')}$ implies $\varphi = \varphi'^{g_\sigma(\varphi')^{-1}g_\sigma(\varphi)}$, so $\varphi \sim \varphi'$. In the other direction, assume $\phi' = \phi^g$ and g_σ is unique. From (3.6), since $[\varphi^g] = [\varphi]$ and $\sigma : [\Phi] \to \Phi$ is a unique embedding :

$$\varphi^{g_\sigma(\varphi)} = \sigma([\varphi]) = \sigma([\varphi^g]) = (\varphi^g)^{g_\sigma(\varphi^g)} = \varphi'^{g_\sigma(\varphi')}, \tag{3.7}$$

which establishes the other direction. And we obtain from (3.7) the following equivariance property for g_σ:[24]

$$g_\sigma(\varphi^g) = g^{-1} g_\sigma(\varphi). \tag{3.8}$$

In possession of a dressing function g_σ, it is convenient to rewrite the dressed field h_σ of (3.4) as $h_\sigma(\bullet) := \bullet^{g_\sigma(\bullet)}$, i.e:

$$\begin{aligned} h_\sigma : \Phi &\to \Phi \\ \varphi &\mapsto h_\sigma(\varphi) := \varphi^{g_\sigma(\varphi)} \end{aligned} \tag{3.9}$$

And of course, we can still change the representational scheme itself, i.e. act with the group on the image of h_σ: because $h : \Phi \to \Phi$ is a projection (as opposed to a reduction $\mathsf{pr} : \Phi \to [\Phi]$),

[24]These equations are schematically identical to the ones we find for Yang-Mills theory in the finite-dimensional case. In the Yang-Mills case, a section σ in the space of models is essentially a section σ of the finite-dimensional principal bundle, P that depends on the isomorphism-class.

and thus $h_\sigma(\varphi) \in \Phi$, this gives $(h_\sigma(\varphi))^g$. But $(h_\sigma(\varphi))^g$ no longer obeys the conditions defining the previous representational scheme, $\mathcal{F}_\sigma(\varphi^{g_\sigma(\varphi)}) = 0$. If these conditions are taken to define, for intra-theoretical, non-empirical reasons, 'acceptable' representations, $(h_\sigma(\varphi))^g$ would not be an 'acceptable' representation of that orbit. Nonetheless, its representation could be acceptable for a different set of intra-theoretical reasons.

Moreover, given any two representational schemes σ, σ', we have

$$\varphi^{g_\sigma(\varphi)} = (\varphi^{g'_\sigma(\varphi)})^{g'_\sigma(\varphi)^{-1}g_\sigma(\varphi)} \tag{3.10}$$

and so we obtain

$$h_\sigma(\varphi) = (h_{\sigma'}(\varphi))^{\mathsf{t}_{\sigma\sigma'}(\varphi)}, \tag{3.11}$$

where $\mathsf{t}_{\sigma\sigma'}(\varphi) := g'_\sigma(\varphi)^{-1}g_\sigma(\varphi)$ is a model-dependent isomorphism (the analog of the transition map between sections of a principal bundle). Although, from the covariance of g_σ, given in (3.8), it immediately follows that the transition map is itself gauge invariant, as expected, depending only on the isomorphism-classes, or orbits $\mathcal{O}_\varphi$. Note also that the dressing according to one scheme of a model dressed according to a different scheme acquires an inverse upon switching the schemes, i.e.

$$g_{\sigma'}(\varphi^{g_\sigma(\varphi)}) = g_\sigma(\varphi)^{-1}g_{\sigma'}(\varphi) = (g_\sigma(\varphi^{g'_\sigma(\varphi)}))^{-1} \tag{3.12}$$

3.3.c Relational reference frames

Finally, as mentioned in Section 3.1.b, dressed variables have also been related to 'material reference frames', where 'material' has nothing to do with 'matter'; it is meant as 'physical', or 'relational' (I'll stick to relational.) Relationism of course has a noble tradition, and the broad idea of a 'relational reference frame' is a staple of research in quantum gravity (cf. e.g. (Rovelli, 2007)). The idea is that we can anchor a particular representation to a physical system, with the ensuing representation being straightforwardly understood in terms of relations to this physical system.

For instance, given a set of four scalar fields obeying functionally independent Klein-Gordon equations, we can understand DeDonder gauge as using these fields as coordinates on a region of spacetime (see Section 4.3). Similarly, in gauge theory, a particular field could select a particular internal frame for a vector bundle, and we would describe other fields relative to this frame. In electromagnetism, unitary gauge can be understood in this way (see Section 4.2.b). In each of these cases, g_σ can be understood as the transformation required to go from an arbitrary representation of the isomorphism-classes to a preferred representation relative to this 'relational reference frame'. In the case of spacetime, by locating points relative to this physical system, we make precise a frequent claim found in the literature on TS about spacetime; namely that 'spacetime points can only be specified by their web of relations to other points.'

But in the case of reference frames we must issue a caveat that does not usually accompany these claims. The caveat is germane to Limitation (B) and it is that this web of relations cannot be too rigid, otherwise each region will bear their specific content essentially. In other words, if the web fixes a region but also entirely fixes the content of that region, no interesting counterfactuals can be expressed for it. In the context of the hole argument, this kind of rigidity leads to (Maudlin, 1988, 1990)'s 'metric essentialist' resolution, which indeed has been criticised for being an argument valid only for the 'actual world' (Butterfield, 1989); and in (Gomes & Butterfield, 2023a) precisely for not allowing the expression of counterfactuals.

The use of relational reference frames to solve Limitation (B), about counterfactuals, is thus contingent on their association with a gauge-fixing, and in particular, its property of Universality. Conceptually, we must use just enough relations between parts of the fields to fix the representation but not limit the physical content.[25]

But note that, if we associate 'relational reference frames' to gauge-fixings in this way, we impose a constraint on what kind of reference frames we are countenancing: only those that are related by the (model-dependent) symmetries of the theory, as illustrated by (3.11). And though certain physical systems may sufficiently anchor the representation, they may fail to fulfill this requirement, and thus fail to form *bona-fide* relational reference-frames in my sense (which should be associated to gauge-fixings). This is not a bad thing, since we should expect that systems described by reference frames not thus associated by symmetries of the laws would not obey the same laws.

We will see specific examples of reference frames associated to gauge-fixings in Section 4 (in particular, in Sections 4.2.b and 4.3), and we will employ this construal explicitly in overcoming Limitation (B) in terms of counterpart theory in Section 6.

3.4 Non locality

This Section deals with aspects of non-locality associated to dressings and gauge-fixings. In Section 3.4.a I will discuss the relationship between dressings, constraints, and non-locality. The kind of non-locality that emerges via dressings in gauge theory forces us to again face the question of *consistency* of partial descriptions of an isomorphism-class. This question was introduced in Section 2.1, where I argued that descriptions via a denumerable list of invariants was generally inconsistent. In Section 3.4.b, I will show that, for representational schemes via gauge-fixing, the question of consistency of piecewise-invariant descriptions, described in Section 2.1, has a neat answer.

3.4.a Dressing and constraints

In Section 4, I will give examples of dressings obtained via representational schemes, most of which are gauge-fixings. And most, whether obtained via a gauge-fixing or not, will display some degree of non-locality. This feature is expected for generic states of gauge theory and general relativity.

For instance, even without appealing to a gauge-fixing, it is well-known that diffeomorphism-invariant functions are usually non-local: this is a notorious problem for quantum gravity (see

[25] Indeed, it is for this reason that relational reference frames associated to gauge-fixings are able to restore a gauge-invariant notion of recursivity for arbitrary subsystems: see footnote 20.

Donnelly & Giddings (2016); Rovelli (2007); Thiemann (2003)). Thus Isham (1992, p. 170) continues the quote of Section 2.2:

> In the present context, the natural objects [that constitute 'observables' in general relativity] are Diff(M)-invariant spacetime integrals [...] Thus the 'observables' of quantum gravity are intrinsically non-local.

Why is it that dressed quantities are generally non-local? Let me try to give a brief conceptual explanation. A local symmetry implies, via Noether's second theorem (cf. Brading & Brown (2000) for a conceptual exposition), that the equations of motion of the theory are not all independent, and thus they only uniquely determine the evolution of a subset of the original degrees of freedom. There is thus a certain freedom in choosing which 'components of the fields' will be uniquely propagated, or will evolve deterministically—each choice corresponds to a gauge-fixing.

For general relativity and Yang-Mills theory, such choices of initial data must satisfy elliptic equations, whose solutions require integrals over an initial spacelike surface. As I argue at further length in Gomes (2021e), in practice, ellipticity means that boundary value problems require only the spatial configuration of the field at the boundary, i.e. they do not also require the field's rate of change at the boundary. Thus the solution of these equations exists *on each simultaneity surface*: they do not describe the propagation of a field, as would a solution to a hyperbolic equation, but valid initial data at the initial surface. Thus non-locality arises because the function that takes the original local degrees of freedom to a uniquely propagated subset is, generally, non-local: the value of an element in the subset at point x depends on the values of the original degrees of freedom at other points on the spatial surface (see Section 4.2.a for a clear example, and (Gomes & Butterfield, 2022, Sec. 1.1, point (3)) for more discussion about this sort of (non-signalling) classical non-locality).[26]

In sum, gauge-fixings or dressed representations generally require this particular (rather benign) type of non-locality, that should rather be called 'holism', or 'non-separability', and that applies to theories with elliptic initial value constraints—such as general relativity and Yang-Mills theories—but does not apply to theories like Klein-Gordon in a fixed Minkowski background, nor to any 'gauge theory' whose gauge transformations involve no derivatives.

I'll finish this Section with a quote from (Harlow & Wu, 2021, p. 3), who, in a recent paper pursuing a symmetry invariant description of a particular gravitational theory, nicely summarise the relationship between dressings, constraints, gauge-invariance, and relationism:

> It has long been understood that diffeomorphism symmetry must be a gauge symmetry, and that physical observables must therefore be invariant under almost all

[26]There are two points to note: (1) it is clear that relativistic causality holds for the isomorphism-invariant facts, since (quasi-)hyperbolicity of the equations of motion ensures causality is respected for one choice of metric or gauge potential representing the isomorphism-equivalence class (e.g. Lorenz or DeDonder gauge for electromagnetism and general relativity, respectively). (2) This argument for non-locality does not rule out that some observables are local! Of course there can be such observables (e.g. $F_I^{\mu\nu}F^I_{\mu\nu}$). But local observables cannot be all of the observables if we cannot parametrise valid initial data with them.

diffeomorphisms. [...] There is an analogous situation in electromagnetism: fields which carry electric charge are unphysical unless they are "dressed" with Wilson lines attaching them either to other fields with the opposite charge or to the boundary of spacetime. One way to think about dressed observables in electromagnetism is that they create both a charged particle and its associated Coulomb field, which ensures that the resulting configuration obeys the Gauss constraint. Similarly in gravity any local observable by itself will not be diffeomorphism-invariant, so we must dress it [...creating] for it a gravitational field that obeys the constraint equations of gravity. In practice such observables are usually constructed by a "relational" approach: rather than saying we study an observable at some fixed coordinate location, we instead define its location relative to some other features of the state.

3.4.b Consistency of partial descriptions and gluing: the advantage of gauge-fixing

A question about non-locality remains: how does such an ubiquituous feature of dressings differ from the type of holism illustrated by denumerable lists of geometric invariants, that will often lead to an inconsistent description, as discussed in Section 2.1? I will now show that, in the case of dressings obtained via gauge-fixing of field theories, the consistency of piecewise-invariant descriptions is straightforward.

