Symmetries as a Guide to the Structure of Physical Quantities

Caspar Jacobs

Magdalen College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2021

Abstract

The aim of this thesis is to extend the success of a novel strategy for the interpretation of symmetry-related models — sophistication — to contemporary physical theories' internal symmetries. In order to do so, I propose that a theory's quantities possess a certain structure. The problematic underdetermination of symmetry-related models arises when we attribute too much structure to the fundamental quantities.

The advantage of sophistication is that it enables a realist interpretation of symmetry-variant quantities, which often figure in the scientific explanation of physical effects. I illustrate sophistication with two case studies: one of mass in Newtonian Gravitation and one of gauge quantities in the Aharonov-Bohm effect. On the traditional approach, symmetries are seen as a guide to superfluous theoretical structure. The alternative I offer demands a reconceptualisation of symmetries as playing a more positive role, namely as a *guide towards the structure of physical quantities*.

This thesis is approximately 65,000 words long.

Acknowledgements

I am most grateful to the British Society for the Philosopy of Science for their generous doctoral scholarship, with which I have been able to spend the past three years in full-time research.

This thesis could not have been written without the continued support and intelligent insights of my primary supervisor, Adam Caulton. I could not have imagined a mentor more kind with his words, or more generous with his time. Adam is responsible for any reference to Carnap in this thesis, but it goes without saying that all remaining mistakes are my own.

In the final year of the DPhil I was lucky enough to have James Read on board as my second supervisor. We have had many productive discussions on topics beyond the scope of this thesis, and I look forward to continuing these conversations in the future.

Within the philosophy of physics graduate community at Oxford, I have enjoyed the friendship of Lydia Drabkin-Reiter, Lia den Daas and Natasha Oughton, among others. It has been a great pleasure to organise both the POP-Grunch and the POP-Gronference with them. I thank my housemate Itzhak Rasooly for the endless games of Backgammon.

I thank Marta Bielińska, Neil Dewar, Henrique Gomes, Niels Martens, Rob Rockwood, David Schroeren and two anonymous reviewers at *Synthese* for invaluable feedback on various parts of this thesis. I also thank Marco Bodnar, Jeremy Butterfield, Patrick Duerr, Guy Hetzroni, Carl Hoefer, Jenann Ismael, Nat Levine, Joanna Luc, Tushar Menon, Ruward Mulder, Sebastián Murgueitio Ramírez, Zee Perry, Oliver Pooley, Imogen Rivers, Carlo Rovelli, Simon Saunders, Daniel Sega, Neil Swanson, David Wallace and James Weatherall for their encouragement and helpful discussions.

Finally, thanks to my parents, Andreas and Judith; and to my partner, Kim — writing this thesis has been so much better with a home to call ours wherever we went.

Contents

Chapter 1

Introduction: Literalism, Reduction and Sophistication

This dissertation concerns one of the most distinctive features of contemporary theories in physics: their symmetries. Symmetries relate mathematical models that may represent physically distinct yet empirically indiscernible possibilities, implying that the quantities which vary under those symmetries are undetectable. Often-mentioned examples of such variant quantities are absolute position, intrinsic mass and the electromagnetic potential. Their undetectability in turn implies a form of underdetermination: the empirical facts fail to determine the physical facts. The puzzle posed by symmetries is that our most successful theories seem to possess empirically redundant structure.

The orthodox response to this puzzle is *reduction*. Reductionism advocates the formulation of a 'reduced' theory whose equations involve some set of new, invariant quantities. For example, a reduced theory of classical mechanics employs distance instead of position and mass ratios instead of intrinsic masses. Since reduction rids the old theory of its variant quantities, it ensures that the new theory has no symmetries left. Symmetries are, in the words of Ismael and van Fraassen (2003), *a guide to* *superfluous theoretical structure.*

Recently, an alternative view called *sophistication* has become popular. Sophistication promises to overcome underdetermination without rejecting the variant quantities as unreal. It does so by assuming that there are no worlds that are numerically distinct yet qualitatively identical: a thesis known as *anti-haecceitism.* Variant quantities often figure in explanations of physical effects, so in retaining them, sophistication gains an advantage over reduction. But sophistication only succeeds once a theory's symmetry-related models are isomorphic, since only then do they agree on all structural facts. The aim of sophistication is thus to restructure a theory's models such that this is the case.

So far, the focus of sophistication has been on 'external' symmetries: roughly, symmetries of spacetime. The objective of this dissertation is to extend sophistication's success to 'internal' symmetries: symmetries that act on a quantity's internal value space. In order to do so, I propose that we think of quantities as possessing a certain structure. For example, the structure of mass consists of the ratios between its determinate values. I aim to show that underdetermination is a consequence of attributing too much structure to variant quantities relative to the theory's dynamics. But when this structure meshes with the dynamics, one can appeal to the analogue of anti-haecceitism for properties, called *anti-quidditism*; the underdetermination then disappears. So, our aim as interpreters of physical theories is to find the correct structures for the theory's fundamental quantities. Contrary to reduction, symmetries here perform a positive role as *guides to the structure of physical quantities.*

In this introductory chapter, I set out the debate between reduction and sophistication. I first offer a relatively informal account of symmetries in the context of physical theories (§1.1). I then define three approaches to the interpretation of symmetries: literalism (§1.2.1), reduction (§1.2.2) and sophistication (§1.2.3). In §1.3, I offer some preliminary reasons against reduction. The chapter concludes with an outline for the remainder of this dissertation (§1.4).

1.1 Symmetries

In short, a symmetry is *a transformation of some object that leaves some of its salient features the same*. These salient features are said to be *invariant* under the symmetry in question. For a simple example, consider a reflection in the mirror. This is a non-trivial transformation that changes an object's 'parity'—in the mirror, a left hand becomes a right hand—yet leaves its shape and size the same. We thus say that shape and size are invariant under reflections.

The symmetries we are interested in do not act on physical objects, such as hands. Rather, they act on a theory's *models*: mathematical structures that represent possible ways the world could be. For example, the models of classical mechanics are structures that represent the trajectories of particles in a three-dimensional Euclidean space. On the *semantic view* of theories, a theory is presented as a family of such models.[1] In this dissertation, models are structures in the sense of Bourbaki (1957): tuples of sets, relations defined over these sets, and functions between them.

In order to define the symmetries we are interested in, it is useful to distinguish between two classes of models. The *kinematically possible models* (KPMs) of a theory are those of the correct 'form', that is, those which contain the right type of mathematical objects. The *dynamically possible models* (DPMs) are those KPMs that in addition satisfy a set of equations of motion. The DPMs represent ways the world could be if the theory were true. For example, the KPMs of Newtonian Gravitation set on Galilean space-time are models of the form $\langle M, t_{ab}, h^{ab}, \nabla, \varphi, \rho, \xi^a \rangle$, where M is a four-dimensional smooth manifold diffeomorphic to $\mathbb{R}^4$; t_{ab} and h^{ab} are compatible temporal and spatial metrics ($t_{an}h^{nb} = 0$); ∇ is a covariant derivative encoding a

[1] See Suppes (1960); van Fraassen (1980); Suppe (1989); French and Ladyman (1999), and references therein.

standard of straight motion compatible with both metrics ($\nabla_a t_{bc} = \nabla_a h^{bc} = 0$); φ and ρ are scalar fields that represent the gravitational potential and the matter distribution respectively; and ξ^a is a timelike vector field whose integral curves represent the four-velocity of test particles.[2] The DPMs of Newtonian Gravitation in addition satisfy the following dynamical equations:

$$R^a{}_{bcd} = 0 \tag{1.1}$$

$$h^{ab}\nabla_a\nabla_b\varphi = 4\pi\rho \tag{1.2}$$

$$-\nabla^a\varphi = \xi^b\nabla_b\xi^a \tag{1.3}$$

where $R^a_{bcn}\xi^n = \nabla_{[b}\nabla_{c]}\xi^a$. Here, (1.1) imposes flatness on ∇; (1.2) is the Newton-Poisson equation; and (1.3) is Newton's second law for test particles. The models that satisfy these equations represent worlds in which Newtonian Gravitation is true.

I now return to the task of defining the symmetries of a theory. As a first approximation, we can think of the symmetries of a theory as bijections on the space of KPMs that preserve the space of DPMs. In other words, if we let $\mathcal{K}$ and $\mathcal{D}$ denote the space of KPMs and DPMs respectively, then a symmetry transformation is an injective function g from $\mathcal{K}$ onto itself such that for any model $\mathfrak{m} \in \mathcal{K}$, $\mathfrak{m} \in \mathcal{D}$ if and only if $g(\mathfrak{m}) \in \mathcal{D}$. Whenever a transformation possesses this feature, we say that it *preserves the dynamics* of the theory. In short, a theory's symmetries are transformations of the theory's models that preserve the dynamics. For example, *static shifts* are symmetries of Newtonian Gravitation: uniform translations of the entire material contents of the universe. If $\mathfrak{m}$ is a model of Newtonian Gravitation, then so is any model $\mathfrak{m}'$ obtained from the first by such a shift, and *vice versa*. Unfortunately, it is well-known that this definition of symmetries is much too liberal: Belot (2013)

[2] For more details, see Malament (2012, §4.2).

calls it 'fruitless'—with good reason. For instance, on this definition any arbitrary permutation of DPMs counts as a symmetry. But this would mean that *any* pair of models is related by some symmetry, trivialising the notion of symmetries altogether. As a consequence, not all symmetries of the above kind lead to the sort of underdetermination discussed above. In what follows, I will only consider those symmetries which in addition relate empirically equivalent models. To be sure, this is true for most (if not all) symmetries of physical interest, so nothing is lost with this proviso. In Chapter 5, I will offer a more detailed definition of symmetries, namely as being maps on the space of models which are the 'lift' of some permutation of these models' common domain.[3]

I will also assume that a theory's symmetry transformations possess a *group structure*. This assumption holds for all concrete instances of symmetries discussed in this dissertation. Consider a set of symmetry transformations labeled G (so each element of G defines a bijection of $\mathcal{K}$ which preserves the dynamics). The claim that G is a group means that:

(i) For any pair of symmetry transformations g and h in G, the successive application of g and h denoted $g \circ h$ forms another element of G;

(ii) For any symmetry transformations $g, h, i \in G$, one has $g \circ (h \circ i) = (g \circ h) \circ i$;

(iii) There exists an identity transformation $e \in G$ such that $g \circ e = e \circ g = g$ for all $g \in G$;

(iv) For any symmetry transformation $g \in G$, there exists an inverse transformation $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.

When these axioms are satisfied, the space of models is partitioned into so-called

[3]Note that this does not amount to an 'epistemic' definition of symmetries, as advocated by Dasgupta (2016). Rather, I retain a 'formal' definition of symmetries, but restrict attention to a subclass of all symmetries defined in this way. For this reason, the approach is closer to that of Ismael and van Fraassen (2003). For more on these issues, see Read and Møller-Nielsen (2020).

equivalence classes of symmetry-related models (SRMs). We can understand equivalence classes as sets of 'connected' models that one can transform into each other via some symmetry transformation. So, if a pair of models lies within the same equivalence class, then there is some symmetry that transforms one into the other; if a pair of models lies within distinct equivalence classes, then there exists *no* symmetry that transforms one into the other. When symmetries relate empirically equivalent models, it follows that all models within an equivalence class are empirically equivalent.