Briefly, suppose two regions, R_+ and R_- (or $R_\pm$ for short) overlap on a region $R_0 := R_+ \cap R_-$. Further suppose that the union of all three regions is a manifold, and that each region is an embedded submanifold within this union. Consider the fields intrinsic to each region, and their intrinsic isomorphisms; call these $\Phi_\pm, \Phi_0$ and $\mathcal{G}_\pm, \mathcal{G}_0$, respectively. For given gauge-fixings $\sigma_\pm$ of $\mathcal{G}_\pm$ on $R_\pm$, we have $h^+(\varphi^+) \in \Phi_+$, and $h^-(\varphi^-) \in \Phi_-$; which we again call $h^\pm$ for short. Then the $h^\pm$ are consistent iff given a representational scheme σ_0 (not necessarily via gauge-fixing) on R_0,

$$h^0(h^+_{|R_0}) = h^0(h^-_{|R_0}), \tag{3.13}$$

where $h^\pm_{|R_0}$ represents the restriction of the $h^\pm$ to the region R_0, and so $h^\pm_{|R_0} \in \Phi_0$. Equation (3.13) means that the $h^\pm_{|R_0}$ are related by a (isomorphism-class-dependent) transformation $\mathfrak{t}_\pm$, as described in (3.11). Thus $h^\pm$ are consistent, or can be 'glued', or composed into a single global model, iff (3.13) holds, for any σ_o.

For a local dressing, $h(\varphi)(x) = h(\varphi(x))(x)$: the dressing at x depends only on the values of the fields being dressed at x. So, for the same representational scheme chosen throughout the manifold:

$$h^+(\varphi_{|R_+}) = h(\varphi)_{|R_+}, \quad \text{thus} \quad h^+_{|R_0} = h_{|R_0} = h^-_{|R_0}. \tag{3.14}$$

So, although (3.13) applies to both local and non-local dressings, it is only non-trivial in the non-local case. In the non-local case, even if the choice of gauge-fixing on R_+ was the same as that in R_- (e.g. Coulomb gauge), the $h^\pm$, when restricted to R_0, may differ.[27] And yet, even if

[27] The non-locality of dressings is equivalent to the non-commutativity between regional restrictions and the

$h^{\pm}_{|R_0}$ differ, they may still be different representations of the same invariant structure for that region. This is why, in the non-local case, it is important that the dressed quantities $h^{\pm}$ are also models of the theory. For, in order to assess consistency of the piecewise dressed quantities, these quantities must enter as arguments of the representational scheme h_0 in (3.13). When σ_0 is also a representational scheme via gauge-fixing, the consistency condition given in (3.13) is checked by composing the solution of two differential equations.

Gluing or composing gauge-invariant quantities in the non-local case requires us to match different representations of the same isomorphism-classes, and to do that, we need transformations that function like gauge transformations: the $\mathfrak{t}_{\sigma\sigma'}$, as given in (3.11) (see also footnote 30). Indeed, because of this requirement, along the lines of an argument originally made by Rovelli (2014) in the case of particles and different sets of fields, in (Gomes, 2019) it was argued that the importance of gauge is tied to non-locality. The 'gluing' and composition of gauge-fixed quantities is described in detail for Yang-Mills theory in (Gomes & Riello, 2021, Sec. 6).

4 Examples of Representational Schemes

Representational schemes can be very general. Here I will provide several examples, mostly based on gauge-fixings, starting from non-relativistic particle mechanics to gauge theory and general relativity.

4.1 Non-relativistic particle mechanics

4.1.a Center of mass for non-relativistic point-particles

Take Newtonian mechanics for a system with N particles, written in configuration space as trajectories $\varphi(t) = q_\alpha(t) \in \mathbb{R}^{3N}$ (with $\alpha = 1, \ldots, N$ labeling the particles), and with shift symmetry:

$$\varphi^g = (q^\alpha(t) + g), \quad \text{with } g \in \mathbb{R}^3. \tag{4.1}$$

Now we take the section $\mathcal{F}_\sigma$ to be defined by the center of mass:

$$\mathcal{F}_\sigma(q) = \sum_\alpha q^\alpha m^\alpha = 0. \tag{4.2}$$

dressing map. That is, suppose we are given a representational scheme σ, which is applicable to any region, e.g. by stipulating a gauge-fixing and boundary condition. Suppose further that the dressing function requires solving an elliptic equation (as in e.g. Coulomb gauge, cf. Section 4.2.a). Then for a generic submanifold R, with boundary ∂R:

$$h(\varphi)_{|R} \neq h(\varphi_{|R}). \tag{3.15}$$

This is easy to see in terms of boundary conditions: although $h(\varphi_{|R})$ will necessarily obey the stipulated boundary condition on ∂R, the values of $h(\varphi)_{|\partial R}$ will be rather arbitrary, and highly dependent on which surface we choose. And the difference won't be constrained to ∂R of course, since solutions of elliptic equations are globally dependent on the boundary condition. This is discussed at lengh, including complications about whether the region-intrinsic isomorphisms are also dynamical symmetries of the theory, in (Gomes, 2022, Ch. 6).

The most straightforward way to find the projection $h(q)$ is displayed in (3.9). It requires us to find the dressing as a functional of the configuration. That is, take

$$h^\alpha(q) = q^\alpha + g_\sigma(q), \tag{4.3}$$

and solve

$$\mathcal{F}_\sigma(h(q)) = 0 \quad \text{ for } g_\sigma : \Phi \to \mathbb{R}^3 \text{ and arbitrary } \; q. \tag{4.4}$$

We obtain:

$$\sum_\alpha (q^\alpha + g_\sigma(q))m^\alpha = 0 \Rightarrow g_\sigma(q) = -\frac{\sum_\alpha q^\alpha m^\alpha}{\sum_\alpha m^\alpha}.$$

Clearly,

$$g_\sigma(\varphi^g) = g_\sigma(q+g) = g_\sigma(q) - g. \tag{4.5}$$

Thus if follows that the new set of dressed configuration variables, written as functionals of arbitrary configuration variables, is gauge-invariant (now including indices for both particles and components in $\mathbb{R}^3$, to be completely explicit):

$$h^\alpha_i(q^\alpha + g) = (q^\alpha_i(t)) + g + g_\sigma(q) - g = h^\alpha_i(q^\alpha). \tag{4.6}$$

And it is also easy to see that, if $h^\alpha(q) = h^\alpha(q')$, then $q^\alpha - q'^\alpha$ is given by a shift, and so $h^\alpha(q) = h^\alpha(q')$ iff q and q' are isomorphic.

This example is still non-local, in the sense of Section 3.4: given two particle subsystems, indexed by $N_+, N_- \subset N$, with $N_+ \cup N_- = N$,[28] the center of mass of either subsystem is generically not identical to the center of mass of their union. But, in the notation of Section 3.4.b, for $N_0 = N_+ \cap N_-$, it is straightforward to verify that, although $h^+_{|N_0} \neq h^-_{|N_0}$ (since the particles in N_0 are being described relative to different centers of mass), if $N_+, N_- \subset N$, with $N_+ \cup N_- = N$, equation (3.13) will hold, since both sides correspond to the description of the N_0 subsystem with respect to its own center of mass.[29]

4.1.b Inter-particle distances

This is the first example of a representational scheme which is *not* equivalent, at least *prima facie*, to any gauge-fixing. We take the same system as in the previous example, described in (4.1). And instead of (4.1), we take, following Rovelli (2014), a new complete set of relational variables:

$$\bar{q}_\alpha(t) := (q_{\alpha+1} - q_\alpha), \quad \text{for } \alpha = 1, \cdots, N-1. \tag{4.7}$$

[28] I am here denoting a set of particle labels with $N_\pm, N$, whereas before N was a number. But this slight innacuracy does not justify the amount of notation that would need to be introduced in order to clarify this point.

[29] Of couse, we need not have $N_+, N_- \subset N$, with $N_+ \cup N_- = N$. That is, we need not have h_+, h_- coming from the restrictions of a consistent global model: equation (3.13) is supposed to assess consistency, not assume it. If $h_0(h^+_{|N_0}) \neq h_0(h^-_{|N_0})$ there is a bijection between the labels of the particles in N_0 according to N_- and N_+, but the intrinsic statess of the restricted subsystems are not symmetry-equivalent.

Thus the new variables provide a shift-invariant description of the system. They are complete, since the dimension of the space of configurations of N particles in e.g. $\mathbb{R}^3$ is $\mathbb{R}^{3N}$, and modulo translations, this is $\mathbb{R}^{3(N-1)}$. So they satisfy Definition 1. But there is a clear disadvantage to this kind of description: because the invariant variables are not elements of the original space of models, Φ, it is not a representational scheme in the sense of Definition 2 (or the more precise Definition 3). So we cannot iteratively apply representational schemes in order to assess consistency, as in (3.13).[30]

Of course, examples for non-relativistic particle mechanics abound: beyond the case of translations and boosts, we could fix a frame in $\mathbb{R}^3$ by diagonalizing the moment of inertia tensor around the center of mass; in this case g_σ would be a state-dependent rotation (see (Gomes, 2021b, Sec. 4) for details).[31]

4.2 Gauge theory

4.2.a Electromagnetism: Coulomb and Lorenz gauge

Electromagnetism in the gauge potential formalism is a gauge theory, and in it, Lorenz gauge is an explicitly Lorentz covariant choice, which has its virtues thoroughly extolled by Mattingly (2006), who argues that it should be considered as Maudlin (1998)'s "ONE TRUE GAUGE". Maudlin (2018) himself endorses a different choice, Coulomb gauge.

Coulomb gauge arises from a split of the electric field into a component that is purely Coulombic, or determined by the synchronic distribution of charges, and another component that is purely radiative. And, apart from worries with *Uniqueness* (described in Section 3.3.b, associated to the Gribov problem (discussed in Appendix A), there are straightforward extensions of these gauge-fixings to the non-Abelian case. This particular example of $\mathcal{F}_\sigma$ (and h_σ, and g_σ) in field theory is fully worked out by Gomes & Butterfield (2022), alongside its physical interpretation. Lorenz gauge is explicitly Lorentz covariant, and has the representation of the gauge potential as determined by the past distribution of currents (see footnote 32).

Although conceptually very different, these two gauges are in fact mathematically very similar: the main difference is that the Lorenz gauge concerns all the components of the potential in a Lorentzian manifold, whereas Coulomb gauge concerns only spatial components of the gauge potential, so is described in a Riemannian manifold. But here I will ignore the

[30] Suppose we were coupling two subsystems for which $N_+ \cap N_- = \emptyset$. Now we have $\mathbb{R}^{3(N_\pm-1)}$ degrees of freedom in each subsystem. But, after coupling, we have $\mathbb{R}^{3(N_++N_--1)} \neq \mathbb{R}^{3(N_+-1)} + \mathbb{R}^{3(N_--1)}$. Indeed, the coupled system has 3 more degrees of freedom than the individual subsystems. Rovelli (2014) argues that these are the gauge degrees of freedom that need to be retained for the coupling of subsystems. Gomes & Riello (2021) show that the same can be applied to gauge theories. The difference is that if we restrict the division of the fields to supervene on a complete division of the manifold by complementary regions, the number of extra, holistic degrees of freedom is the dimension of the stabilliser group of the states at the boundaries between the complementary regions. See Gomes (2021a) for a conceptual exposition.

[31] Configurations that are collinear, or are spherically symmetric, etc. would be unable to fix the representation: these are *reducible* configurations; they have stabilizers that cannot be fixed by any feature of the structure that is isomorphism-invariant.

difference by construing Lorenz gauge in a spacetime manifold with Euclidean signature, and Coulomb gauge to be a choice of gauge in the Hamiltonian version of electromagnetism.[32]

In the Coulomb case for Hamiltonian electromagnetism, the field φ is a doublet that transforms only in one component, i.e.

$$\varphi = (E^i, A_i), \quad \varphi^g = (E^i, A_i + \nabla_i g), \tag{4.8}$$

with $g \in C^\infty(M)$, where M is a spatial (Cauchy) surface, E_i is the electric field, and A_i is the spatial gauge potential, with spatial indices i, j etc. The Lorenz case would be similar, but M would be the spacetime manifold and we would have:

$$\varphi = A_\mu, \quad \varphi^g = A_\mu + \nabla_\mu g, \tag{4.9}$$

with μ, ν, etc, spacetime indices. In the following, I will subsume both cases using the abstract index notation A_a.

First, following the definition of the dressed function, in (3.9), we write:

$$h(\mathbf{A})_a := A_a + \nabla_a g_\sigma(\mathbf{A}), \tag{4.10}$$

where $g_\sigma : \Phi \to C^\infty(M)$, and M is either the space or the spacetime manifold (with Euclidean signature), and Φ is the space of gauge potentials over M satisfying appropriate boundary conditions (which I will not here discuss: for more details, about boundary conditions and the point that this gauge is not complete in the Lorentzian signature: see footnote 32 and (Gomes & Riello, 2021)). The condition to be satisfied by g_σ will be obtained from our definition of $\mathcal{F}_\sigma : \Phi \to C^\infty(M)$, with:

$$\mathcal{F}_\sigma(\mathbf{A}) = \nabla^a A_a = 0. \tag{4.11}$$

That is, we obtain:

$$\mathcal{F}_\sigma(h(\mathbf{A})) = \nabla^a(A_a + \nabla_a g_\sigma(\mathbf{A})) = 0, \tag{4.12}$$

and thus:

$$g_\sigma(\mathbf{A}) = -\nabla^{-2}\nabla^b A_b, \tag{4.13}$$

where ∇^{-2} is the inverse Laplacian (in the Lorentzian case, given sufficiently strict boundary conditions, it would be a Green's function, $\Box^{-1}$, with $\Box$ the D'Alembertian). Equation (4.13) clearly satisfies the covariance equation (3.8), namely:

$$g_\sigma(\mathbf{A}^g) = -\nabla^{-2}\nabla^b(A_b + \nabla_b g) = g_\sigma(\mathbf{A}) - g. \tag{4.14}$$

[32]The difference matters: the Lorenz choice in a Lorentzian manifold gives rise to a hyperbolic, and Coulomb always gives rise to an elliptic, partial differential equation. In the Lorentzian case with a Lorenz choice, the Maxwell equations of motion for A_a become the hyperbolic equation (which gives a well-posed initial value problem: see Gomes (2021e) for more details about the relation between the IVP and gauge freedom): $\Box A_a = j_a$, This equation still has the gauge-freedom corresponding to ϕ such that $\Box g = 0$. To fix g uniquely one requires strict boundary conditions, which could still carry the non-local behavior we have alluded to in Section 3.4 (cf. Gomes (2021e)). Thus, in the Lorentzian setting, this gauge-fixing is not complete and we would require additional imput about the initial state. In contrast, the kernel of the elliptic equation $\nabla^2 g = 0$ corresponding to Coulomb gauge can be defined in the absence of spatial boundaries: it is zero. So Coulomb, but not Lorenz, define a *bona-fide* gauge-fixing condition without the need of auxiliary conditions.

Since the inverse Laplacian is determined only up to a constant, so is g_σ, but this is a happy case in which the degeneracy in g_σ is a stabiliser of the gauge potential (and the stabilisers form a normal subgroup of $\mathcal{G}$; cf. Appendix A), and so has no effect on the dressed variable.

The corresponding dressed functional is:

$$h(\mathbf{A})_a := A_a^{g_\sigma(A)} = A_a - \nabla_a g_\sigma(A) = A_a - \nabla_a(\nabla^{-2}\nabla^b A_b) \tag{4.15}$$

Gauge-invariance of the dressed function is implied by (4.14) and can be immediately verified in (4.15). And of course, from (4.10), it is clear that if two dressed potentials match, they are dressings of gauge-related potentials. Thus two dressed variables match iff what is being dressed is related by an isomorphism.

4.2.b Maxwell Klein-Gordon: unitary gauge

This gauge is only available for some sectors of Abelian theories like electromagnetism with a nowhere vanishing charged scalar (Maxwell Klein-Gordon). As we will see, this choice of gauge is *sui generis* for being local (see Wallace (2024) for an in-depth analysis of the merits and shortcomings of this gauge).

The Maxwell-Klein-Gordon theory has models $\varphi = (A_\mu, \psi)$, with A_μ the gauge potential and ψ a complex scalar function. The gauge transformation of the models is defined as:

$$\varphi^g = (A_\mu + \partial_\mu g, \psi e^{ig}), \tag{4.16}$$

with $g \in C^\infty(M)$, and M representing spacetime.

In the unitary gauge, we have the condition

$$\mathcal{F}_\sigma(\varphi) = |\psi| - \psi = 0. \tag{4.17}$$

The dressed variables are:

$$h(A, \psi) = (A_\mu + \partial_\mu g_\sigma(\psi), \psi e^{ig_\sigma(\psi)}), \tag{4.18}$$

where we solve $F(h(\varphi)) = 0$ for $g_\sigma : \Phi \to C^\infty(M)$ and arbitrary φ. Assuming $|\psi|(x) \neq 0, \forall x \in M$, we write:

$$\psi = |\psi|e^{i\theta}, \text{ with } \theta = -i \ln \frac{\psi}{|\psi|}.^{33} \tag{4.19}$$

Finally, (4.17) and (4.18) imply that:

$$|\psi|e^{i(\theta + g_\sigma(\psi))} = |\psi|, \tag{4.20}$$

so we find

$$g_\sigma(\psi) = -\theta, \tag{4.21}$$

[33] There is an issue here that the logarithm of complex functions is not well-defined over the entire $\mathbb{C} - \{0\}$, because of the periodicity of solutions. Thus we should choose a subset of $\mathbb{C} - \{0\}$ that contains a single branch; but nothing depends on the branch.

which clearly has the right gauge-covariance properties. Thus,

$$h(A, \psi) = h(A, |\psi|) = (A_\mu - \partial_\mu\theta, |\psi|), \tag{4.22}$$

which is gauge-invariant. But subtracting $\partial_\mu\theta$ from A_μ imposes no constraint on the values that $A_\mu - \partial_\mu\theta$ can take. Thus we can redefine $A_\mu - \partial_\mu\theta \to \tilde{A}_\mu$ as an unconstrained, invariant 1-form. We thus rewrite the dressed fields as $(\tilde{A}_\mu, \rho)$, where $\rho \in C^\infty_+(M)$ (it is an everywhere positive smooth scalar function) and $\tilde{A}_\mu \in C^\infty(T^*M)$. So this is an explicitly local representation of the isomorphism-class; in particular, the consistency of piecewise-invariant descriptions as in (3.14), becomes trivial to assess.

A natural extension of this unitary gauge to the non-Abelian case takes a nowhere vanishing matter field to define an internal direction in the vector spaces where the fields take their values.[34]

In closing, this list of gauge-fixings for Yang-Mills theory is of course not exhaustive. There are many other, physically motivated choices we can make. For instance, we may also want to highlight the helicity degrees of freedom of the theory, in which case we would use temporal gauge. But temporal gauge, like the Coulomb one, depends on a foliation of spacetime by spacelike surfaces, and so, strictly speaking, cannot be given a univocal physical interpretation unless coupled to a representational scheme about spacetime foliations. We turn to this in Section 4.3. Before getting there, I will now exhibit a choice of dressed variables, that, like the choice of inter-particle distances for non-relativistic mechanics described in Section 4.1.b, is *not* obtained by a gauge-fixing section.

4.2.c Holonomies

A holonomy basis for Abelian gauge theories associates to each loop in spacetime a gauge-invariant phase, namely:

$$hol_\gamma(A) = \exp i \int_\gamma A, \quad \text{for each} \quad \gamma : S^1 \to M. \tag{4.23}$$

More generally, the holonomy thus defined is Lie-group-valued, and, in the non-Abelian case, it is not entirely gauge-invariant, but transforms covariantly in the adjoint representation. Nonetheless, there is a straightforward way to regain gauge-invariant variables by taking the trace of the holonomy: these are called *Wilson-loops*.[35]

[34] In the context of covariant perturbative gauge-fixings, this was labeled 'the Higgs relational connection' in (Gomes et al., 2019, Sec. 7). In the case of SU(2), a closely related decomposition of an arbitrary field using an everywhere non-zero internal direction is known as the The Cho-Duan-Ge (CDG) decomposition (Duan & Ge, 1979). It was rediscovered at about the turn of the century by several groups who were readdressing the stability of the chromomonopole condensate: the ensuing decomposition of the gauge potential helps in identifying the topological structures like monopoles and vortices in the gauge field configuration (see (Walker & Duplij, 2015) for a more recent concise summary of the literature and an application to QCD).

[35] Closed loops suffice in the vacuum case, but charges can be accomodated by integrating a gauge potential along a segment with charges at both its ends.

The Wilson-loop basis for gauge-invariant quantities is a representational scheme in the sense of Definition 1, but not in the sense of 3: although they 'dress' the original variables, they do not take values in the original space of models (they assign numbers to loops in spacetime). Thus the consistency of the composition of piecewise-invariant descriptions (discussed in Section 2.1 in the case of denumerable lists of invariants), cannot be settled by a simple criterion such as (3.13), which applies only for representational schemes via gauge-fixing.

In more detail: unlike dressed quantities obtained from gauge-fixings, the gauge-invariant variables derived from holonomies is vastly overcomplete. Thus the composition of these variables must obey certain constraints (cf. footnote 9). And whereas the consistency of the composition of piecewise-invariant descriptions via dressed variables obtained via gauge-fixing is settled by numerical coincidence of solutions of partial differential equations (cf. (3.13)), the constraints to be satisfied by the composition of holonomies are known as *Mandelstam identities*, and they can be rather complicated in the case of non-Abelian gauge theories (I briefly discuss this in Appendix B). Thus, in order to assess the consistency of piecewise invariant descriptions via holonomies we need to know whether all the values of two distinct sets of Wilson-loops, beloging to two regions that partially overlap, can be obtained from the holonomies of a single connection.[36] And to assess that we need to be able to solve these constraints. This is, again, if not intractable, much harder than solving the differential equations in the gauge-fixing case.

Here we see complications for describing gluing or composition for representational schemes according to Definition 1 (but not satisfying Definition 3).

4.3 General relativity

In the case of general relativity in the Lagrangian formalism, I will start with De Donder gauge, which fixes coordinates so that the densitized metric is divergence-free:

$$\mathcal{F}_\sigma(g) = \partial_\mu(g^{\mu\nu}\sqrt{g}) = 0. \tag{4.24}$$

I will not bore the reader with yet another demonstration that this gauge gives rise to dressed gravitational fields with the expected properties, described in Section 3.3.[37]

[36] Going in the other direction, if one assumes that all Wilson loops already satisfy the Mandelstam constraints, it *is* possible to reconstruct the corresponding isomorphism-classes, cf. (Barrett, 1991).

[37] As in footnote 44, we are here also in effect assuming the active-passive correspondence of (Gomes, 2021d, Sec. 5), to say that fixing a choice of coordinates is equivalent to picking out a unique model within an isomorphism class. In fact, as in the case of Lorenz gauge in gauge theory (cf. footnote 32), De Donder or harmonic gauge does not completely fix coordinate freedom in general relativity. Given the coordinates in an initial surface, the gauge uniquely defines coordinates to the past and future of that surface. Moreover, the use of coordinates implies that in general such conditions are only local. Giving a global, or geometric version of deDonder gauge can accomplished using an auxiliary metric, a measure of distance between metrics, and a diffeomorphism that is used to compare them (see (Landsman, 2021, Sec. 7.6)). This strategy—of using auxiliary metrics—is often employed in converting results originally expressed in particular coordinate systems into explicitly covariant results.