Finally, recall that symmetries are informally defined as transformations of an *object* that leave invariant some salient *features*. For symmetries of a physical theory, the relevant objects are the models. What are the salient features that remain the same? Clearly, one such feature is 'solutionhood', or the property of being a solution to the theory's equations of motion. But one can also consider the theory's *quantities*: variables such as location, mass or force. Invariant quantities are quantities whose values remain the same under any symmetry transformation. For example, recall that constant shifts of all matter content are symmetries of Newtonian Gravitation. The distances between particles are invariant under the group of shift transformations, since a uniform translation of all particles in space does not alter the distances between them. On the other hand, the positions of particles vary under translations. Generally, we can divide the theory's quantities into two disjoint classes: the variant and the invariant ones. The former must remain the same after a symmetry transformation, whereas the latter may vary under such transformations. As Caulton (2015) notes, there is a duality between a theory's symmetries and its invariant quantities: the more symmetries a theory possesses, the fewer in number its invariant quantities are, and *vice versa*. As we will see below, one reason philosophers are interested in symmetries is the belief that only invariant quantities are physically real. If this is so, then one can employ so-called 'symmetry-to-(un)reality' inferences from the variance of certain quantities under the theory's symmetries to their non-existence (Dasgupta, 2013). One of the main aims of this dissertation is to resist such inferences.

1.2 Three Approaches

I will now consider the *interpretation* of symmetries. Recall that a theory's dynamically possible models represent ways the world could be if the theory were true. But there is no requirement that they represent possibilities in a one-to-one fashion. When models represent the same state of affairs, I will call them *physically equivalent*. The first question I wish to address is:

(1) Should we interpret symmetry-related models as physically equivalent?

This question arises for symmetry-related models specifically because symmetries are systematic transformations of a theory's models that leave some of their salient features the same. So, if the salient features are just those that are representationally significant, then it follows that symmetry-related models represent the same state of affairs.

I will discuss three approaches to the interpretation of symmetry-related models. The first, *literalism*, answers 'No' the the first question: it says that SRMs, if distinct, are physically *in*equivalent. However, literalism implies an unwelcome form of empirical underdetermination, and in some cases also leads to radical indeterminism. For this reason, my focus is on approaches which answer 'Yes' to (1), that is, approaches on which an entire *equivalence class* of symmetry-related models jointly represents a single physical possibility. It may seem as if this answer implies that only a theory's invariant content—the salient features that remain the same after a symmetry transformation—has representational significance. But as we will see below, the issue is subtler than that. Call the physical content common to all models within an equivalence class its 'representational core'. If one answers 'Yes' to the first question, then this raises a second question:

(2) What is the representational core of an equivalence class of symmetry-related models?

As we will see below, reduction and sophistication disagree over the answer to this second question.

With these questions in mind, I will now discuss the three approaches in turn.

1.2.1 Literalism

The first position is called *literalism*.[4] Literalism answers 'No' to the first question: a theory's SRMs represent distinct states of affairs. In particular, literalism argues that since the values of a theory's variant quantities differ across SRMs, they must represent distinct possibilities. For example, models related by a static shift disagree over the value of any particle's position function. So, literalism infers that such models represent the same particle as located in different places.

The problem with literalism is that it seems to lead to a particularly puzzling form of underdetermination. This is a consequence of the claim that a theory's SRMs are *empirically equivalent*: the possibilities that these models represent are identical insofar as their empirical content is concerned. Now, whether this is true for *all* symmetries of a certain class is a contentious issue. But it is true for many symmetries of philosophical interest, from the space-time symmetries of Newtonian Gravitation and General Relativity to the gauge symmetries of electrodynamics. With respect to this class of symmetries, literalism faces several related problems. Firstly, the existence of distinct models that represent the same empirical state of affairs implies

[4] The terminology is inspired by Rynasiewicz (1994, 1996), who uses the term 'model literalism'. The idea that literally interpreted, SRMs represent distinct states of affairs is widespread. For example, Pooley (2006, 102) discusses (but rejects!) the view that "the theoretical treatment of spacetime, read literally, strongly support haecceitism". Similarly, Read and Møller-Nielsen's (2018, 270) claim that "[s]traightforwardly—or 'literally'—understood, the world represented by [a model of Newtonian Gravitation] differs from that represented by [another model related to the first by a translation symmetry] with regard to which particular points of space are underlying various parts of the matter field".

a form of underdetermination. The theory draws distinctions that are more fine-grained than our empirical evidence. What seems particularly suspicious here is that it is the theory *itself* that tells us that the difference between these possibilities is undetectable. Secondly, the quantities that vary under symmetries are undetectable.[5] Occam's Razor tells us that all else being equal, we are better off without such undetectable structure. Finally, when symmetries are *local* (roughly, when we can apply them for an arbitrary stretch of time) the existence of SRMs seems to imply a particularly harmful form of indeterminism. This last claim is most familiar from the so-called 'Hole Argument' in General Relativity. These are all strong reasons against literalism which motivate an alternative approach on which SRMs are physically equivalent.[6] If SRMs are physically equivalent the above issues do not arise, since putatively distinct models in fact represent the same state of affairs. In what follows, I will not further consider literalism.[7]

1.2.2 Reduction

The main alternatives to literalism, reduction and sophistication, both answer 'Yes' to the first question: physical possibilities are represented by equivalence classes of SRMs. This leads to the second question: what is the representational core of such an equivalence class of SRMs? Reduction and sophistication offer different answers to

[5] For arguments to this effect, see Roberts (2008); Dasgupta (2016). For critical responses, see Wallace (2019b); Middleton and Murgueitio Ramírez (2020), and for a critical response to these responses see Jacobs (2020).

[6] I write 'motivate' here with reason. According to a view called *interpretationalism*, one can *invariably* interpret SRMs as physically equivalent, even without a perspicuous account of the possibility that equivalence classes of SRMs are supposed to represent (i.e. their representational core). But on the *motivational* approach, defended by Møller-Nielsen (2017), one can only interpret SRMs as physically equivalent once one has provided a perspicuous metaphysical picture in terms of which we can understand their common representational content. I subscribe to the motivational approach, so in the absence of a perspicuous metaphysical picture literalism is the default option.

[7] As far as I am aware, the only defence of literalism is Maudlin's (1988) essentialism (Maudlin's (1998) 'ONE TRUE GAUGE' is also a version of literalism, but not one Maudlin himself defends; see also Chapter 7). On this view, SRMs represent physically inequivalent states of affairs, but only one of those states is physically possible. For a recent criticism of essentialism, see Teitel (2019). In addition, one could also interpret Belot (2018) as advocating a qualified form of literalism.

this second question. Both approaches have an *interpretational* and a *formal* aspect. I discuss reduction in this subsection, and sophistication in the next.[8]

According to reduction, the theory's representational core consists of its invariant quantities.[9] In other words, models represent the physical values of those quantities—and nothing else. By construction, these values are constant within equivalence classes of SRMs, so such equivalence classes jointly represent the same state of affairs. For example, the invariant quantities of Newtonian Gravitation are distance, relative velocity and acceleration. So, on a reductionist interpretation the models of this theory represent the distances and relative velocities between particles, as well as their accelerations, but not their (absolute) positions or velocities. The models within an equivalence class of shift-related models are physically equivalent because they agree on the values of the former set of quantities.

The claim that only invariant quantities are physically real concerns the *interpretation* of a theory. But so far, reduction does not alter the theory's *formalism*. This is unsatisfactory. After all, the formalism still involves variant quantities such as position or velocity. In Arntzenius' (2012) words, reduction 'piggy-backs' on the old theory's success. In the same vein, Earman (1989) calls such interpretations 'cheap instrumentalist knock-offs', while Belot (2000) uses the phrase 'arrant knavery'. In addition to the identification of the theory's representational core with its invariant quantities, a satisfactory form of reduction will also have to offer a novel formalism which solely involves these quantities. Dewar (2019, 492) outlines a general procedure: "(i) identify some collection of invariants of the original theory; (ii) specify a theory in terms of those invariants; and (iii) show that the new theory captures all

[8] As far as I am aware, Dewar (2019) was the first to explicitly define reduction. But this is not to say that the idea of reduction wasn't already around—indeed, Dewar describes it as the orthodoxy in the literature on symmetries. For example, Caulton (2015); Møller-Nielsen (2017) seem to have had a similar concept in mind. Likewise, the mathematical strategy of 'quotienting' as found in Belot (2003a) is intimately related to reduction.

[9] Dewar (2019) offers a different characterisation of reduction in terms of a distinction between syntax and semantics. In the Appendix to this chapter, I show that this distinction does not capture the content of these positions.

the symmetry-invariant content of the old theory". For example, the reductionist could identify inter-particle distances as a basis of invariant quantities and rewrite Newton's equations in terms of them.[10] The models of the new theory are then automatically invariant under the old theory's symmetries. Indeed, there is a sense in which the old theory's symmetries have simply vanished, since those symmetries act on the variant quantities that have been removed in the process of reduction. For example, if distances are all there is then there are no positions left to shift! The *formal* result of reduction is a theory whose models stand in a one-to-one correspondence with equivalence classes of SRMs of the old theory. In what follows, I will only consider reduced theories in this sense.

There is a further question, namely *how* to formulate a reduced theory. We can either formulate a theory *intrinsically* or *extrinsically.* On the former approach, we build its mathematical objects axiomatically 'from the ground up'. On the latter approach, we define the same mathematical objects in terms of the models of the old theory. For example, in the process of 'quotienting' one defines the space of models of the new theory as the quotient of the space of models of the old theory. In effect, this approach *defines* the models of the reduced theory as equivalence classes of the old theory. But this is unsatisfactory insofar as some of the old theory's terms do not refer, namely the variant quantities. For this reason, I prefer intrinsic approaches. However, this distinction is independent from the choice between reduction and sophistication, so it is of little relevance to the central issue of this dissertation (I discuss the analogous distinction for sophistication below and in Chapter 4).

Example of reduction

With this characterisation of reduction in mind, let me offer an example: the move from Newtonian Gravitation set on Galilean spacetime to *Geometrised Newtonian*

[10] This is the project carried out in Saunders (2013). For another attempt, see Dasgupta (2013).