One can also see De Donder gauge as anchoring coordinate systems on waves of massless scalars: four different solutions of the relativistic wave equation. That is, given coordinates in an initial hypersurface Σ (or a portion thereof), we *define* coordinates in some region of spacetime using four non-local scalar functionals of the metric, that solve four wave-equations:

$$\Box \mathcal{R}^\mu_g(p) = 0; \tag{4.25}$$

the variables depend functionally on the metric (possibly non-locally), but take values at each point of the manifold. In a less coordinate-centric language, the idea here is to define point $p_x(g_{ab})$ as the point in which a given list of scalars $\mathcal{R}^\mu_g(p)$, $\mu = 1, \cdots, 4$ takes a specific list of values, $(x_1, \cdots, x_4)$'). In other words, fixing the metric and initial values, the four scalar quantities, $\mathcal{R}^\mu_g$ define a map:

$$\mathcal{R}_g = (\mathcal{R}^1_g, \cdots \mathcal{R}^4_g) : M \to \mathbb{R}^4. \tag{4.26}$$

To pick out points $p \in M$ by the value of the quadruple we invert the map (4.26). Assuming that the map (4.26) is a diffeomorphism—in general it is only locally one—there is a unique value, for all of the models, of $g_{ab}(\mathcal{R}^{-1}_g)(x)$, for any $x \in \mathbb{R}^4$.[38] That is, applying the chain rule for the transformations of $\mathcal{R}_g$,

$$\forall f \in Diff(M), \quad g_{ab} \circ \mathcal{R}^{-1}_g = f^* g_{ab} \circ \mathcal{R}^{-1}_{f^*g}. \tag{4.27}$$

So, using this quadruple, we have a unique representation of the metric on $\mathbb{R}^4$. Given some metric tensor $g^{\kappa\gamma}$ in coordinates x^κ, we can compute the metric in the new, $\mathcal{R}^\mu$ coordinate system as:

$$h^{\mu\nu} = \frac{\partial \mathcal{R}^\mu_g}{\partial x^\kappa} \frac{\partial \mathcal{R}^\nu_g}{\partial x^\gamma} g^{\kappa\gamma}. \tag{4.28}$$

This is just a family of 10 scalar functions indexed by μ and ν: the left-hand-side is the dressed variable, $h^{\mu\nu}_\sigma$, and the $\mathcal{R}^\mu$ are the dressing functions.

Of course, equations (4.26) and (4.27) would hold for any viable gauge-fixing: those choices also describe a local physical coordinate system, or a relational reference frame, on a patch of spacetime. And indeed, any such relation specifying points in terms of their 'qualitative properties' will be seen to explicitly furnish a specific counterpart relation across isomorphic and non-isomorphic models, thus resolving Limitation (B); cf. (Gomes & Butterfield, 2021).[39]

And there are many such examples. One that is widely used to study black hole mergers and initial value problems in general, is (the partial) 'CMC gauge' (cf. e.g. respectively (Pretorius, 2005; York, 1971)): this choice picks out clocks and simultaneity surfaces

[38] More commonly, for each g_{ab}, there will be only a subset $U \subset M$ that is mapped diffeomorphically to $\mathbb{R}^4$.

[39] Curiel (2018, p. 468) construes qualitative identity of points similarly:

> Once one has the identification of spacetime points with equivalence classes of values of scalar fields, one can as easily say that the points are the objects with primitive ontological significance, and the physical systems are defined by the values of fields at those points, those values being attributes of their associated points only per accidens.

But he does not consider the element of choice of the schemes, as we do.

so that simultaneous observers measure the same local expansion of the universe. In this case, $\mathcal{F}_\sigma(\gamma_{ij}, \pi^{ij}) = \gamma^{ij}\pi_{ij}$ =const, where γ_{ij} is the spatial metric and π^{ij} is its conjugate momentum, obtained in the Hamiltonian (3+1) formulation of general relativity (Arnowitt et al., 1962).

Another example that does not require any fields other than the gravitational ones, is known as *Komar-Kretschmann* variables. For this example, we must first restrict our attention to spacetimes that are not homogeneous, i.e. generic spacetimes (i.e. excluding Pirani's type II and III spaces of pure radiation, in addition to excluding symmetric type I spacetimes). Once this is done, we consider $\mathcal{R}^\mu_g(p))$, $\mu = 1, \cdots 4$, formed by certain real scalar functions of the Riemman tensor. Since the spacetimes considered here are suitably inhomogeneous, they all contain points in which these functions are linearly independent, and so can be used to specify location without limiting the physical content of the spacetime region *as a whole*.[40] This choice is exceptional for not requiring auxiliary structure: no fields other than the metric, nor fixed spacetime curves or points; the relational reference frame that it provides supervenes only on properties of the metric tensor.[41]

But given particular spacetime points or curves—taken to have, in Kripke's famous terminology, a *rigid designation*, across a suitably restricted set of models[42]—we can find adapted coordinate systems or reference frames. Perhaps the most famous choice of coordinates for general relativity, called 'Riemann-normal' coordinates, is obtained in this way: having rigidly fixed a tangent frame at an event across a certain set of isomorphism-classes (or physical possibilities), the local coordinates are obtained by applying the Riemann exponential map for the tangent space of that event. This choice is particularly useful for a set of possibilities contaning material systems whose scale is small compared to the curvature scale: in these coordinates the metric is described as almost flat along the trajectory of the system; the Riemann curvature appears only at second order in the proper distance to the freely-falling trajectory of the material system.

Another common example, indeed the most applied for navigation on Earth, is a coordinate choice based on GPS sattelites. This requires four sattelites to be rigidly designated (taken to be small, and non-back-reacting, and thus following timelike geodesics), which cross at an initial event and emit timed light-like signals. Those signals cover a spacetime region, and their values provide it with a coordinate system (cf. (Rovelli, 2002) for details). [43]

[40]Komar (1958) finds these real scalars through an eigenvalue problem:

$$(R_{abcd} - \lambda(g_{ac}g_{bd} + g_{ad}g_{bc}))V^{cd} = 0,$$

where V^{cd} is an anti-symmetric tensor. The requirement ensures solutions λ, whose existence we assume, are independent real scalar functions. Komar (1958, p.1183) takes these scalars to be preferred, "since they are the only nontrivial scalars which are of least possible order in derivatives of the metric, thus making them the simplest and most natural choice."

[41]See (Bamonti, 2023) for a classification of different types of reference frames in general relativity, according to their coupling to the metric and to the inclusion of back-reaction.

[42]A term is said to be *a rigid designator* when it designates (picks out, denotes, refers to) the same thing in all possible worlds in which that thing exists.

[43]Similarly to GPS coordinates, the 'dressed' diffeomorphism-invariant observables of Donnelly & Giddings

5 How to choose a representational scheme?

In the previous Section, I exhibited several possible choices of representational schemes via gauge-fixing. How should we choose amongst them? Before answering the question of this Section, I will motivate it with an argument from Healey (2007). More specifically, I take issue with one argument in Healey's (otherwise outstanding!) book on the philosophical interpretation of gauge theories. It occurs in his Chapter on classical gauge theories (Chapter 4). Here, by answering Healey, I will (superficially) connect the ideas to a popular topic in analytic philosophy: functionalism.

I will start in Section 5.1 by recapitulating Healey's argument. Then in Section 5.2 I will show that, *contra* Healey, intra-theoretic resources enable us to pick out representational schemes for gauge theory.

5.1 Healey's argument from functional roles

(Healey, 2007, Section 4.2)'s describes what he takes to be a fundamental difference between the tenability of gauge-fixing in general relativity and in Yang-Mills theory. That is, a difference between specifying, amongst the infinitely many physically equivalent representatives, a particular spacetime distribution of the gauge potentials or of the metric. As I understand Healey, he posits that this specification is easy for the metric, but impossible for the gauge potential. It is easy for the metric because Lewis's ideas about functionalism apply to it, but, allegedly, they don't apply to gauge potentials. And with this alleged contrast I will disagree. My disagreement will shed light on how representational schemes are chosen.

To spell out Healey's argument in more detail, I will now indulge in a bit of 'Healey exegesis'. Healey admits that within a theory there may be many terms standing for unobservable items, such as those variables that are not gauge-invariant. But unobservability by itself is not bad news, since Lewis (1970, 1972)'s construction, which I will summarize shortly below, applies to the observable-unobservable distinction. The construction allows us to specify the meaning of unobservable items via their relations to observable ones and to each other. In more detail, the idea is that we can simultaneously specify what several theoretical items—labeleled T-terms—refer to, by their each uniquely satisfying some description (usually called "functional role") that can be formulated in terms of each other and of the better understood—perhaps even observable—terms, labeled O-terms.

But how could we use a theory to functionally single out any item if that theory treats that item and others, at least in certain respects, as being on a par, i.e. related by some symmetry? As described by Lewis (2009), the response is to appeal to patterns of facts of "geography" to break the underdetermination:

> Should we worry about symmetries, for instance the symmetry between positive and negative charge? No: even if positive and negative charge were exactly alike in their nomological roles, it would still be true that negative charge is found in

(2016) are anchored on non-null spacetime curves.

> the outlying parts of atoms hereabouts, and positive charge is found in the central parts. O-language has the resources to say so, and we may assume that the postulate mentions whatever it takes *to break* such symmetries. Thus the theoretical roles of positive and negative charge are not purely nomological roles; *they are locational roles as well.* [my italic]

But Healey argues that, in gauge theories, even this Lewisian strategy is bound to be plagued by under-determination. I interpret Healey as saying that one can functionally specify a representational scheme for a spacetime metric using only O-terms, but cannot specify a representational scheme for a gauge potential. More precisely, here is Healey's argument that the functionalist methodology applies so as to single out spacetime metrics, but not to single out gauge potentials:

> The idea seems to be to secure unique realization of the terms in face of the assumed symmetry of the fundamental theory in which they figure by adding one or more sentences stating what might be thought of as "initial conditions" to the laws of that theory [...] to break the symmetry of how these terms figure in T. They would do this *by applying further constraints* Those constraints would then fix the actual denotation of the [...symmetry-related terms...] in T so that, subject to these further constraints, T is uniquely realized. [...But] *The gauge symmetry of the theory would prevent us from being able to say or otherwise specify which among an infinity of distinct distributions so represented or described is realized in that situation. This is of course, not the case for general relativity.* (Healey, 2007, p. 93) [my italics]

But I would ask: why does Healey see here a contrast between general relativity and gauge theory? I do not see in this entire passage an attempt to draw a distinction regarding Leibniz-equivalence and TS about symmetries between the two cases. I think the interesting question being alluded to here is whether we can use features of the world around us to single out a unique model of the theory using unique features of that model.

This question is interesting because, in practice, we *do* select some model over others when we represent a given physical situation, and therefore we must somehow 'break the symmetry' between all of the models. In other words, if we accept that the basic ontology is symmetry-invariant, we are faced with a mystery: how do we select particular representatives of the ontology—i.e. models—over others? Under what conditions could we be justified in choosing for the gauge potential one spacetime distribution over another, isomorphic one? Selecting such a representative involves a tension between: (a) taking physical properties as invariant under the symmetries in question, and (b) in practice selecting unique representative distributions, among the infinitely many isomorphic representatives of the same situation. At first sight, these two requirements, (a) and (b), are inimical, if not contradictory, for (a) implies we can have no physical guidance for accomplishing (b)!

Getting representational schemes off the ground of course depends on a positive resolution of this paradox. For though no particular choice is mandatory, each must be based on something: physical features, indexicals, ostension, etc.

Thus I take the more interesting interpretation of Healey's passage here to be that this 'singling out' of particular models is possible for gravity, but not for gauge theory, where the choice is unthethered from any feature of the physical world: it is entirely arbitrary, according to him. Answering this question is thus crucial in order to clarify the scope of representational schemes.

5.2 Refuting Healey's alleged distinction using representational schemes

Contra Healey, I say that having some physical "hook" with which to choose representatives does *not* imply that we are breaking the symmetry at a fundamental level. Different choices of representational schemes are equally capable of representing a given state of affairs; some may just be more cumbersome than others. Or perhaps some representations obscure matters for the purposes at hand, even while they may shine a light on complementary aspects of that state of affairs. And, in this sense, by choosing different schemes we shift our focus to different features of the world, according to our interest.

So we construct a particular representative of the gauge potential as fulfilling a given role. We dissolve the tension between (a) and (b) described above, with explicit examples; by, in Healey's words: '*breaking* the symmetries', by providing '*further constraints*'.

Note to begin with that, according to Healey's standards, we are justified in including in our O-vocabulary all the 'locational roles', which describe contingent, happenstantial facts about 'where and when' specified events happen; and which I will loosely interpret as 'referring to spacetime'. Thus I free myself to include in Lewis' O-vocabulary, and thereby use in the specification of the roles, the differential geometry of spacetime. In short, I will assume reference to spacetime is 'old' or already understood.

In the simple example of electromagnetism, we require the model to satisfy certain relations among the parts of the field. For example, in our hierarchy of extra-empirical theoretical virtues, we could place Lorentz covariance very highly (cf. Mulder (2021) and Mattingly (2006) for advocacy of this criterion and choice of gauge) and therefore prefer an explicitly Lorentz-covariant choice of scheme. The only extra constraints that we have imposed in (**??**), namely, that the spacetime divergence of the particular representative of the gauge potential vanishes, use only the O-vocabulary that Healey would grant us.[44]

Of course, no empirical fact can pick out a unique choice of σ. The representational scheme cannot be empirically *mandated*, since, by assumption, the symmetries leave all empirical matters invariant, in both the gravitational and the gauge cases. But a different choice would not satisfy the original condition that uniquely specified σ.

Each choice is at most suggested by being *suitable* for certain types of questions one might want to ask about the system. Different choices merely represent different lenses through which we capture the invariant structure of the states.

[44]There is a subtle point here: we are in effect assuming the active-passive correspondence of (Gomes, 2021d, Sec. 5), to say that fixing a choice of section of the bundle is equivalent to picking out a unique model within an isomorphism class.

We should not be disappointed by this lack of uniqueness. It should be seen instead as a *flexibility* that is also explanatory. Just as it is easy to explain the Larmor effect by a Lorentz boost between different frames, we use different gauges to explain that a given process in quantum electrodynamics involves just two physical polarisation states or that it is Lorentz invariant. In both the special relativistic and the gauge scenarios, two different 'frames'—for quantum field theory, the temporal and Lorenz gauge—are necessary to more easily explain two different aspects of a given phenomenon. In the words of Tong (2018, p. 1):

> The [gauge] redundancy allows us to make manifest the properties of quantum field theories, such as unitarity, locality, and Lorentz invariance, that we feel are vital for any fundamental theory of physics but which teeter on the verge of incompatibility. If we try to remove the redundancy by fixing some specific gauge, some of these properties will be brought into focus, while others will retreat into murk. By retaining the redundancy, we can flit between descriptions as is our want, keeping whichever property we most cherish in clear sight.

In sum, the representational scheme can be seen as specifying a functional role, using different pragmatic, explanatory, and theoretical requirements. The choice of gauge that it implies finds 'sufficient reason' in the satisfaction of this role; there is no empirical breaking of gauge-symmetry. This construction thus explicitly contradicts the letter of Healey's underdetermination argument. and concludes my response to what I called Healey's 'interesting challenge': showing that intra-theoretic resources enable us to choose representational schemes.

6 Representational schemes overcoming Limitation (B)

I take Limitation (A)—the issue described in Section 2: that TS left open the matter of providing sufficient and tractable conditions of identity for the isomorphism-classes—to have been essentially overcome in Section 3, exemplified in Section 4, and clarified in Section 5.

But I still have not explicitly addressed how representational schemes also overcome Limitation (B). This is the aim of this Section.

To recap, the second Limitation is that (cf. Section 2.1.b):

B. TS denies that objects have *primitive* identity across physical possibilities, or, equivalently, across different isomorphism-classes. But such identities were useful in order to express counterfactuals. So how are counterfactuals expressed using only isomorphism-invariant properties?

For concreteness, and since most of the literature focuses on Limitation (B) in the case of spacetime, I will here also focus in that application; but most conclusions will also apply to the gauge case (by replacing 'points' by 'internal frames' of a vector space).

More specifically, I will address Limitation(B) through another Lewisian topic: countepart theory. Counterpart theory is the philosophical doctrine, due to David Lewis (Lewis (1968), (Lewis, 1973, p. 38-43) and (Lewis, 1986, Ch. 4)) that any two objects—in particular, spacetime points—in two different possibilities (in philosophical jargon: possible worlds) are never strictly identical.[45] They are distinct, though of course similar to each other in various, perhaps many, respects. Representational schemes, specified by extra-empirical criteria as in Section 5, can be construed as picking out which respects are important for the issue at hand.[46]

As to counterfactuals, what makes true a proposition that the object a of the actual world could have had property F (though in fact it lacks F) is not that in another possible world, a *itself* is F, but that in another possible world, which is sufficiently similar to the actual world, an object appropriately similar in certain respects to a, is F. That object is called the *counterpart*, at this other possible world, of a.

Summing up the Limitation and the strategy to overcome it: though denying trans-world identity fits TS's denial of any sort of primitive identity, counterparthood furnishes a relation that TS lacks, allowing us to discuss counterfactuals. Each representational scheme via gauge-fixing provides explicit and invariant relations between non-isomorphic models, that are sufficiently flexible to define local counterparts. It does this most clearly in its guise as a relational reference frame. That is because, as I said in Section 3.3.c, relational reference frames make precise a frequent claim found in the literature on TS about spacetime: that 'spacetime points can only be specified by their web of relations to other points'. Each gauge fixing fixes a web of relations that is maximally rigid with respect to locations but also maximally loose with respect to the physical content of those locations. Thus we can specify spacetime points in different possibilities—in non-isomorphic models—by their location within such a web of relations and compare the values of these physical quantities that are not fixed by that web.[47]

In Section 6.1 I provide the basic formalism for conceiving of counterpart relations as given by representational schemes and discuss how changing representational schemes affects the counterpart relations; in Section 6.2 I describe obstructions to construing counterpart relations in this manner—obstructions due to homogeneous models.

6.1 Basic formalism

Mathematically, I will construe a counterpart relations very narrowly: as group elements $g \in \mathcal{G}$ that relate two different models φ_1 and φ_2, not necessarily isomorphic. For a defence of this narrow mathematical interpretation of counterparthood, see Gomes & Butterfield (2023a).

Given two models for vacuum general relativity, $\langle M, g^1_{ab}\rangle$ and $\langle M, g^2_{ab}\rangle$, a diffeomorphism $f \in \text{Diff}(M)$ will give us a counterpart relation between the spacetime points of M in each of

[45]For Lewis, 'possibilities' and 'possible worlds' are not exactly the same thing: possibilities require the stipulation of a world and a choice of counterpart relation.

[46]This idea of using gauge-fixings as counterpart relations was introduced in (Gomes, 2022) and developed in fullness in Gomes & Butterfield (2023a); here I will give only a sketch.

[47]In this Section, since we will be comparing objects across different possibilities, and using different comparisons, I prefer to talk about *specifying* a point or region, rathern than individuating it, since the latter term's connotation of identity is stronger than is required here, and could lead to regions bearing their content essentially, and not leading to an interesting notion of counteparthood; cf. Section 2.3.

the two models.[48]

Given a section $\mathcal{F}_\sigma$, as described in Section 3.3, we have just such an element: the dressing $g_\sigma : \Phi \to \mathcal{G}$, given in Equation (3.6). Using this convenient mathematical operator, the counterpart relation between φ_1 and φ_2 is given by:

$$\text{Counter}_\sigma(\varphi_1, \varphi_2) := g_\sigma(\varphi) g_\sigma(\varphi_2)^{-1}. \tag{6.1}$$

Generally, i.e. for models that are not along the section, the relation is given by the group element that takes the first model down to the gauge-fixing section and then back up towards the second model. Importantly, by being associated to a gauge-fixing, we don't limit the physical content that we seek to compare via counterparthood. Thus each section gives a 'qualitative', i.e. diffeomorphism-invariant, counterpart relation between the dressed spacetime points of distinct physical possibilities.

Since g_σ is generally a non-local functional of the model, the counterpart relation may also depend non-locally on the models. But the counterpart relation between any two states that already lie at the gauge-fixing section is just the identity: they are both already in their preferred representation relative to the reference frame that is associated to the gauge-fixing, so these counterpart relations are trivially local. It is the reference frame associated to the gauge-fixing that defines locality: though forming the reference frame using properties of the fields may involve non-local operations, once we have it, it fixes physically relevant, isomorphism-invariant, locations.

Such a notion of spacetime counterparts have interesting properties. Due to covariance (see (3.8)),

$$\text{Counter}_\sigma(\varphi_1^g, \varphi_2^g) = g^{-1} g_\sigma(\varphi_1)(g^{-1} g_\sigma(\varphi_2)^{-1} = g^{-1}\,\text{Counter}_\sigma(\varphi_1, \varphi_2) g\ . \tag{6.2}$$

Moreover, two models that lie in the same orbit will always be related by the unique isomorphism that connects them. That is: if $\varphi_2 = \varphi_1^g$, then $\text{Counter}_\sigma(\varphi_1, \varphi_2) = g$, even if neither φ_1 nor φ_2 lie in the section σ. For with the definitions above,

$$\text{Counter}_\sigma(\varphi_1, \varphi_1^g) = g_\sigma(\varphi_1)(g^{-1} g_\sigma(\varphi_1))^{-1} = g \tag{6.3}$$

Thus, for example, in the spacetime case, if the unique counterpart of p in model $\langle M, g_{ab}\rangle$ is q in model $\langle M, g'_{ab}\rangle$, then the counterpart of p in model $\langle M, f^* g'_{ab}\rangle$ will be $f(q)$. And this particular property is independent of which scheme σ we choose.

In other words, though there are several distinct choices of counterparts, each choice must identify the same spacetime points across isomorphic models. In the spacetime context, this is essentially Field (1984, p. 77)'s observation that

> "individuation of objects across possible worlds" is sufficiently tied to their qualitative characteristics so that if there is a unique 1-1 correspondence between the

[48] Similarly, given a vector bundle, E, the bundle of admissible frames over E, $P \simeq L_G(E)$, and two sections of E, $\varphi, \varphi' \in \Gamma(E)$, the counterpart relations tells which frame over x for φ corresponds to which frame over x for φ'. It thus can be seen as a vertical automorphism between two models of Yang-Mills theory: $\langle P, \omega_1\rangle$ and $\langle P, \omega_2\rangle$.

space-time of world A and the space-time of world B that preserves all geometric properties and relations (including geometric relations among the regions, and occupancy properties like being occupied by a round red object), then it makes no sense to suppose that identification of space-time regions across these worlds goes via anything other than this isomorphism.

Indeed, this criterion, that we can here call 'the drag-along' interpretation of isomorphic models (cf. Butterfield & Gomes (2022)), was the centerpiece of Weatherall (2018): the paper that sparked a recent surge of interest in the hole argument in general relativity.[49] And here we see that this criterion is satisfied by our notion of qualitative counterpart relations using representational schemes. But we do not always have universal, unique, counterpart relations, and so Field's criterion does not always apply. We now turn to this.

6.2 A snag for counterparts but not for schemes

Given any representational scheme via gauge-fixing, in accord with Field's criterion, above, as long as the web of properties and relations in each of our models is sufficiently complex, it suffices to uniquely specify points across models.

But if this web is not sufficiently complex, the specification will fail. To be blunt: we cannot stipulate a scheme for some region in spacetime or frame in a vector space if there aren't enough specific physical features to single out that region or frame.

As described in Appendix A (see also end of Section 3.2), models that are 'too homogeneous', or symmetric, have stabilizers, which represent a degeneracy in the web of relations that they can furnish. That is, on homogeneous states, $\tilde{\varphi}$, a representational scheme—seen as a map $\sigma : [\Phi] \to \Phi$—will fail to uniquely specify the g_σ of (3.6). The reason is that g_σ fails to satisfy the property of *Uniqueness* of Section 3.3. There are particular $\tilde{g} \in \mathcal{G}$ such that $\tilde{\varphi}^{\tilde{g}} = \tilde{\varphi}$, and so, in particular, we cannot fulfill the covariance Equation (3.8). In the words of Gomes et al. (2019, Sec. 7):[50]

[49] Bradley & Weatherall (2022); Weatherall (2018) associate the 'the drag-along' idea less with Field and more with Mundy (1992), who distinguishes a theory's synthetic language—that is only able to expresses qualitative, non-singular facts—from a theory's metalanguage, in which we are able to talk about points singularly, i.e. without a definite description (see (Gomes, 2021c, Sec. 3.A)). But the gist is the same: isomorphisms—expressed in the metalanguage in Mundy's case—will map objects singled out by the same description into each other. Here is Mundy (1992, p. 520):

> Thus an element p of M satisfies a description D of L_R [i.e. a qualitative description] in S iff the corresponding element $f(p)$ satisfies D in Sf: p and fp occupy the same structural roles in these two isomorphic models, so every statement true of p in S is true of $f(p)$ in Sf. Therefore, since a theory identifies elements of its domain only by descriptions expressed in its language, two such models describe the same theoretical world.