Gravitation set on Newton-Cartan spacetime. Recall that models of Newtonian Gravitation are of the form $\langle M, t_{ab}, h^{ab}, \nabla, \varphi, \rho, \xi^a \rangle$. These models have a well-defined standard of acceleration but no objective notion of absolute velocity, concomitant with the theory's invariance under boosts. But there is still a sense in which Galilean spacetime has excess structure, since so-called 'dynamic shifts' are *also* symmetries of Newtonian Gravitation. In more detail, the dynamical equations (1.1)-(1.3) are invariant under the joint transformation $(\nabla, \varphi) \to (\nabla', \varphi')$, where $\nabla' = (\nabla, t_b t_c \nabla^a(\varphi' - \varphi))$ and $\nabla^a \nabla^b (\varphi' - \varphi) = 0$.[11] This transformation is not a symmetry of Galilean spacetime, since $\nabla \neq \nabla'$. Physically, a dynamical shift corresponds to an acceleration of all material bodies in the same direction, offset by a change in the gravitational potential. Since such transformations leave all distances the same, they relate empirically equivalent models.

This is where the reduction of Newtonian Gravitation to Geometrised Newtonian Gravitation comes in. In order to formulate Geometrised Newtonian Gravitation, one first defines a new, invariant quantity, $\tilde{\nabla} = (\nabla, t_{bc} \nabla^a \varphi)$. The models of the new theory are of the form $\langle M, t_{ab}, h^{ab}, \tilde{\nabla}, \rho, \xi^a \rangle$: all elements of these models are invariant under dynamic shifts. Secondly, one formulates a new set of laws in terms of $\tilde{\nabla}$:

$$R_{ab} = 4\pi \rho t_{ab} \tag{1.4}$$

$$R^{a\ c}_{\ b\ d} = R^{c\ a}_{\ d\ b} \tag{1.5}$$

$$R^a_{bcd} = 0 \tag{1.6}$$

$$\xi^b \tilde{\nabla}_b \xi^a = 0 \tag{1.7}$$

[11]The notation here derives from Malament (2012, Prop 1.7.3): $\tilde{\nabla} = (\nabla, C^a_{bc})$ iff for all smooth tensor fields $\alpha^{a_1 \ldots a_r}_{b_1 \ldots b_s}$,

$$(\tilde{\nabla}_m - \nabla_m)\alpha^{a_1 \ldots a_r}_{b_1 \ldots b_s} = \alpha^{a_1 \ldots a_r}_{n b_2 \ldots b_s} C^n_{m b_1} + \ldots + \alpha^{a_1 \ldots a_r}_{b_1 \ldots b_{s-1} n} C^n_{m b_1} - \alpha^{n a_2 \ldots a_r}_{b_1 \ldots b_s} C^{a_1}_{mn} - \ldots - \alpha^{a_1 \ldots a_{r-1} n}_{b_1 \ldots b_s} C^{a_r}_{mn}$$

In particular, note that (1.4) is the 'geometrised' version of the Newton-Poisson equation (1.2). Finally, one can prove that there exists a one-to-one correspondence between unique models of Geometrised Newtonian Gravitation and equivalence classes of models of Newtonian Gravitation related by dynamic shifts (Malament, 2012, Prop. 4.2.1 and 4.2.5). This gives us a sense in which the new theory captures all of the symmetry-invariant content of the old theory. Thereby, the reduction of Newtonian Gravitation to Geometrised Newtonian Gravitation is completed.

1.2.3 Sophistication

Reduction is the orthodox approach to the interpretation symmetries. Recently, Dewar (2019) has recently offered an alternative to reduction, called *sophistication*. In this section I explain the nature of sophistication. The remainder of this dissertation constitutes a sustained defence and development of that position.

For sophistication, the representational core of an equivalence class of SRMs is the set of symmetry-invariant *sentences*. In other words, models represent invariant *facts*—and nothing else. By construction, those sentences' truth values are constant within equivalence classes of SRMs, so such equivalence classes jointly represent the same state of affairs. Crucially, some of these sentences involve *variant* quantities. For example, the sentence "The Sun has a position in space" is invariant under translations, despite the fact that position itself varies under translations. Sentences such as these are part of the representational core for sophistication, but not for reduction. On the other hand, any sentence that solely concerns invariant quantities is itself invariant, so any sentence that is part of the representational core of reduction is also part of the representational core of sophistication. Therefore, sophistication has strictly more content than reduction (I will elaborate on this claim in Chapter 2).

As before, to simply state our commitment to the theory's invariant sentences is not enough: we also require a formalism that meshes with this interpretation. In the case of sophistication, this is a formalism in which the theory's SRMs are *isomorphic to each other*. In more detail, a sophisticated theory is such that (i) there is a one-to-one correspondence between models of the old theory and models of the sophisticated theory, and (ii) symmetry-related models of the sophisticated theory are isomorphic (if this second condition is already satisfied in the old theory, for instance in General Relativity, then that theory is 'born sophisticated'). The notion of isomorphism here is meant to capture a principle of structural equivalence, and varies from theory to theory. For example, in the context of spacetime symmetries the relevant notion of structural equivalence is that of *isometry*, while for classical gauge theories the relevant notion is that of a so-called *vertical bundle isomorphism* (Weatherall, 2018; Fletcher, 2020). The reason that sophistication demands isomorphisms is that the only sentences not invariant under isomorphisms are those that concern particular individuals. For instance, the claim that the Sun occupies *this* particular point varies under isomorphisms. Therefore, when symmetries are isomorphisms a rejection of facts that concern particular individuals, or *haecceitistic* facts, automatically entails a sophisticated semantics. The rejection of haecceitism thus functions to explain the physical equivalence of SRMs. We will see an example of this below.

As with reduction, one can formulate a sophisticated theory intrinsically or extrinsically. This corresponds to what Dewar (2019) calls 'internal' and 'external' sophistication. The latter approach is associated with the mathematics of category theory: one can simply stipulate isomorphisms between models by drawing additional arrows. As I mentioned above, the choice between intrinsic and extrinsic approaches is independent from the choice between reduction and sophistication. However, Dewar (2019) has recently defended an extrinsic approach to sophistication. In Chapter 5, I argue against his approach and in favour of intrinsic sophistication.

Example of sophistication

The shift from Newtonian Gravitation set on a Newtonian spacetime to the same theory set on a Galilean spacetime is an example of sophistication. Recall that models of Newtonian Gravitation set on Newtonian spacetime are of the form $\langle M, t_{ab}, h^{ab}, \nabla, \sigma^a, \varphi, \rho, \xi^a \rangle$. Here, σ^a is a time-like vector field that represents the identity of points across time and so entails a standard of rest. The following transformation is a symmetry of this theory (where d_* is the pushforward map of a diffeomorphism on M which encodes a kinematic shift):

$$\langle M, t_{ab}, h^{ab}, \nabla, \sigma^a, \varphi, \rho, \xi^a \rangle \to \langle M, t_{ab}, h^{ab}, \nabla, \sigma^a, d_*\varphi, d_*\rho, d_*\xi^a \rangle \tag{1.8}$$

The second model differs from the first by a constant boost of all matter content. Since a boost leaves all distances the same, boost-related models are empirically equivalent. Crucially, note that we can express the physical difference between these possibilities in terms of the theory's mathematical structures. Specifically, since σ is *not* invariant under boosts ($d_*\sigma^a \neq \sigma^a$), the models in (1.8) are not isomorphic. As a result, there are structural differences between them which correspond to facts about the absolute state of motion of material bodies.

The aim of sophistication is to modify the models of Newtonian Gravitation such that the transformation in (1.8) is an isomorphism. It does so by replacing Newtonian with Galilean spacetime, as defined above. The models of Newtonian Gravitation set on Galilean spacetime are simply those obtained by removing σ^a. The boost transformation then becomes:

$$\langle M, t_{ab}, h^{ab}, \nabla, \varphi, \rho, \xi^a \rangle \to \langle M, t_{ab}, h^{ab}, \nabla, d_*\varphi, d_*\rho, d_*\xi^a \rangle \tag{1.9}$$

Unlike (1.8), (1.9) *is* an isomorphism. This means that the models related by such transformations are structurally equivalent.

Does this mean that we are allowed to interpret SRMs as physically equivalent? Not yet, because the models in (1.9) are still *distinct*. Therefore, in addition to the formal requirement that SRMs are isomorphic, sophistication also has a metaphysical component: anti-haecceitism. As I will understand it, anti-haecceitism involves the denial of 'primitive identities': non-qualitative properties that apply to certain individuals in virtue of being *that* individual.[12] The rejection of primitive identities implies that there exist no qualitatively identical yet numerically distinct possible worlds. On this view, individuals are no more than 'nodes' in a web of relations. So, consider the putatively different boost-related trajectories. Those trajectories are qualitatively identical with respect to their relations to spacetime and to matter. From anti-haecceitism it follows that those trajectories are in fact identical. There is no *further* question of *which* points a particle occupies apart from their structural features. Therefore, anti-haecceitism implies that the transformation in (1.9) represents a 'distinction without a difference'.

1.3 Against Reduction

In this section I offer some preliminary reasons against reduction. These are not decisive reasons to favour sophistication, but they motivate the defence of the latter in the remainder of this dissertation.

Reduction is hard

There is a sense in which reduction is more difficult than sophistication. In particular, in order to successfully reduce a theory, one usually has to both find a new set of invariant quantities *and* reformulate the theory's laws in terms of those quantities.

[12]In fact, there are (at least) two distinct definitions of anti-haecceitism: a rejection of primitive identities, as above, and a rejection of *trans-world* identities. Møller-Nielsen (2017) adopts the former definition, which originates with Adams (1979), whereas Pooley (2006), following Lewis, adopts the latter. The exact difference between these two is an interesting question for further research.

In the case of sophistication, on the other hand, the quantities remain the same: one *only* has to restructure the theory's models. Put differently, sophistication only requires us to *remove* structure, whereas reduction often also requires us to construct *novel* structures. It is the latter task that seems especially difficult.

The claim that reduction is more difficult than sophistication is borne out by the history of modern physics. For example, consider electrodynamics. Formulated in terms of a one-form, A_a, the theory has distinct models related by gauge symmetries that seem to leave all physical facts the same. The invariant quantity of this theory is given by the Faraday tensor, F^{ab}, so one can try to formulate a reduced theory in terms of F^{ab}. As is well-known, this is easily done—but the resulting theory involves action-at-a-distance![13] An alternative approach involves a different set of quantities, namely so-called *holonomies* of A_a around closed paths. These holonomies act locally on fields and so avoid action-at-a-distance. However, as far as I am aware there exist no equations of motion that describe this interaction directly in terms of the holonomies. Instead, the action for this theory indispensably contains A_a. So, reduction of electrodynamics remains unsuccessful.[14] On the other hand, physicists know how to sophisticate the theory, namely by formulating it in terms of *fibre bundles.* In this mathematical framework, gauge symmetries are vertical bundle automorphisms. If these examples were to point towards a broader theme, then, it is that sophistication is often easier than reduction. (To further support this assertion, note that a generally-covariant theory such as General Relativity is *born* sophisticated. On the other hand, the only serious attempt at a *reduction* of General Relativity—Earman's (1989) 'Einstein algebras'—is universally seen as a failure, as seen from Rynasiewicz's (1992) criticisms.)