[50] (Gomes et al., 2019)'s discussion is in the context of the 'relational connection-form', to be briefly described in Section 7.2. But in Section 7 (ibid), they are in particular describing a case in which that connection is integrable, and thus recovers a gauge-fixing section.

> The fact that [the reference frame] is left undetermined in the directions belonging to the stabilizer should have been expected from the simple fact that a field insensitive to some transformations cannot be used as a reference to [discern those transformations].

More broadly, this type of obstruction due to homogeneity is well known, and related to Black (1952)'s criticism of Leibniz's Principle of the Identity of Indiscernibles—the principle that motivates Leibniz equivalence (see (Pooley, 2022, Sec. 3.3) for a thorough exposition).

One might have thought that, precisely because the degeneracy in g_σ acts trivially on the respective models, it does not yield any degeneracy about which models belong to the gauge-fixing surface. But by being more mathematicaly rigorous about counterpart relations, we uncover a snag which has not been noticed in the philosophical literature. In the non-Abelian case, there may be no well-defined quotient of the entire gauge group by the stabiliser subgroup. This will occur if the stabiliser subgroup is not a normal subgroup of the gauge group. And so, in these cases, it makes no sense to define a counterpart relation 'up to stabilisers'.[51]

In sum, for very homogeneous states, g_σ is not well-defined, so the counterpart relation (6.1) is no longer valid and will not recover Field's criterion. Nonetheless, the interpretation of isomorphic models through TS survives unscathed and, when the stabiliser group is a normal subgroup of $\mathcal{G}$, this interpretation is preserved by representational schemes, which are still well-defined. The investigation of counterpart relations for non-normal stabiliser groups will be left for future work.

7 Summing up

In this Section, I conclude, by first, in Section 7.1, offering a very brief summary. Then, in Section 7.2, I list two idealizations involved in representational schemes, which I have so far skatted over.

7.1 Representational schemes conceptually summarised

In both the gravitational and the gauge cases, there are many ways in which we can represent structurally identical patterns; each of these ways correspond to one of many *isomorphic models* of the theory. But we can pick particular representations by resorting to a *representational scheme*.

Representational schemes in the context of these theories will introduce a notion of 'relationism' into the formalism. This is not necessarily the notion that is usually attributed to Mach, Poincaré, Barbour, etc. In their camp, this notion also requires relative values of fields or particle positions, but these relative values are subsets of all the relations, and need not correspond to the 'fundamental ontology'.

[51] But since stabilisers themselves form a group, one could still plausibly construct a counterpart relation for that group, valid for all configurations that share that stabiliser group (up to conjugacy).

Let us take general relativity as an example: I accept that the implied ontology is firmly about distances between spacetime points. In my view, there is not much more to say about that: the fundamental ontology is interpreted as per TS and anti-haecceitism. But even being clear about the fundamental ontology, or the type of structure in question, we have at first glance no concise way of qualitatively describing how these distances are distributed, since we have no concise way of qualitatively designating the points that stand in these relations. Here then is the role of relationism as embodied by representational schemes: each representational scheme is a choice of a set of relations used to describe physical content without redundancy, i.e. satisfying Definition 1. In the case of spacetime, one can think of each representational scheme as providing a particular qualitative description of these distributions of distances as described relative to some reference frame.

Thus one main difference between mine and the traditional notion of relationism is that I explicitly admit multiple choices of sets of relations, leading to many distinct possible reductions of the theory. One set can only be *pragmatically* preferred to others. Indeed, when Einstein famously overcame his doubts about the 'hole argument' by appealing to 'point coincidences' (cf. Giovanelli (2021)), he assumed there was no further question as to what they were coincidences of. But, as we have seen, different choices can yield very different physical descriptions of a given state of affairs: very different coincidences, which equally well coordinitize a given region of spacetime. So there is the matter of which scheme to choose. To pick a scheme we must employ intra-theoretic resources by the lights of which the scheme is appropriate.

In sum, representational schemes can be construed as (non-unique) choices of a relational set of quantities that:

(I) provide tractable, complete characterizations of the individual isomorphism-classes, i.e. of the structure that, given a fixed notion of isomorphism, is: common to each class of isomorphic models, and different for different classes of isomorphic models. They give a tractable criterion of individuation of isomorphism-classes.

(II) allow us to describe qualitative counterpart relations for points, regions, frames, etc, across isomorphic and non-isomorphic models, while entirely denying any primitive—i.e. scheme-independent—identification across models.

And so, to complement my first two Desiderata for TS of Gomes (2021c), I here introduce Desideratum (iii): that the theory's symmetries admit representational schemes. Fulfillment of this Desideratum overcomes the first two Limitations of Section 2.1.b, left over from Gomes (2021c). But is Desideratum (iii) a non-trivial requirement?

7.2 Limitations of representational schemes.

Representational schemes overcome some of the limitations of TS, but they are also limited in a way which we have only touched on in passing so far. Take De Donder variables as defined by four massless scalar fields satisfying a wave equation (cf. Equation (4.25)). If their domain is a generic spacetime, that spacetime will (generically) not be homeomorphic to $\mathbb{R}^4$. As we know, there are topological limitations to the existence of such homeomorphisms, and there might be other obstructions as well. The upshot is that representational schemes aren't usually global, either in spacetime or in Φ.

Recall from Section 2 that individuating isomorphism-classes is a very common goal for theorists working in quantum gravity, and in the (non-perturbative) quantization of Yang-Mills theory. And indeed, the global limitations of representational schemes are major obstacles to achieving this goal. For instance, in quantum gravity, the infamous "problem of time" is often portrayed as a problem of finding suitably general 'internal clocks' (cf. e.g. (Isham, 1992, Sec. 3.4)).

This is why adding the Desideratum to the theory that it admits a representational scheme is highly non-trivial. In fact, it is not perfectly satisfied even for non-Abelian Yang-Mills and general relativity, due to the 'Gribov obstruction' (see Appendix A). But it is satisfied in the Abelian case, and, hopefully, its general features should be present even with the caveats above.

There is an alternative route that overcomes some of the limitations of representational schemes. In the original treatment of Gomes et al. (2019); Gomes & Riello (2017, 2018, 2021), one avoids the problem of Gribov ambiguities by introducing a more flexible tool than representational schemes, called a 'relational connection-form'. In brief, this is a distribution of infinitesimal gauge-fixing surfaces, with appropriate covariance properties. These connection forms generalize parallel transport in principal fiber bundles to parallel transport in the space of models, where the gauge group forming the orbits is the entire group of gauge transformations. In this new context, connection-forms provide counterpart relations between members of any one-parameter sets of models—i.e. along 'histories' of models—and so generalise Barbour's 'best-matching' (see e.g. (Mercati, 2017, Ch. 4)). Heuristically, this can be understood as picking out, along two neighboring orbits, the two models that are the closest, according to some choice of gauge-invariant notion of 'closeness' that applies to entire models.

The transversal, covariant distribution defined by a relational connection form is not necessarily integrable into 'slices' (see Appendix A), which means it usually has curvature in Φ, and that it is inequivalent to any gauge-fixing. In this case, since they are history-dependent, counterpart relations given by connection-forms will not generally satisfy 'Field's criterion' (see (Gomes et al., 2019) for a pedagogical introduction).[52]

Acknowledgements

I would like to thank Jeremy Butterfield for many detailed comments on several versions of this paper, and Frank Cudek, for comments and pointing me to Curiel's paper on the problem of counterfactuals for general relativity. I would also like to thank the British Academy for financial support.

[52]This more flexible tool (which was technically developed in Gomes et al. (2019); Gomes & Riello (2017, 2018, 2021) has been conceptually appraised in Gomes (2019, 2021a) and (Gomes & Butterfield, 2023a, Appendix A).

APPENDIX

A Infinite-dimensional principal fiber bundles

In the case of field theories, such as general relativity and Yang-Mills, even admitting a free and proper action of an infinite-dimensional Lie group on an infinite-dimensional smooth manifold, it is not guaranteed that a local product structure exists everywhere: so Φ is not necessarily a *bona-fide* infinite-dimensional principal $\mathcal{G}$-bundle, i.e. $\mathcal{G} \hookrightarrow \Phi \to \Phi/\mathcal{G}$. The obstacle is that there are special states—called *reducible*—that have stabilizers, i.e. elements $\tilde{g} \in \mathcal{G}$ such that $\tilde{\varphi}^{\tilde{g}} = \tilde{\varphi}$. Moreover, the statement holds for the entire group orbit, as it is easy to show that for some $\tilde{\varphi}' := \tilde{\varphi}^g \in \mathcal{O}_{\tilde{\varphi}}$,

$$\tilde{\varphi}'^{g^{-1}\tilde{g}g} = \tilde{\varphi}^{\tilde{g}g} = \tilde{\varphi}^g = \tilde{\varphi}'$$

and so all the elements of the orbit are also reducible (with stabilizers related by the co-adjoint action of the group). And so orbits are of 'different sizes', and not isomorphic to the structure group. Nonetheless, there is a generalization of a section, called a slice, that provides a close cousin of the required product structure (see (Gomes & Butterfield, 2023a, Sec. 2.1, and footnotes 4,5)). As has been shown using different techniques and at different levels of mathematical rigour, Diez & Rudolph (2019); Ebin (1970); Isenberg & Marsden (1982); Kondracki & Rogulski (1983); Mitter & Viallet (1981); Palais (1961); Wilkins (1989) both the Yang-Mills configuration space and the configuration space of Riemannian metrics (called Riem(M)), admit slices. These slices endow the quotient space with a *stratified* structure. That is, the space of models can be organised into orbits of models that possess different numbers of stabilizers; with the orbits with more stabilizers being at the boundary of the orbits of models with fewer stabilizers. For each stratum, we can find a section and form a product structure as in the standard picture of the principal bundle.

For both general relativity and non-Abelian gauge theories, reducible configurations form a meagre set. *Meagre* sets are those that arise as countable unions of nowhere dense sets. In particular, a small perturbation will get you out of the set (and this is true of the reducible states in the model spaces of those theories, according to the standard field-space metric topology (the Inverse-Limit-Hilbert topology cf. e.g. Fischer & Marsden (1979); Kondracki & Rogulski (1983)). In this respect, Abelian theories, such as electromagnetism, are an exception: *all* their configurations are reducible, possessing the constant gauge transformation as a stabilizer.

And apart from this obstruction to the product structure—i.e. even if we were to restrict attention to the generic configurations in the case of non-Abelian field theories—one can have at most a *local* product structure: no representational scheme, or section, giving something like (3.3), is global (this is known as the *Gribov obstruction*; see Gribov (1978); Singer (1978)). Unfortunately, the space of Lorentzian metrics is not known to have such a structure: it has only been shown for the space of Einstein metrics that admit a constant-mean-curvature (CMC) foliation.

In the infinite-dimensional case, both the dimension and the co-dimension of a regular value surface can be infinite, and it becomes trickier to construct a section: roughly, one starts by endowing Φ with some $\mathcal{G}$-invariant (super)metric, and then finds the orthogonal complement to the orbits, $\mathcal{O}_\varphi$, with respect to this supermetric. But here the intersection of the orbit with its orthogonal complement cannot be assumed to vanish, as it does in the finite-dimensional case. Nonetheless, in the cases at hand, that intersection is given by the kernel of an elliptic operator, and one therefore can invoke the 'Fredholm Alternative' (see (Gilbarg & Trudinger, 2001, Sec. 5.3 and 5.9)) to show that that intersection is at most finite-dimensional, but generically is zero, and thus the generic orbit has the 'splitting' property: the total tangent space decomposes into a direct sum of the tangent space to the orbit and its orthogonal complement. Now we must extend the directions transverse to the orbit, so as to construct a small patch that intersects the neighboring orbits only once. But the space of Riemannian metrics is a cone inside a vector space, so it is not even affine and we cannot just linearly extend the directions normal to the orbit and hope for the best. And so we extend the normal directions by using the Riemann normal exponential map with respect to the supermetric (cf. Gil-Medrano & Michor (1991)), and thus conclude that, for a sufficiently small radius, the resulting submanifold is transverse to the neighboring orbits and has no caustics. Finally, to show that this 'section' is not only transverse to the orbits, but indeed that it intersects neighboring orbits only once, the orbits must be embedded manifolds, and not just local immersions: this is guaranteed if the group action is proper, cf. Ebin (1970).

B The holonomy representation

We can assign a complex number (matrix element in the non-Abelian case) $hol(C)$ to the oriented embedding of the unit interval: $C : [0,1] \mapsto M$, by integration of a phase:

$$hol_C(A) = \exp i \int_C A. \tag{B.1}$$

Under a gauge transformation, we obtain:

$$hol_C(A^g) = g^{-1}(C(0))hol_C(A)g(C(1)). \tag{B.2}$$

If the endpoint of C_1 coincides with the starting point of C_2, we define the composition $C_1 \circ C_2$ as, again, a map from $[0,1]$ into M, which takes $[0,1/2]$ to traverse C_1 and $[1/2,1]$ to traverse C_2. The inverse C^{-1} traces out the same curve with the opposite orientation, and therefore $C \circ C^{-1} = C(0)$. Following this composition law, it is easy to see from (B.1) that

$$hol(C_1 \circ C_2) = hol(C_1)hol(C_2), \tag{B.3}$$

with the right hand side understood as complex multiplication in the Abelian case, and as composition of linear transformations, or multiplication of matrices, in the non-Abelian case. For both Abelian and non-Abelian groups, given the above notion of composition, holonomies are conceived of as smooth homomorphisms from the space of loops into a suitable Lie group.

One obtains a representation of these abstractly defined holonomies as holonomies of a connection on a principal fiber bundle with that Lie group as structure group; the collection of such holonomies carries the same amount of information as the gauge-field A (cf. (Belot, 1998, Sec. 3) for a philosophical exposition). However, only for an Abelian theory can we cash this relation out in terms of gauge-invariant functionals. That is, while (B.1) is gauge-invariant, the non-Abelian counterpart (with a path-ordered exponential), is not. For non-Abelian theories the gauge-invariant counterparts of (B.1) are Wilson loops, see e.g. (Barrett, 1991), $W(\gamma) := \mathrm{Tr}\,\mathcal{P}\exp\left(i\int_\gamma A\right)$, where one must take the trace of the (path-ordered) exponential of the gauge-potential. It is true that all the gauge-invariant content of the theory can be reconstructed from Wilson loops; (see also Rosenstock & Weatherall (2016), for a category-theory based derivation of this equivalence). But, importantly for our purposes, it is no longer true that there is a direct homomorphism from the composition of loops to the composition of Wilson loops. That is, it is no longer true that the counterpart (B.3) holds: $W(\gamma_1\circ\gamma_2) \neq W(\gamma_1)W(\gamma_2)$. The general composition constraints—named after Mandelstam—come from generalizations of the Jacobi identity for Lie algebras, and depend on N for SU(N)-theories; e.g. for $N = 2$, they apply to three paths:

$$W(\gamma_1)W(\gamma_2)W(\gamma_3) - \frac{1}{2}(W(\gamma_1\gamma_2)W(\gamma_3) + W(\gamma_2\gamma_3)W(\gamma_1) + W(\gamma_1\gamma_3)W(\gamma_2) \\ +\frac{1}{4}(W(\gamma_1\gamma_2\gamma_3) + W(\gamma_1\gamma_3\gamma_2) = 0. \tag{B.4}$$

References

Arnowitt, R., Deser, S., & Misner, C. (1962). The dynamics of general relativity, pp. 227-264. In *in gravitation: an introduction to current research, l. witten, ed.* Wiley, New York.

Bagan, E., Lavelle, M., & McMullan, D. (2000). Charges from dressed matter: construction. *Annals of Physics*, *282*(2), 471–502.

Bamonti, N. (2023). *What is a reference frame in general relativity?*

Barrett, J. W. (1991, Sep 01). Holonomy and path structures in general relativity and yang-mills theory. *International Journal of Theoretical Physics*, *30*(9), 1171–1215. Retrieved from https://doi.org/10.1007/BF00671007 doi: 10.1007/BF00671007

Belot, G. (1998). Understanding Electromagnetism. *The British Journal for the Philosophy of Science*, *49*(4), 531-555. Retrieved from https://doi.org/10.1093/bjps/49.4.531 doi: 10.1093/bjps/49.4.531

Belot, G. (2013). Symmetry and Equivalence. In *The Oxford Handbook of Philosophy of Physics.* Oxford University Press. Edited by Batterman, R.

Belot, G. (2018). Fifty Million Elvis Fans Can't be Wrong. *Nous*, *52*(4), 946-981. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/nous.12200 doi: 10.1111/nous.12200

Belot, G., & Earman, J. (1999). From metaphysics to physics. In J. Butterfield & C. Pagonis (Eds.), *From physics to philosophy* (p. 166–186). Cambridge University Press. doi: 10.1017/CBO9780511597947.009

Belot, G., & Earman, J. (2001). Pre-Socratic quantum gravity. In C. Callender & N. Huggett (Eds.), *Physics meets philosophy at the planck scale: Contemporary theories in quantum gravity* (p. 213-255). Cambridge University Press. doi: 10.1017/CBO9780511612909.011

Bergmann, P. G. (1961). Observables in General Relativity. *Rev. Mod. Phys.*, *33*, 510–514. doi: 10.1103/RevModPhys.33.510

Black, M. (1952). The Identity of Indiscernibles. *Mind*, *61*(242), 153–164. Retrieved 2022-07-15, from http://www.jstor.org/stable/2252291

Brading, K., & Brown, H. R. (2000). Noether's theorems and gauge symmetries.

Bradley, C., & Weatherall, J. O. (2022). Mathematical responses to the hole argument: Then and now. *Philosophy of Science*, *89*(5), 1223–1232. doi: 10.1017/psa.2022.58

Butterfield, J. (1989, 03). The Hole Truth. *The British Journal for the Philosophy of Science*, *40*(1), 1-28. Retrieved from https://doi.org/10.1093/bjps/40.1.1 doi: 10.1093/bjps/40.1.1

Butterfield, J., & Gomes, H. (2022). The Hole Argument: a Drama in Five Acts. *in preparation*.

Carrozza, S., & Höhn, P. A. (2022, feb). Edge modes as reference frames and boundary actions from post-selection. *Journal of High Energy Physics*, *2022*(2). doi: 10.1007/jhep02(2022)172

Curiel, E. (2015). *If metrical structure were not dynamical, counterfactuals in general relativity would be easy.*

Curiel, E. (2018). On the Existence of Spacetime Structure. *The British Journal for the Philosophy of Science*, *69*(2), 447-483. Retrieved from https://doi.org/10.1093/bjps/axw014 doi: 10.1093/bjps/axw014

Dewar, N. (2017, 09). Sophistication about Symmetries. *The British Journal for the Philosophy of Science*, *70*(2), 485-521. Retrieved from https://doi.org/10.1093/bjps/axx021 doi: 10.1093/bjps/axx021

Diez, T., & Rudolph, G. (2019, Aug). Slice theorem and orbit type stratification in infinite dimensions. *Differential Geometry and its Applications*, *65*, 176–211. Retrieved from http://dx.doi.org/10.1016/j.difgeo.2019.03.005 doi: 10.1016/j.difgeo.2019.03.005

Donnelly, W., & Freidel, L. (2016). Local subsystems in gauge theory and gravity. *JHEP*, *09*, 102. doi: 10.1007/JHEP09(2016)102

Donnelly, W., & Giddings, S. B. (2016, Jan). Diffeomorphism-invariant observables and their nonlocal algebra. *Physical Review D*, *93*(2). Retrieved from http://dx.doi.org/10.1103/PhysRevD.93.024030 doi: 10.1103/physrevd.93.024030

Duan, Y.-S., & Ge, M.-L. (1979). SU(2) Gauge Theory and Electrodynamics with N Magnetic Monopoles. *Sci. Sin.*, *9*(11). doi: 10.1142/9789813237278_0001

Earman, J. (1989). *World Enough and Spacetime.* MIT press.

Ebin, D. (1970). The manifold of riemmanian metrics. *Symp. Pure Math., AMS, 11*, *11,15*.

Einstein, A. (1916). Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik, 49. pp 769-822*.

Field, H. (1984). Can we dispense with space-time? *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1984*, 33–90. Retrieved 2022-08-30, from http://www.jstor.org/stable/192496

Fischer, A. E., & Marsden, J. E. (1979). The initial value problem and the dynamical formulation of general relativity. In *General relativity : an einstein centenary survey. cambridge university press , new york, pp. 138-211.*

Fletcher, S. C. (2020). On Representational Capacities, with an Application to General Relativity. *Foundations of Physics*, *50*(4), 228–249. doi: 10.1007/s10701-018-0208-6

François, J. (2021, mar). Bundle geometry of the connection space, covariant Hamiltonian formalism, the problem of boundaries in gauge theories, and the dressing field method. *Journal of High Energy Physics*, *2021*(3). doi: 10.1007/jhep03(2021)225

François, J., Parrini, N., & Boulanger, N. (2021, dec). Note on the bundle geometry of field space, variational connections, the dressing field method, presymplectic structures of gauge theories over bounded regions. *Journal of High Energy Physics*, *2021*(12). doi: 10.1007/jhep12(2021)186

Geiller, M. (2017). Edge modes and corner ambiguities in 3d Chern–Simons theory and gravity. *Nucl. Phys.*, *B924*, 312-365. doi: 10.1016/j.nuclphysb.2017.09.010

Geiller, M., & Jai-akson, P. (2020, Sep). Extended actions, dynamics of edge modes, and entanglement entropy. *Journal of High Energy Physics*, *2020*(9). Retrieved from http://dx.doi.org/10.1007/JHEP09(2020)134 doi: 10.1007/jhep09(2020)134

Gilbarg, D., & Trudinger, N. (2001). *Elliptic Partial Differential Equations of Second Order.* Springer.

Gil-Medrano, O., & Michor, P. W. (1991). The Riemannian Manifold Of All Riemannian Metrics. *Quarterly Journal of Mathematics, (42), 183-202*.

Giovanelli, M. (2021, may). Nothing but coincidences: the point-coincidence and Einstein's struggle with the meaning of coordinates in physics. *European Journal for Philosophy of Science*, *11*(2). doi: 10.1007/s13194-020-00332-7

Giulini, D. (1995). Asymptotic symmetry groups of long-ranged gauge configurations. *Modern Physics Letters A*, *10*(28), 2059-2070. Retrieved from https://doi.org/10.1142/S0217732395002210 doi: 10.1142/S0217732395002210

Gomes, H. (2019). Gauging the boundary in field-space. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*. Retrieved from http://www.sciencedirect.com/science/article/pii/S1355219818302144 doi: https://doi.org/10.1016/j.shpsb.2019.04.002

Gomes, H. (2021a). Holism as the significance of gauge symmetries. *European Journal of Philosophy of Science, vol 11, 87*.

Gomes, H. (2021b). The role of representational conventions in assessing the empirical significance of symmetries. *(to appear in Studies in History and Philosophy of Modern Physics)*.

Gomes, H. (2021c). Same-diff? Conceptual similarities between gauge transformations and diffeomorphisms. Part II: Challenges to sophistication. *Arxiv: 2110.07204. Submitted.*.

Gomes, H. (2021d). Same-diff? Conceptual similarities between gauge transformations and diffeomorphisms. Part I: Symmetries and isomorphisms. *Arxiv: 2110.07203. Submitted.*.

Gomes, H. (2021e). Same-diff? Part IV: A compendium of similarities between gauge transformations and diffeomorphisms. *Arxiv: 2110.07204. Submitted.*.