[13]For further discussion of this point, see §7.2.1.

[14]I discuss these matters in more detail in Chapter 7, where I will also argue that an attempt to reduce electrodynamics due to Wallace (2014b) is unsuccessful.

There isn't always a unique set of invariant quantities

The aim of reduction is to define a new theory in terms of some set of invariant quantities. However, there is no guarantee that there exists a *unique* set of invariant quantities in terms of which to formulate the reduced theory. Consider a Machian theory of gravity with the following law for a particle's escape velocity from a body with mass M:[15]

$$v_e = \sqrt{\frac{\gamma}{r \sum_k \frac{m_k}{m}}} \tag{1.10}$$

Uniform mass scalings are a symmetry of this theory (intuitively, the strength of the gravitational force in this theory depends on the sum of all masses in the universe, so that a uniform mass scaling simultaneously decreases the force of gravity). However, there are (at least) two distinct sets of invariant quantities. The first set consists of mass ratios $\frac{m_i}{m_j}$, distances r and a new quantity $v/\sqrt{m}$; the second consists of mass ratios $\frac{m_i}{m_j}$, velocities v and a new quantity m/r. But it seems that any choice between these distinct sets of fundamental quantities is essentially arbitrary: it is not at all obvious, for example, whether $v/\sqrt{m}$ is in some sense more 'natural' than m/r. Therefore, reduction may involve a radical form of ontological underdetermination, which seems at least as bad as the empirical underdetermination that literalist approaches face. We will see exactly this issue come up for Wallace's (2014a) deflationary account of the Aharonov-Bohm effect in the discussion in Chapter 7.

Reduced theories can incur an explanatory loss

Finally, reduced theories often seem to have less explanatory strength than their sophisticated counterparts. Dewar (2019) lists some examples of this phenomenon, but I will briefly illustrate the claim with an example due to Martens and Read

[15]This example is due to Niels Martens and James Read; I thank them for sharing a manuscript with me in which it is discussed.

(2020). Consider, again, uniform mass scalings of our Machian theory. As we have seen above, mass ratios are invariant under such scalings. Now, in the non-reduced theory—that is, the theory expressed in terms of intrinsic masses—mass ratios satisfy the equation

$$m_{ik} = m_{ij} \times m_{jk} \tag{1.11}$$

We can easily explain why this is so if we rewrite this equation in terms of intrinsic masses, since then $m_{ij} = m_i/m_j$ and the equation reduces to a mathematical theorem:

$$m_{ik} = \frac{m_i}{m_k} = \frac{m_i}{m_j}\frac{m_j}{m_k} = m_{ij} \times m_{jk} \tag{1.12}$$

However, the reductionist renounces intrinsic mass, which is a variant quantity of the theory. Therefore, this explanation is unavailable to her. Put differently: *if* mass ratios are fundamental, then presumably they could have had any value whatsoever. But if mass ratios are really ratios of intrinsic masses, then they are constrained in the way dictated by (1.11). The fact that (1.11) holds in the actual world is thus strong evidence that mass ratios are not fundamental.

In my opinion, this third objection is the most serious. I cover it in much more detail in Chapter 3, and again in the context of gauge theories in Chapter 7.

1.4 Conclusion

The aim of this dissertation is to offer a sustained defence and development of sophistication. Specifically, I will extend sophistication to internal symmetries of both the global and local variety. Here follows an overview of the dissertation.

In this chapter, I introduced the debate between sophistication and reduction. In Chapter 2, I further characterise the difference between these views in terms of two well-known principles: Leibniz Equivalence and the Invariance Principle. I will show that while both approaches adopt Leibniz Equivalence, sophistication rejects the Invariance Principle. This accounts for the sense in which sophisticated theories have more ontological content than reduced theories. In order to defend sophistication, I then refute three arguments in favour of the Invariance Principle. Chapter 3 develops the 'conspiracy' argument against reductionism. I show that under a reasonable set of assumptions, comparativist forms of reduction invariably involve 'cosmic conspiracies': unexplained patterns in the instantiation of the fundamental quantities, such as the transitivity of mass ratios discussed above. I believe this is a serious issue with reduction, and so motivates the consideration of sophistication in what follows. Having found reduction undesirable, I turn my attention to sophistication for the remainder of this dissertation. In Chapter 4, I consider Dewar's (2019) distinction between two forms of sophistication, namely 'internal' and 'external' sophistication. I defend the latter from Martens and Read's (2020) criticisms, but argue that it is ultimately unsatisfactory. Specifically, it fails to offer *intrinsic* formulations of theories. As a result, it remains opaque what an externally sophisticated theory is metaphysically committed to.

With the distinction between symmetry-first and structure-first sophistication in mind, I offer two detailed case studies of the latter. In preparation, Chapter 5 defines the central concepts of this dissertation more precisely: symmetries and quantities. Specifically, I formulate an extension of Earman's famous symmetry principles for internal symmetries. These symmetry principles will guide us in the case studies that follow. The first case study (Chapter 6) concerns the mass symmetries of Newtonian Gravitation. I argue that mass scalings *are* symmetries of that theory, but that this does not imply that mass relations are more fundamental than intrinsic masses. As an alternative to comparativism, I develop a version of sophistication about mass. The second case study (Chapter 7) concerns gauge symmetries. I argue that the appropriate setting for gauge theories is the fibre bundle formalism, and that in this setting the theory's symmetry-related models are isomorphic. This allows for a sophisticated interpretation of gauge theories, according to which symmetry-related models are physically equivalent. On such an interpretation, the Yang-Mills field is physically real and provides a local explanation of the Aharonov-Bohm effect.

In the concluding chapter, I take a step back and consider Rovelli's (2014) question: why do our theories possess (gauge) symmetries? I argue that the advocate of reduction cannot offer a satisfactory answer, since on that view symmetries are merely redundant structure. But the advocate of sophistication does have an answer: for her, symmetries are a consequence of the metaphysical thesis of anti-haecceitism. Far from being a symptom of 'descriptive fluff', a theory's symmetries in fact reveal the rich structure of the value spaces of the quantities it posits—and these sorts of facts provide genuine insight into what the world is made of.

1.A Syntax versus Semantics

In the introduction, a reduced theory was defined as a theory whose models stand in a one-to-one correspondence with equivalence classes of SRMs of the old theory, and a sophisticated theory as one whose symmetries are natural isomorphisms. Dewar (2019) offers an alternative characterisation: reduction modifies the *syntax* of a theory (i.e. its dynamical equations), while sophistication only modifies the *semantics* (i.e. its models). As Martens and Read (2020) point out, with 'semantics' Dewar does not mean the *mathematical* interpretation of the theory but rather the *logical* interpretation of its symbols in terms of Tarski-style models. However, Dewar's distinction is orthogonal to the one I have drawn: there are cases of reduction (as I have defined it) without a change in syntax, and sophistication (as I have defined it) *with* a change in syntax. Since the definitions above most naturally capture the *physical* content of reduction and sophistication—with which Dewar concurs—I will prefer them over Dewar's own characterisation.

On the one hand, it is possible to construct a reduced theory without a syntactical reformulation. In particular, it is at least conceptually possible that the reduced theory's equations have the same *form* as the old theory's. In that case, their symbols are merely *reinterpreted* as referring to some new, invariant quantities. It's easiest to see this with an (admittedly trivial) example. Suppose that a theory contains a quantity φ that assigns each particle a complex number. Furthermore, suppose that the dynamics of this theory only concern the magnitude $\rho := |\phi|$, i.e. are of the form:

$$f(|\varphi|) = 0 \tag{A.1}$$

Since the dynamics only 'care' about the absolute value of φ, any transformation of the form $\varphi \to \varphi e^{i\theta}$ is a symmetry of the theory. We may then try to formulate a reduced theory in terms of the invariant quantity ρ. But since the dynamical equations of the old theory were in any case only concerned with ρ, the dynamics of the reduced theory are of the form:

$$f(|\rho|) = 0 \tag{A.2}$$

Now, (1.4) and (1.5) are syntactically equivalent: the only difference is that they use distinct characters. Put differently, we can obtain a reduced theory from (1.4) by simply *reinterpreting* φ as a real scalar quantity. Of course, the situation here is highly unusual: it is a coincidence that the old laws were already fit for purpose.[16] But the example suffices to show that Dewar's gloss on reduction can come apart from ours.

On the other hand, sophistication sometimes includes a change in syntax, since it is possible that the theory's laws themselves involve some piece of structure that is removed in the process of sophistication. For instance, Martens and Read (2020, fn. 26) point out that one can write the laws of Newtonian Gravitation set on Newtonian spacetime in terms of the 'standard of rest' σ^a: the affine connection ∇^a is then defined in terms of σ^a. But since the move to Galilean spacetime removes σ^a from the theory's spacetime structure, one has to rewrite the theory's laws. Specifically, one can write the laws directly in terms of ∇, as we had already done in the above. As Martens and Read point out, this would mean that the difference between reduction and sophistication depends on a conventional choice of which objects to choose as basic. But then the distinction is unrelated to the issue of the physical interpretation of our theories. For this reason I will adopt the above definitions of reduction and sophistication in what follows. In sum (and in accordance with Dewar's account),

[16] Arguably, an example of such a lucky coincidence is Wallace's account of the AB effect. Wallace introduces two new invariant quantities, $D\theta = \nabla\theta - \mathbf{A}$ and $\rho = |\psi|$, and rewrites the laws of scalar electrodynamics in terms of these two quantities. However, it turns out that these laws are identical to the original laws, except that the symbols $\mathbf{A}$ and ψ have been replaced by $D\theta$ and ρ. In other words, Wallace's account of the AB effect requires no syntactic reformulation of the laws of scalar electrodynamics. Yet it is a form of reduction, since—because Wallace's account solely traffics in invariant quantities—each unique model of the new theory corresponds to an equivalence class of gauge-related models of the original theory.

the difference between reduction and sophistication is that the models of a reduced theory are *invariant* under symmetry transformations, whereas for a sophisticated theory symmetries act as *isomorphisms* of the models.

Chapter 2

Invariance or Equivalence: a Tale of Two Principles

As explained in the previous chapter, both reduction and sophistication hold that symmetry-related models represent the same state of affairs. This thesis is known as *Leibniz Equivalence*. With Leibniz Equivalence, there simply is no underdetermination: symmetry-related models are just different ways of describing the same possibility. In the case of shifts, this means that shift-related models really represent the same state of affairs: shifts are 'distinctions without a difference'. Reductionism in addition claims that only a theory's invariant quantities are physically real: this is called the *Invariance Principle*. The Invariance Principal codifies so-called 'symmetry-to-(un)reality inferences': inferences from the variance of some quantity to its non-reality.[1] For example, (absolute) positions vary under shifts, and hence the Invariance Principle rejects them as unphysical.

Leibniz Equivalence and the Invariance Principle are often seen as part of the same package (§2.2). This is a mistake: Leibniz Equivalence and the Invariance Principle are orthogonal. It is no surprise that this fact has gone unnoticed, for the

[1]For more on the symmetry-to-reality inference, see Dasgupta (2016); Read and Møller-Nielsen (2020).

principles *are* equivalent if one assumes that a quantity's values always represent the same magnitude. I understand 'values' here as mathematical entities that represent magnitudes, which are physical properties (§2.3). Call this claim the *Value-Magnitude Link*. If we accept the Value-Magnitude Link, Leibniz Equivalence does entail the Invariance Principle. We are then naturally led to reduction, which aims to formulate 'reduced' theories in terms of these invariant quantities. However, it is also possible to reject the Value-Magnitude Link. Indeed, a popular position in the philosophy of spacetime called *sophisticated substantivalism* does just that. If we reject this link, Leibniz Equivalence does not entail the Invariance Principle. This means that it is possible to take symmetry-related models to represent the same state of affairs whilst having a realist attitude towards variant quantities. In other words, these doctrines can come apart. Sophistication, of which sophisticated substantivalism is an instance, aims to fill this space.[2] In §2.4, I explain the nature of sophistication in terms of a method of supervaluation over symmetry-related states.[3]

This means that those who accept Leibniz Equivalence as a solution to the problems of underdetermination and indeterminism face a choice: reject the Value-Magnitude Link, or accept it and hence commit to the Invariance Principle? Various arguments have been presented in favour of the Invariance Principle: a rejection of the principle is supposed to cause indeterminism (Earman and Norton, 1987), undetectability (Dasgupta, 2016) and failure of unique reference (Healey, 2009; Caulton, 2015). As I will argue in §2.5, however, those arguments at best support Leibniz Equivalence. The inference to the Invariance Principle is unwarranted since this additionally requires that one assumes the Value-Magnitude Link, which proponents of sophistication re-

[2]For more on the reduction/sophistication distinction, see Dewar (2019); Martens and Read (2020).

[3]This naturally raises the question: is there also a coherent position that denies Leibniz Equivalence yet affirms Invariance Principle? It has been suggested to me that Belot's (2011) 'modal relationism' might fulfil this role. As a relationist, Belot accepts the Invariance Principle: only the invariant distances are physically real. But Belot's relational facts are endowed with rich modal structures that allow for *more* possibilities than standard relationism, *contra* the collapse of possibilities entailed by Leibniz Equivalence.

ject. Therefore, these arguments cannot offer support for reductionism, contrary to the intentions of their authors.

2.1 Two Principles

In this section I discuss Leibniz Equivalence and the Invariance Principle in turn. I will adopt a loose definition of symmetries as transformations that in some sense 'preserve the dynamics'. In order to explain this notion, it is useful to distinguish between two types of models. The *kinematically possible models* (KPMs) of a theory are those that are of the correct 'form', or contain the right sort of mathematical objects. The *dynamically possible models* (DPMs) are those KPMs that in addition satisfy a set of equations of motion. The DPMs represent ways the world could be if the theory were true. Symmetries are maps from the space of KPMs onto itself that preserve the space of DPMs. Since Belot (2013), it has been well-known that not all symmetries of this type relate empirically—let alone physically—equivalent states. For this reason I restrict discussion to just those symmetries that in addition satisfy a condition of empirical equivalence. The main symmetries of philosophical interest, from 'shifts' in Newtonian Mechanics and diffeomorphisms in General Relativity to the gauge symmetries of electrodynamics, satisfy this requirement.

2.1.1 Leibniz Equivalence

Leibniz Equivalence was first suggested by Earman and Norton (1987) as a response to the Hole Argument, which states that General Relativity contains pairs of models related by diffeomorphisms that agree on all physical fact up until some time t but diverge thereafter. Because of this fact, it seems that substantivalist interpretations of General Relativity are indeterministic. Earman and Norton (1987, 522) phrase Leibniz Equivalence as the claim that *diffeomorphic models represent the same physical* *situation*. With this assumption, General Relativity is *not* indeterministic, since the symmetry-related models that seem to represent diverging futures are in fact merely distinct representations of the same future. Since then, the term has been used more broadly to refer to the view that *any* symmetry (subject to the proviso of empirical equivalence) relates physically equivalent models, whether or not the symmetry in question is a diffeomorphism. I will understand Leibniz Equivalence in this broader sense:

Leibniz Equivalence: Symmetry-related models are physically equivalent.

The aforementioned shifts are a paradigmatic example: shifts preserve all distances between bodies and so relate empirically equivalent states of affairs. From Leibniz Equivalence, it follows that shift-related models are also physically equivalent: they merely represent the same state of affairs in different ways. Proponents of Leibniz Equivalence include Saunders (2003b), Baker (2010), Weatherall (2018) and Greaves and Wallace (2014). The latter write:

> There is a widespread consensus that 'two states of affairs related by a symmetry transformation are really just the same state of affairs differently described'. That is, if two mathematical models of a physical theory are related by a symmetry transformation, then those models represent one and the same physical state of affairs. (60)

This is not to say that there is no dissent; see for instance Roberts (2020). But it seems clear that Leibniz Equivalence is the orthodox view with respect to symmetry-related models.

Rynasiewicz (1994) calls the opposite of Leibniz Equivalence *model literalism*. According to model literalism (or literalism, for short), distinct models represent distinct states of affairs, whether or not they are symmetry-related. There is a one-to-one relation between models and possibilities. As we will see below (§2.5), in the presence of non-trivial symmetries literalism implies exactly those vices that Leibniz Equivalence is supposed to solve. I will argue that the acceptance of Leibniz Equivalence, and hence the denial of literalism, suffices to avoid these: the further posit of the Invariance Principle is unnecessary.[4]

2.1.2 The Invariance Principle

The *Invariance Principle* is the claim that "only quantities that are invariant under the symmetries of our theories are physically real" (Møller-Nielsen, 2017, 1253). In other words, only those quantities that have the same value across equivalence classes of symmetry-related states represent a feature of physical reality:

> *Invariance Principle.* A quantity is physically real only if it is invariant under the symmetries of our theories.

Examples of variant quantities include absolute position under shifts, but also the values of the electrostatic potential in electromagnetism under gauge transformations, and (arguably) intrinsic mass under mass scalings.[5] Examples of invariant quantities are distances and relative velocities, so-called 'loop holonomies' of the potential field, and mass ratios. The Invariance Principle suggests that the latter set of quantities is more fundamental than the former.

The Invariance Principle has gathered assent from a broad range of physicists and philosophers throughout the 20th and 21st century. Here is a collection of representative quotes:

> "Physically real quantities are invariant under exact symmetries—this is the general lesson." (Saunders, 2003b, 300)

> "[O]nce we possess the covariant representation under which the equations stay the same for all coordinate systems, the quantities in the (covariant) equations are the real and objective quantities." (Nozick, 2001, 82)

[4]Literalism and LE are not mutually exhaustive, since one may also hold that some but not all SRMs represent distinct state of affairs; this is the view of Belot (2018).

[5]The latter issue is controversial: see Dasgupta (2013) for the case for comparativism, and Baker (2014); Martens (2019a) for responses. I discuss mass symmetries in detail in Chapter 6.

> "[T]he important things in the world appear as the invariants [...] of these transformations." (Dirac, 1930, vii)

Endorsements of the Invariance Principle can also be found in Saunders (2007), Baker (2010), Dewar (2015), Caulton (2015) and Dasgupta (2016), among others. The Invariance Principle is perhaps even more orthodox than Leibniz Equivalence.

The Invariance Principle encodes the idea of so-called *symmetry-to-reality inferences*: the inference from the variance of a quantity under the theory's symmetries to its physical unreality. Dasgupta (2016) justifies such inferences on the basis that variant quantities are undetectable. It then follows from an application of Occam's razor that theories that excise these quantities are preferable, all else being equal. But as I will argue below (§2.5), realism about such quantities does not entail empirical underdetermination. Therefore, Occam's razor does not justify the Invariance Principle, *pace* Dasgupta.

2.2 One Principle?

Leibniz Equivalence and the Invariance Principle are closely related. Nevertheless the two principles *are* distinct, a fact which has often gone unnoticed.[6] For example, Saunders (2007) equivocates between them in the following passage:

> [Relationalism] says that *only quantities invariant under exact symmetries are real*—thus relative directions, relative distances, and so on, under rotations and translations, etc. [...]. Call [this principle] the invariance principle. The distinctions among representations then correspond to nothing physically real. *Equivalently, such models, as goes their physical content, can be simply identified.* (453, emphases mine)

The first italicised part of this passage is a canonical statement of the Invariance Principle. But towards the end of the quote Saunders equates this with Leibniz

[6] Read and Møller-Nielsen (2020, §6) are an exception: they rightly point out that these principles are not logically equivalent.

Equivalence, the claim that symmetry-related models are physically equivalent. Similarly, Baker (2010) writes:

> [I]f changes in surplus structure are generally (as in geometry) mere descriptive changes, it follows that physical situations related by symmetries must be qualitatively identical [Leibniz Equivalence]. And if this is right, then physical quantities that change under symmetry transformations (i.e. that are not invariant) must not be fundamental quantities [Invariance Principle]. (1158)

In this passage, Baker moves freely from Leibniz Equivalence to the Invariance Principle.

As far as I am aware, however, there is no explicit argument for the claim that Leibniz Equivalence implies the Invariance Principle to be found in the literature. It seems plausible that the following line of reasoning is implicitly assumed. Let $\mathfrak{m}$ and $\mathfrak{m}'$ denote symmetry-related models of a theory T, such that there is a variant quantity Q which takes on different values across these models. For instance, $\mathfrak{m}$ and $\mathfrak{m}'$ could denote shift-related models of Newtonian Gravitation, with Q the (variant) position of the Earth. Then:

(1) The models $\mathfrak{m}$ and $\mathfrak{m}'$ represent the same physical possibility (Leibniz Equivalence);

(2) Distinct models represent the same physical possibility iff they agree on all physical features;

(3) Therefore, $\mathfrak{m}$ and $\mathfrak{m}'$ agree on all physical features (from (1) and (2));

(4) The models $\mathfrak{m}$ and $\mathfrak{m}'$ disagree on the value of Q;

(5) Therefore, if Q is physically real then $\mathfrak{m}$ and $\mathfrak{m}'$ disagree on some physical feature (from (4));

(C) Conclusion: Q is not physically real (Invariance Principle, from (3) and (5)).

This argument seems valid, but in fact it is not: (5) does not follow from (3) and (4). The reason is that a quantity's values are *representational devices*, rather than physical features. The fact that two models disagree on those values does not imply that they disagree on any physical fact. It is possible that the values of Q represent *different* physical properties (or *magnitudes*) across models. In particular, it is possible that Q's distinct values in $\mathfrak{m}$ and $\mathfrak{m}'$ denote the *same* property. In that case the two models are in agreement on all physical features after all, so the Invariance Principle does not follow.