Gomes, H. (2022). *Why gauge? Conceptual Aspects of Gauge theories.* arXiv. Retrieved from https://arxiv.org/abs/2203.05339

Gomes, H. (2024). Gauge theory as the geometry of internal vector spaces.

Gomes, H., & Butterfield, J. (2021). Counterpart relations in general relativity (forthcoming).

Gomes, H., & Butterfield, J. (2022, aug). How to Choose a Gauge? The Case of Hamiltonian Electromagnetism. *Erkenntnis*. doi: 10.1007/s10670-022-00597-9

Gomes, H., & Butterfield, J. (2023a, jun). The Hole Argument and Beyond: Part II: Treating Non-isomorphic Spacetimes. *Journal of Physics: Conference Series*, *2533*(1), 012003. Retrieved from https://dx.doi.org/10.1088/1742-6596/2533/1/012003 doi: 10.1088/1742-6596/2533/1/012003

Gomes, H., & Butterfield, J. (2023b, jun). The Hole Argument and Beyond: Part I: The Story so Far. *Journal of Physics: Conference Series*, *2533*(1), 012002. Retrieved from https://dx.doi.org/10.1088/1742-6596/2533/1/012002 doi: 10.1088/1742-6596/2533/1/012002

Gomes, H., Hopfmüller, F., & Riello, A. (2019). A unified geometric framework for boundary charges and dressings: Non-abelian theory and matter. *Nuclear Physics B*, *941*, 249 - 315. Retrieved from http://www.sciencedirect.com/science/article/pii/S0550321319300483 doi: https://doi.org/10.1016/j.nuclphysb.2019.02.020

Gomes, H., & Riello, A. (2017). The observer's ghost: notes on a field space connection. *Journal of High Energy Physics (JHEP)*, *05*, 017. Retrieved from https://link.springer.com/article/10.10072FJHEP0528201729017 doi: 10.1007/JHEP05(2017)017

Gomes, H., & Riello, A. (2018, Jul). Unified geometric framework for boundary charges and particle dressings. *Physical Review D*, *98*, 025013. Retrieved from https://link.aps.org/doi/10.1103/PhysRevD.98.025013 doi: 10.1103/PhysRevD.98.025013

Gomes, H., & Riello, A. (2021). The quasilocal degrees of freedom of Yang-Mills theory. *SciPost Phys.*, *10*, 130. Retrieved from https://scipost.org/10.21468/SciPostPhys.10.6.130 doi: 10.21468/SciPostPhys.10.6.130

Gribov, V. N. (1978). Quantization of Nonabelian Gauge Theories. *Nucl. Phys.*, *B139*, 1. ([,1(1977)]) doi: 10.1016/0550-3213(78)90175-X

Hall, N. (2015). Humean Reductionism about Laws of Nature. In *A companion to david lewis* (pp. 262–277). John Wiley and Sons, Ltd. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118398593.ch17 doi: https://doi.org/10.1002/9781118398593.ch17

Harlow, D., & Wu, J.-Q. (2021). *Algebra of diffeomorphism-invariant observables in jackiw-teitelboim gravity.*

Healey, R. (2007). *Gauging What's Real: The Conceptual Foundations of Gauge Theories.* Oxford University Press.

Henneaux, M., & Teitelboim, C. (1992). *Quantization of gauge systems.* Princeton University Press.

Isenberg, J., & Marsden, J. E. (1982). A slice theorem for the space of solutions of einstein's equations. *Physics Reports*, *89*(2), 179–222.

Isham, C. J. (1992). Canonical quantum gravity and the problem of time. In *19th International Colloquium on Group Theoretical Methods in Physics (GROUP 19) Salamanca, Spain, June 29-July 5, 1992.*

Jacobs, C. (2022). Invariance or Equivalence: A Tale of Two Principles. *Synthese*, 1–21. doi: 10.1007/s11229-021-03205-5

Jaramillo, J. L., & Lam, V. (2021). Counterfactuals in the Initial Value Formulation of General Relativity. *The British Journal for the Philosophy of Science*, *72*(4), 1111-1128. Retrieved from https://doi.org/10.1093/bjps/axy066 doi: 10.1093/bjps/axy066

Kabel, V., de la Hamette, A.-C., Apadula, L., Cepollaro, C., Gomes, H., Butterfield, J., & Časlav Brukner. (2024). *Identification is pointless: Quantum reference frames, localisation of events, and the quantum hole argument.*

Komar, A. (1958, Aug). Construction of a Complete Set of Independent Observables in the General Theory of Relativity. *Physical Review*, *111*, 1182–1187. Retrieved from https://link.aps.org/doi/10.1103/PhysRev.111.1182 doi: 10.1103/PhysRev.111.1182

Kondracki, W., & Rogulski, J. (1983). *On the stratification of the orbit space for the action of automorphisms on connections. on conjugacy classes of closed subgroups. on the notion of stratification.* Inst., Acad. Retrieved from https://books.google.co.uk/books?id=LK0JrgEACAAJ

Landsman, K. (2021). *Foundations of General Relativity.* Radboud University Press.

Lavelle, M., & McMullan, D. (1995). Observables and gauge fixing in spontaneously broken gauge theories. *Phys. Lett.*, *B347*, 89-94. doi: 10.1016/0370-2693(95)00046-N

Lavelle, M., & McMullan, D. (1997). Constituent quarks from QCD. *Phys. Rept.*, *279*, 1-65. doi: 10.1016/S0370-1573(96)00019-1

Lewis, D. (1968). Counterpart Theory and Quantified Modal Logic. *Journal of Philosophy*, *65*(5), 113–126. doi: 10.2307/2024555

Lewis, D. (1970). How to Define Theoretical Terms. *Journal of Philosophy,* **67***,* pp. 427-446..

Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*.

Lewis, D. (1973). *Counterfactuals.* Cambridge, MA, USA: Blackwell.

Lewis, D. (1986). *On the Plurality of Worlds.* Blackwell Publishers.

Lewis, D. (2009). Ramseyan Humility. In D. Braddon-Mitchell & R. Nola (Eds.), *Conceptual analysis and philosophical naturalism* (pp. 203–222). MIT Press.

Martens, N., & Read, J. (2020). Sophistry about symmetries? *Synthese*. Retrieved from http://philsci-archive.pitt.edu/17184/

Mattingly, J. (2006). Which gauge matters? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *37*(2), 243-262. Retrieved from https://www.sciencedirect.com/science/article/pii/S1355219806000037 doi: https://doi.org/10.1016/j.shpsb.2005.08.001

Maudlin, T. (1988). The Essence of Space-Time. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1988*, 82–91.

Maudlin, T. (1990). Substances and space-time: What Aristotle would have said to Einstein. *Studies in History and Philosophy of Science Part A*, *21*(4), 531-561. Retrieved from https://www.sciencedirect.com/science/article/pii/0039368190900324 doi: https://doi.org/10.1016/0039-3681(90)90032-4

Maudlin, T. (1998). Healey on the aharonov-bohm effect. *Philosophy of Science*, *65*(2), 361–368. Retrieved from http://www.jstor.org/stable/188266

Maudlin, T. (2018). Ontological clarity via canonical presentation: Electromagnetism and the aharonov–bohm effect. *Entropy*, *20*(6). Retrieved from https://www.mdpi.com/1099-4300/20/6/465 doi: 10.3390/e20060465

Mercati, F. (2017). *Shape Dynamics: Relativity and Relationalism.* Oxford University Press.

Mitter, P. K., & Viallet, C. M. (1981). On the Bundle of Connections and the Gauge Orbit Manifold in Yang-Mills Theory. *Commun. Math. Phys.*, *79*, 457. doi: 10.1007/BF01209307

Mulder, R. (2021). Gauge-Underdetermination and Shades of Locality in the Aharonov–Bohm Effect. *Foundations of Physics*.

Mundy, B. (1992). Space-Time and Isomorphism. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1992*(Volume One: Contributed Papers), 515–527.

Palais, R. (1961). On the existence of slices for the actions of non-compact groups. *Ann. of Math.*, *73*, 295-322.

Penrose, R. (1996, May). On Gravity's role in Quantum State Reduction. *General Relativity and Gravitation*, *28*(5), 581-600. doi: 10.1007/BF02105068

Pooley, O. (2022). *The Reality of Spacetime.* Unpublished manuscript.

Pooley, O. (2023). The Hole Argument. In E. Knox & A. Wilson (Eds.), *The routledge companion to the philosophy of physics.* Routledge. Retrieved from http://philsci-archive.pitt.edu/18142/

Pooley, O., & Read, J. (2022). On the Mathematics and Metaphysics of the Hole Argument. *The British Journal for the Philosophy of Science*. doi: 10.1086/718274

Pretorius, F. (2005, sep). Evolution of binary black-hole spacetimes. *Physical Review Letters*, *95*(12). Retrieved from https://doi.org/10.1103%2Fphysrevlett.95.121101 doi: 10.1103/physrevlett.95.121101

Ramirez, S., & Teh, N. (2019). Abandoning galileo's ship: The quest for non-relational empirical signicance. *preprint*.

Riello, A. (2021). Symplectic reduction of Yang-Mills theory with boundaries: from superselection sectors to edge modes, and back. *SciPost Phys.*, *10*, 125. Retrieved from https://scipost.org/10.21468/SciPostPhys.10.6.125 doi: 10.21468/SciPostPhys.10.6.125

Rosenstock, S., & Weatherall, J. O. (2016). A categorical equivalence between generalized holonomy maps on a connected manifold and principal connections on bundles over that manifold. *Journal of Mathematical Physics*, *57*(10), 102902. (http://philsci-archive.pitt.edu/11904/)

Rovelli, C. (2002, January). GPS observables in general relativity. *Physical Review D*, *65*(4). Retrieved from http://dx.doi.org/10.1103/PhysRevD.65.044017 doi: 10.1103/physrevd.65.044017

Rovelli, C. (2007). *Quantum Gravity*. Cambridge University Press.

Rovelli, C. (2014). Why Gauge? *Found. Phys.*, *44*(1), 91-104. doi: 10.1007/s10701-013-9768-7

Singer, I. M. (1978). Some Remarks on the Gribov Ambiguity. *Commun. Math. Phys.*, *60*, 7-12. doi: 10.1007/BF01609471

Thiemann, T. (2003). Lectures on loop quantum gravity. *Lecture Notes in Physics*, 41–135. Retrieved from http://dx.doi.org/10.1007/978-3-540-45230-0_3 doi: 10.1007/978-3-540-45230-0_3

Tong, D. (2018). *Lecture notes on gauge theory.* (Available online at http://www.damtp.cam.ac.uk/user/tong/gaugetheory.html)

Walker, M. L., & Duplij, S. (2015). Cho-Duan-Ge decomposition of QCD in the constraintless Clairaut-type formalism. *Phys. Rev. D*, *91*(6), 064022. doi: 10.1103/PhysRevD.91.064022

Wallace, D. (2002). Time-Dependent Symmetries: The Link Between Gauge Symmetries and Indeterminism. In K. Brading & E. Castellani (Eds.), *Symmetries in physics: Philosophical reflections* (pp. 163–173). Cambridge University Press.

Wallace, D. (2019). Observability, redundancy and modality for dynamical symmetry transformations. *Forthcoming*. Retrieved from http://philsci-archive.pitt.edu/18813/ (Revised 3/2021 to correct a few typos and add a section on Noether's Theorem.)

Wallace, D. (2024). *Gauge invariance through gauge fixing.*

Weatherall, J. (2018). Regarding the 'Hole Argument'. *The British Journal for the Philosophy of Science*, *69*(2), 329-350. Retrieved from https://doi.org/10.1093/bjps/axw012 doi: 10.1093/bjps/axw012

Wilkins, D. R. (1989). Slice theorems in gauge theory. *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, *89A*(1), 13–34. Retrieved from http://www.jstor.org/stable/20489307

Witten, E. (2023). *A background independent algebra in quantum gravity.*

York, J. W. (1971). Gravitational degrees of freedom and the initial-value problem. *Phys. Rev. Lett.*, *26*, 1656–1658. doi: 10.1103/PhysRevLett.26.1656