In order to turn the above into a valid argument, an assumption that connects a quantity's values with the magnitudes they are supposed to represent is required. I will discuss such an assumption in the next section.

2.3 The Value-Magnitude Link

The above argument is valid *if* we assume the following principle, which I call the *Value-Magnitude Link*:

> *Value-Magnitude Link.* The values of a quantity invariably represent the same magnitude across models.

I am not aware of any explicit statement of the Value-Magnitude Link in the literature, let alone a defence. But in light of the quotes above, I contend that the Value-Magnitude Link is a widely-shared implicit assumption in the symmetries literature.

It is important to be clear on what is meant by *values* here. In the first instance, it is tempting to think of values as *numbers*. For example, the mass value of a 5 kg object is '5'. But on this definition the Value-Magnitude Link is obviously false, since the same magnitude may be represented by different numbers in different unit systems. For example, we can also represent '5 kg' as '5000 g'. The Value-Magnitude Link clearly fails for *passive transformations*. When I write 'value', then, I don't mean the numerical value of a quantity in a certain system of units. Rather, I mean a mathematical element of the theory's models: a particular point of a differentiable manifold, or an element of an internal 'value space', i.e. the values *come with* a dimension/unit. For example, a model of Newtonian Gravitation consists of an assignment of massive particle trajectories to a manifold M, over which some spacetime structure is defined. We can endow the manifold M with coordinates: a passive transformation then is simply a change in these coordinates. But we can also leave the coordinates fixed, and assign the particle trajectories to different points of the manifold itself. This is an *active transformation.*[7] The values in this case are the point of the manifold M, and the Value-Magnitude Link asserts that the same point of M represents the same point of spacetime in all of the theory's models. We can present a similar set-up for quantities such as mass. Here, quantities are functions from the theory's domain—say, particles—into a *value space*: a mathematical structure whose elements represent determinate magnitudes of said quantity. For instance, mass is a function from particles into a mass value space, whose elements represent determinate mass magnitudes such as '5 kg'. As with spacetime, we can 'coordinatise' a value space via an assignment of numbers to its elements. This amounts to a choice of unit. Alternatively, we can leave the units fixed and consider a different mapping from the domain into value space. For instance, an (active) mass scaling maps each particle onto a different element of mass value space. In this case, values are simply elements of a value space, and the Value-Magnitude Link asserts that these elements represent the same magnitudes across the theory's models.[8]

Let's consider the Value-Magnitude Link in more detail. In the previous section, I argued that Leibniz Equivalence implies the Invariance Principle *if* we assume the

[7]Note that on this definition an active transformation need not correspond to any *physical* transformation: if shift-related models of Newtonian Gravitation are physically equivalent, then an active transformation relates models that represent the same state of affairs.

[8]Chapter 5 further elaborates on the relation between value spaces and their numerical representation.

Value-Magnitude Link. In fact, these two principles are *equivalent* on that assumption, as I will now show.

Proof. Left to right. Consider a variant quantity Q. Suppose that in some model $\mathfrak{m}_1$, the values of Q for x and y are $Q(x) = a$ and $Q(y) = b$. Now, consider a symmetry-related model $\mathfrak{m}_2$ in which $Q(x) = b$ and $Q(y) = c$.[9] By Leibniz Equivalence, $\mathfrak{m}_1$ and $\mathfrak{m}_2$ are physically equivalent, so $Q(x)$ and $Q(y)$ each represent the same magnitude in both models. But since $Q(y) = b$ in $\mathfrak{m}_1$ while $Q(x) = b$ in $\mathfrak{m}_2$, this means that the value b represents *distinct* magnitudes across those models. This contradicts the Value-Magnitude Link, so Q cannot represent a physically real quantity. Since Q is an arbitrary variant quantity, it follows that only invariant quantities are physically real.

For an example, consider again the static shift. Since static shifts are symmetries of Newtonian Gravitation, Leibniz Equivalence implies that shift-related models are physically equivalent. In that case, the Invariance Principle implies that position, which varies under shifts, is *not* a physically real quantity. Consider the following two shift-related models, where x_A and x_B denote the locations of particles A and B respectively:

$$\mathfrak{M}_1 : \begin{cases} x_A = (1, 0, 0) \\ x_B = (2, 0, 0) \end{cases} \qquad \mathfrak{M}_2 : \begin{cases} x_A = (2, 0, 0) \\ x_B = (3, 0, 0) \end{cases}$$

If these models are physically equivalent, then $(1, 0, 0)$ and $(2, 0, 0)$ both represent the same spacetime point: A's location. Similarly, $(2, 0, 0)$ and $(3, 0, 0)$ both represent B's location. This means that the same value, $(2, 0, 0)$, represents B's location in $\mathfrak{m}_1$, and A's location in $\mathfrak{m}_2$. But both models represent those bodies as occupying *distinct* locations. It follows that the same values of x_A and x_B must denote different spatial

[9]This relies on the assumption that there exists some symmetry-related model in which $Q(x) = b$; see the shift example below.

locations across models. This contradicts the Value-Magnitude Link, and so variant location is not physically real.

Right to left. Consider any pair of symmetry-related models. By definition, the invariant quantities have the same values across these models. The Invariance Principle implies that only those quantities are physically real. From the Value-Magnitude Link, it follows that their values denote the same magnitudes across models. Therefore, these models agree on all physical features. Symmetry-related models are thus physically equivalent, and so Leibniz Equivalence holds. For example, consider once more two shift-related models. From Invariance Principle, it follows that only the invariant quantities—distances, relative velocities and accelerations—are physically real. By definition, those quantities have the same value in both models. It follows from the Value-Magnitude Link that those values represent the same magnitude. Therefore, both models represent the same state of affairs. □

On the assumption of the Value-Magnitude Link, Leibniz Equivalence and the Invariance Principle are equivalent. If that is the case, then "symmetry-to-reality" inferences have a clear structure: symmetry-related models are physically equivalent, and hence variant quantities are not physically real. This naturally leads to the view that we ought to formulate new theories that are solely expressed in terms of the invariant quantities, so-called reduced theories. But if one were to insist that variant quantities *are* physically real, then under the assumption of the Value-Magnitude Link this entails the denial of Leibniz Equivalence, or literalism. On the assumption of the Value-Magnitude Link, the interpretation of symmetry-related models leads to a dilemma: either variant quantities are physically real and symmetry-related models represent distinct states of affairs (literalism), or only invariant quantities are real and so symmetry-related models are physically equivalent. In the latter case, a new theoretical formalism is appropriate (reduction). There is no room for an intermediate view. But as we will see in the next section, there is no need to accept the Value-

Magnitude Link.

2.4 Conventional Values and Supervaluation

The Value-Magnitude Link seems unassailable, but is denied by one of the most common views on symmetries: sophistication. In this section, I will first illustrate sophistication with a couple of examples: sophisticated spacetime substantivalism, and sophistication for the electrostatic potential. The fact that the latter example concerns an internal symmetry shows that the failure of the Value-Magnitude Link is not merely a particularity of spacetime. I will then undertake a more general discussion of the relation between sophistication, reduction and literalism.

The most well-known example of sophistication is *sophisticated substantivalism*.[10] According to sophisticated substantivalism, spacetime is a real substance in which bodies are located; hence, there is a sense in which (variant) positions are physically real.[11] Contrast this with relationism, which says that only distances between bodies are real. Despite the fact that sophisticated substantivalism is realist about position, it affirms the physical equivalence of symmetry-related models. It does so via an appeal to *anti-haecceitism* regarding spacetime points. Rather than possessing primitive identities, spacetime points are individuated via their qualitative relations to each other and to the universe's matter content. Since spacetime symmetries relate qualitatively identical models, symmetry-related models represent the exact same points in the same configurations. As a result, such models have identical physical content.

Sophisticated substantivalism thus accepts Leibniz Equivalence and rejects the

[10]See Pooley (2013, §7) for an overview.

[11]The claim that sophisticated substantivalism is realist about (variant) position assumes a particular form of that view, namely what Saunders (2003a) calls *non-reductive relationism*. On the other hand, an *eliminative relationist* may understand positions as non-fundamental quantities. To confuse matters, both types of view have sometimes been called 'structuralism'; cf. Greaves (2011). I will eschew the latter term in what follows.

Invariance Principle. It follows from the previous section that sophistication must also reject the Value-Magnitude Link. This is indeed the case. Here is a simple example. Suppose as before that models related by shifts are physically equivalent. The following three diagrams then represent the same state of affairs:

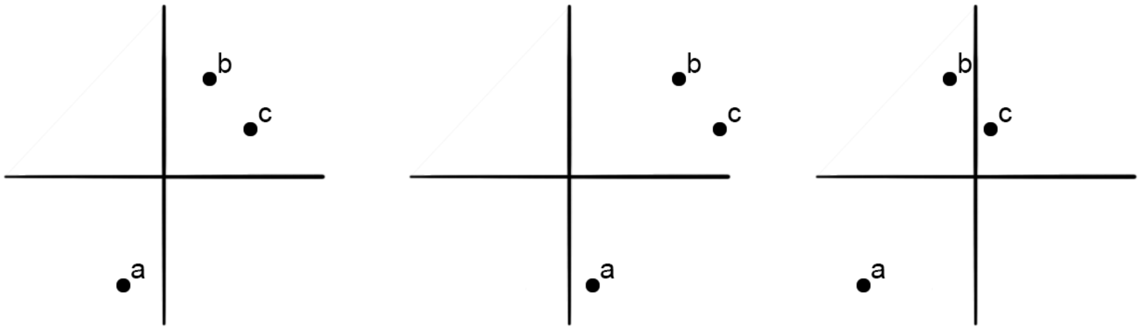

Figure 2.1: Examples of shift-related models.

Here, the dots represent the locations of three bodies labeled 'a', 'b' and 'c' in two-dimensional position space. The location value of the dots is different in each diagram (the dots are shifted on the x-axis). On the assumption of the Value-Magnitude Link this means that spatial location itself is not a physically real quantity, since it varies across the theory's shift-symmetries. According to sophistication, however, there *are* physically meaningful statements about spatial position. Consider the claim that particle b has *some* location in space (i.e. b is assigned *some* value in two-dimensional position space). This statement is invariant under shifts: b is assigned a location on all three diagrams in Fig. 2.1. Robert Stalnaker has summarised the point as follows: "in the relational theory of space, spelled out in this way, no intrinsic spatial properties would be real (*except the property of being located somewhere in space*)" (Stalnaker, 1979, 353-4, emphasis mine). This is a direct consequence of the sophisticated substantivalist's commitment to space as a real substance in which particles are located.

Sophistication is not just applicable to spacetime symmetries. Dewar (2019) has recently argued that sophistication can be seen as a general approach to the interpretation of symmetries. He mentions the example of the electrostatic potential, ϕ. The empirical consequences of ϕ are due to its gradient $E := \nabla\phi$. Therefore, uniform shifts of ϕ are symmetries of electrostatics. According to Leibniz Equivalence, models related by shifts in the potential are physically equivalent. It makes no difference if we increase ϕ everywhere by the same amount. The Invariance Principle in turn implies that the electrostatic potential is not physically real, since it is variant under the symmetries of electrostatics. Instead, the electric field E, which *is* invariant, represents the theory's fundamental field. This motivates the demand for a reduced theory formulated in terms of E alone. But as Dewar points out, such a reduced theory cannot explain the fact that E is a *conservative* force, i.e. that $\nabla \times E = 0$.[12] On the other hand, if E is *defined* as the gradient of ϕ then this conservativeness rolls out as a mathematical theorem, since $\nabla \times (\nabla\phi) = 0$ for any ϕ. For this reason, Dewar argues that it is better to affirm the physical reality of ϕ, contrary to the Invariance Principle. This implies a rejection of the Value-Magnitude Link. For consider two symmetry-related models, one in which $\phi(x) = a$ and one in which $\phi(x) = b$. When $a \neq b$, the value of ϕ at x differs across these models. But since symmetry-related models are physically equivalent, both models must represent the same physical *magnitude* of the electrostatic potential at x. Hence, it is not the case that values invariably represent the same magnitude across models. Dewar's sophistication achieves this balance via an appeal to *anti-quidditism*, the analogue of anti-haecceitism for properties: physical properties are individuated via their position in a structure of qualitative relations. Since the distinct values a and b occupy the same structural role across symmetry-related models, they also represent the same magnitude.[13]

These examples illustrate how sophistication steers a course between equivalence

[12] On this point, see also Roberts (2016).

[13] For more on anti-quidditism in the context of symmetries, see Martens and Read (2020).

and invariance. But in order to obtain a more general characterisation of the difference between sophistication and reduction, it is necessary to provide a more specific criterion for the physical content of a sophisticated theory. What we need is a statement of the *semantics* of a sophisticated theory.[14] I propose the following criterion:

> *Sophistication.* A theoretical statement is *physically meaningful* if and only if it is invariant across physically equivalent models.[15]

From left to right, this says that a statement is physically meaningful only if it is invariant across physically equivalent models. This is uncontroversial: physically equivalent models agree on all physical facts, so whatever varies across such models cannot have a consistent meaningful interpretation. But read from right to left, this claim is non-trivial: it states that *any* invariant statement is physically meaningful, *even if* such a statement involves variant quantities. In other words, sophistication locates physical meaning in the full sentence, rather than in its terms. For an example, borrowed from van Fraassen (1989, 284), consider the statement "total momentum is conserved". In Newtonian Gravitation, total momentum is a variant quantity: it varies under boosts of our reference frame. But it is true in any frame that the total momentum of all particles is conserved, and hence the statement "total momentum is conserved" is still physically meaningful under a sophisticated semantics.

It is possible to understand this semantic criterion in terms of the logic of supervaluation. This will provide us with another, more formal way of understanding the space between literalism and reduction. In formal semantics, the idea of supervaluation is to consider the truth-value of a sentence with respect to a class of interpretations, rather than within any particular interpretation. In the present context, the equivalent idea is to evaluate the truth-value of a sentence of the theory with

[14] This is different from the *formalism* of a sophisticated theory, which is what Dewar (2019) is chiefly concerned with. As Dewar argues, sophistication requires that a theory's symmetry-related models are *isomorphic*. On the 'internal' approach to sophistication, this usually requires a reformulation of the theory's models.

[15] Cf. Mundy (1986). See also Barrett's (2017) proof that any piece of symmetry-invariant structure is *implicitly definable* from the theory's models.

respect to an equivalence class of symmetry-related models, rather than within any particular model. We call a statement *supertrue* (respectively superfalse) iff it is true (respectively false) in all models within a class of symmetry-related models.[16] Importantly, *supervaluation does not commute with logical composition.* The truth-value of a composite statement under supervaluation is not a function of the truth-values of its components under supervaluation. We can use this fact to explain the difference between reduction and sophistication: the former *first* supervaluates over 'atomic' theoretical statements and then composes complex ones, whereas the latter first composes complex sentences and *then* supervaluates. Let $[\phi(\mathfrak{m})]_{\mathcal{S}}$ denote the truth value of a sentence ϕ in a model $\mathfrak{m}$ when we supervaluate it with respect to a group of symmetries $\mathcal{S}$ (compare the role of *interpretations* in logic). The truth conditions of $[\phi(\mathfrak{m})]_{\mathcal{S}}$ are:

$$[\phi(\mathfrak{m})]_{\mathcal{S}} = \begin{cases} T \text{ iff } \phi(s(\mathfrak{m})) = T \text{ for all } s \in \mathcal{S} \\ F \text{ iff } \phi(s(\mathfrak{m})) = F \text{ for all } s \in \mathcal{S} \\ \# \text{ otherwise} \end{cases} \tag{2.1}$$

where $s(\mathfrak{m})$ denotes the model obtained from an application of s to $\mathfrak{m}$. In simple terms, this says that ϕ is true (false) in $\mathfrak{m}$ iff it is true (false) in all models related to $\mathfrak{m}$ by one of the symmetry transformations in $\mathcal{S}$. If neither situation obtains—if ϕ is true in some model, but false in another, symmetry-related one—then the statement is meaningless. For instance, the claim that a certain body is at absolute rest is true in some models of Newtonian Gravitation, but false in any model related to the first by a uniform boost. On a supervaluationist semantics it is physically meaningless. But the claim that total momentum is conserved is true in all models of the theory, so it comes out as true under the supervaluationist semantics.

[16] Cf. Dewar's (2019) supervaluationist semantics for sophistication. See also Russell's (2018) discussion of determinacy.

For an example of a statement that is meaningful (indeed, true) for sophistication but meaningless for reduction, consider the claim that a quantity Q has *some* value in $\mathfrak{m}$: $\exists x(Q(\mathfrak{m}) = x)$. For example, Q can stand for the Earth's position, so the claim is that the Earth is located somewhere in space. We can write this claim as a disjunction:[17]

$$Q(\mathfrak{m}) = x_1 \vee ... \vee Q(\mathfrak{m}) = x_n \tag{2.2}$$

The crucial insight here is that there is a difference in truth value between the supervaluation of this sentence as a whole, and the composition of its individually supervaluated disjuncts. In other words,

$$[Q(\mathfrak{m}) = x_1 \vee ... \vee Q(\mathfrak{m}) = x_n]_{\mathcal{S}} \neq [Q(\mathfrak{m}) = x_1]_{\mathcal{S}} \vee ... \vee [Q(\mathfrak{m}) = x_n]_{\mathcal{S}} \tag{2.3}$$

Proof. Since Q varies under $\mathcal{S}$, for any x the claim $[Q(\mathfrak{m}) = x]_{\mathcal{S}}$ is neither true nor false. Under standard trivalent semantics (such as Kleene), the disjunction of sentences with truth-value # itself has truth-value #. Therefore, the rhs of (2.3) is neither true nor false.[18] On the other hand, $[Q(\mathfrak{m}) = x_1 \vee ... \vee Q(\mathfrak{m}) = x_n]_{\mathcal{S}}$ *does* have a truth-value, since in any model Q has *some* value between x_1 and x_n. Therefore, the lhs of (2.3) is true. This completes the proof. □

Because of this fact, sophistication is different in terms of its physical content from both literalism and reduction. Using the example of the Earth's location, the difference is as follows. Reduction says that no statement of the form "The Earth is located at position x" is meaningful, since no such statement has the same truth-value within a class of shifted models. Therefore, the reductionist argues, *composites* of such

[17] I will gloss over any issues arising from the fact that for most quantities, this disjunction is infinitary

[18] Note that the symbol $\vee$ on the rhs of (2.3) really functions as a binary operator on truth-values, rather than as a sentential connective. The standard rules of Kleene semantics therefore strictly don't apply, and so the idea here is rather to employ analogous rules for this operator.

statements are also meaningless. The sentence "The Earth has *some* location"—which we have interpreted as a disjunction of ascriptions of particular locations—has no physical relevance. But sophistication evaluates the truth-value of sentences *as a whole*. And as we have seen, the sentence "The earth has some location" *is* true across shift-related models. The semantics of supervaluation offers a formal way of understanding the sense in which sophistication has *more* physical content than reduction.

Let me briefly say something about the benefits of this additional content; I will return to this matter in more detail in the next chapter. The rough idea is that the theory's variant quantities contribute towards its *explanatory strength*. Consider the fact that the electric force is conservative, i.e. that $\nabla \times E = 0$. On a sophisticated semantics, the definition $E \coloneqq \nabla\phi$ is invariant across symmetry-related models, hence physically meaningful. Therefore, we can use it to *explain* the conservativeness of E as a result of the fact that $\nabla\times\nabla\phi = 0$ for any scalar field ϕ. On a reductionist semantics, on the other hand, *any* statement that involves ϕ is physically meaningless as a result of the fact that ϕ itself varies under the theory's symmetries. Therefore, reductionism cannot appeal to ϕ to explain the conservativeness of E. For reductionism it is a brute fact that E is conservative. Generally, what this means is that sophistication but not reduction allows for explanations of physical phenomena in terms of variant quantities. In fact, such explanations abound in physics. For an example, we can explain the Twin Paradox in Special Relativity in terms of the phenomenon of time dilation.[19] But this is a frame-dependent (hence symmetry-variant) phenomenon, and so reductionism accords it less reality than sophistication. Here, too, it seems that sophistication increases the explanatory strength of the theory. I'll briefly return to these ideas in the conclusion.

[19] See, for instance, Rindler (1977, §2.14). But note that Debs and Redhead (1996) argue that such explanations are invalid exactly because they appeal to frame-dependent facts.

2.5 Against Invariance

If one accepts the Value-Magnitude Link, one faces a stark binary choice between literalism and reduction. On the first view, symmetry-related models represent distinct possibilities: the consequence is underdetermination and possibly indeterminism. On the second view, symmetry-related models are physically equivalent. With the Value-Magnitude Link in place, this implies the Invariance Principle: only invariant quantities are physically real. This principle motivates the search for a reduced theory formulated solely in terms of these invariant quantities. Clearly, reduction is preferable over literalism. But if we reject the Value-Magnitude Link a third option becomes available: sophistication. Sophistication accepts Leibniz Equivalence, but not the Invariance Principle. As I illustrated above, this means that it remains committed to the physical reality of variant quantities. The possibility of sophistication shows that Leibniz Equivalence in itself cannot motivate the Invariance Principle.

But there are many other arguments in the literature in favour of the Invariance Principle. Earman and Norton (1987) argue that the Invariance Principle is necessary to avoid indeterminism; Dasgupta (2016) claims that variant quantities are undetectable; and according to Caulton (2015), such quantities fail to refer uniquely. These arguments, if successful, *a fortiori* support reduction over sophistication. But I will argue that they fail. Specifically, I claim that all three arguments only support Leibniz Equivalence, rather than the Invariance Principle. It is only when one tacitly assumes the Value-Magnitude Link that the latter follows. But that assumption begs the question against sophistication, which explicitly rejects this link. Therefore, I conclude that these purported arguments for the Invariance Principle cannot support reduction over sophistication, contrary to their author's explicit claims.

I will now discuss the arguments in favour of the Invariance Principle one by one.

2.5.1 Determinism

The claim that symmetries lead to indeterminism of our theories is well-known. Earman and Norton (1987) argue that this happens in General Relativity due to the Hole Argument, and Belot (1998) applies a similar line of reasoning to electromagnetism. Wallace (2002) generalises the argument to any time-dependent local symmetry (for our purposes, a local symmetry is one that can act non-trivially over some finite period of time). The thrust of these arguments is that there are distinct models that agree on all physical facts before some time t, but which differ by a non-trivial symmetry transformation after t. If these models represent distinct physical possibilities, indeterminism fails: the history of the universe up to time t does not determine its state at later times.

Earman and Norton (1987) argue that we should avoid this sort of radical indeterminism not because indeterminism is in principle unacceptable, but because "[determinism] should fail for a reason of physics, not because of a commitment to substantival properties which can be eradicated without affecting the empirical consequences of the theory" (524). In a spacetime context, the 'substantival properties' to which Earman and Norton refer here are variant positions in spacetime. They claim that it is necessary to accept Leibniz Equivalence in order to avoid indeterminism, and that therefore the fundamental spatiotemporal quantities are relational rather than intrinsic. In other words, Earman and Norton advocate a form of reduction. Indeed, Earman (1989) tried to formulate such a reduced theory in terms of Einstein algebras, but Rynasiewicz (1992) showed that this was unsuccessful.

I agree with Earman and Norton that Leibniz Equivalence is necessary to avoid this sort of indeterminism. But it does not follow that we have to 'eradicate' substantival properties, or variant quantities more generally. This inference is only valid on the assumption of the Value-Magnitude Link. On the other hand, if the Value-Magnitude Link is denied then Leibniz Equivalence implies that we can remain realist about variant properties. In the case of spacetime, for example, sophisticated substantivalism accepts that symmetry-related models are physically equivalent, but nevertheless gives a realist interpretation of spacetime. Moreover, Dewar (2019) suggests that one can use the same strategy for non-spatiotemporal, or 'internal' symmetries, such as the gauge symmetry of electrodynamics. There, too, we can simply interpret gauge-related models as distinct representations of the *same* physical fields, rather than representations of distinct possible evolutions of those fields for a given initial condition. This suffices to avoid indeterminism, since the symmetry-related models that were supposed to represent distinct possible futures now considered as distinct representations of the same future. In Chapter 7 I will consider the case of gauge symmetries in more detail. The key point is that we need not accept the Invariance Principle in order to avoid radical indeterminism.

2.5.2 Detectability

The second argument for the Invariance Principle is epistemic. According to Dasgupta (2016), we should renounce variant quantities because they are in principle *undetectable*. On his account, the symmetry-to-reality inference goes as follows (Dasgupta, 2016, 843):

(i) Laws L are the complete laws of motion governing our world.

(ii) Feature X is variant under the symmetries of L.

(iii) Therefore, X is undetectable (from (i) and (ii)).

(C) Therefore, X is not real (from (iii) and an Occamist norm that we dispense with undetectable structure).

Like Earman and Norton, Dasgupta claims that such reasoning motivates the search for an alternative theory that solely trades in the invariant quantities.

I will once again argue that we can avoid the undetectability of variant quantities by assuming Leibniz Equivalence, which only implies the Invariance Principle if we also accept the Value-Magnitude Link. But first let me comment on the 'Occamist norm' to which Dasgupta refers. This norm states that *all else being equal* we ought to dispense with undetectable structure. The 'all else equal' clause is important here: Dasgupta believes that there may be good reasons to posit undetectable structure. Such structure may yield a simpler theory, or one with greater explanatory strength. "In that case, we would have empirical evidence of sorts that the feature is real, in the sense that our all-things-considered best empirically confirmed theory implies that it is real" (Dasgupta, 2016, 854). Dasgupta thus believes it is possible that variant quantities *are* physically real, contrary to the Invariance Principle. Nevertheless, the spirit of Dasgupta's argument is clearly against variant quantities. It is therefore worthwhile to show that there is a sense in which such quantities are *not* undetectable on a sophisticated view.

Here is how Dasgupta initially defines undetectability:

> [S]omething is undetectable in my sense if, roughly speaking, it is physically impossible for it to have an impact on our senses. (Dasgupta, 2016, 853)

This definition of detectability is rather weak: even absolute velocities are detectable on this definition. After all, in Newtonian mechanics relative velocities supervene on absolute velocities, and since the former impact our senses, so do the latter. For example, suppose that we increase the absolute velocity of a ship at sea. Since this also increases the relative velocity of the ship with respect to the shore, the difference is clearly visible. Therefore, the absolute velocity of the ship has an impact on our senses. On this definition, the symmetry-to-reality inference is inconsequential.

However, Dasgupta quickly revises his initial definition in favour of something stronger:

> It then follows [...] that the feature is undetectable [if] there is no physically possible process by which we might discover which determinate values are actually instantiated. (Dasgupta, 2016, 871)

This is a stronger criterion, because it demands that we can measure the particular *value* of a quantity (presumably up to some degree of accuracy), rather than just its effect. Let us call this notion *measurability* (cf. Ismael and van Fraassen's (2003) distinction between the observable and the measurable). For instance, absolute velocity is unmeasurable in the sense that there is no unique numerical value we can assign to the motion of any particular object: we can only express velocities relative to some other object. What is undetectable are not the variant quantities themselves, but their particular values. Suppose that the current mass of the Eiffel Tower is 7 million kilograms. And suppose that there exists a physically possible world qualitatively identical to ours, but in which the masses of all bodies are doubled—hence the mass of the Eiffel tower is 14 million kilograms in that world. Since these two worlds are (by stipulation) empirically equivalent, there is no experiment that could reveal the *actual* mass of the Eiffel Tower. For such an experiment, if it were an accurate measurement of the Eiffel Tower's mass, would have to have *different* outcomes in these worlds.[20] Therefore, we cannot know, even in principle, whether the Eiffel Tower has the mass denoted by '7 million kilograms' rather than the mass denoted by '14 million kilograms'.

If we follow Dasgupta's Occamist advice and reject the physical reality of intrinsic mass, there simply are no mass values to measure. But such a solution to the problem of undetectability is unnecessarily strong. As we have seen, the problem arises when our theories imply the existence of symmetry-related worlds that are empirically equivalent yet physically distinct. The issue is a failure of Leibniz Equivalence, rather than the Invariance Principle. If symmetry-related models represent the same

[20] This, in a nutshell, is Roberts' (2008) argument for the claim that variant quantities are unmeasurable.

possible worlds, then there are no differences between them, so *a fortiori* no undetectable differences. As I explained in §2.4, Leibniz Equivalence is consistent with the rejection of the Invariance Principle once we reject the Value-Magnitude Link. On such a view, the value '7 million kilograms' does not denote the *same* mass across models. Instead, sophistication adopts an anti-quidditist picture on which physical mass magnitudes are *qualitatively identified*. Since the values '7 million kilograms' in the first model and '14 million kilograms' in the second instantiate the same qualitative profile (for example, the same mass ratios to external bodies), the Eiffel Tower is represented as having the same mass in both models. Sophistication implies that the Eiffel Tower has *some* mass, whereas reduction renounces intrinsic mass quantities altogether. Nevertheless, it may seem as if there is still a further question: *which* mass does the Eiffel Tower have? If this is a question about the primitive identities of mass values, there simply are none on the sophisticated account—so the question is void. If, on the other hand, this is a question about the *qualitative* features of the Eiffel Tower's mass, the answer is the same for both models: the Eiffel Tower's mass is 100 000 times my mass, a mere fraction of the Sun's mass, and so on. On this view, we *can* measure which mass the Eiffel Tower (or any other object) has, since we can measure a mass value's qualitative features.

Therefore, sophistication avoids undetectable properties without the Invariance Principle; Leibniz Equivalence is all that is needed. Dasgupta's epistemic argument does not justify the latter once we recognise the Value-Magnitude Link as a tacit but unwarranted assumption.

2.5.3 Reference

The previous section was concerned with an epistemic argument. The final argument, found in Healey (2006) and Caulton (2015), is semantic in nature. The claim I will contest is that variant quantities cannot be given a realist interpretation because they have no unique referent.

Suppose we are trying to interpret a physical theory. Part of the aim of interpretation is to assign referents to the terms that occur in the theory in question. Some of these terms will obviously correspond to physical quantities with which we are directly acquainted, such as colour or distance. We call these *observables*. Since we already know what the observables *are*, we can simply stipulate which theoretical quantities denote them. Caulton (2015) calls this the first phase of interpretation. But theories usually also introduce novel terms for *unobservable quantities*, such as 'electric field'. We have no grasp on what these unobservables are apart from what the theory says about them. It seems that we cannot know what these theoretical terms refer to, unless we have *already* interpreted the theories in which they occur. This raises the worry that the interpretation of theoretical terms is an impossible mission.

Fortunately, the situation is not hopeless. Following Lewis (1970), we can construct *implicit definitions* of the theory's theoretical terms once we have interpreted the observable terms by stipulation; this is Caulton's (2015) second phase of interpretation. Instead of directly connecting a theoretical term to its referent, we declare that the term refers to *whatever* it is that is related to the observables in the manner described by the theory. For example, the electric field is just whatever is generated by, and exerts a force on, charged particles in accordance with Maxwell's laws. For this approach to work, our theories must be *uniquely realised*. There must exist in nature a unique set of unobservable quantities that the theoretical terms of our theories refer to. If this condition fails to hold, then the theory is either not realised at all (i.e. it is false), or it is *multiply realised*. In the latter case it is *indeterminate* what its theoretical terms refer to.

Both Healey and Caulton argue that if we interpret variant quantities as physically real, theories with non-trivial symmetries are multiply realised. In order to avoid indeterminacy of reference we ought to embrace the Invariance Principle. But as with the two previous arguments, Leibniz Equivalence is enough to avoid a failure of reference. The further step from Leibniz Equivalence to the Invariance Principle is unwarranted. I will borrow an example from Healey (2006) to illustrate this claim. Suppose that we have a toy theory according to which sub-atomic particles called 'quarks' cluster together to form protons and neutrons. Quarks can have one of three colours: red, green and blue. Colour is a dynamically efficacious property which couples to various other quantities in the laws. But the theory has a colour permutation symmetry, such that if one changes (for example) all red quarks to green, all green quarks to blue, and all blue quarks to red, the difference is empirically indiscernible. Finally, the dynamics of our toy theory are such that quarks are strongly confined in colour-neutral combinations of red, green and blue.

Healey argues that in this set-up the terms 'green', 'red' and 'blue' are referentially indeterminate. For simplicity, let's label these properties R, G and B. The toy theory has a partially interpreted model of the form $\mathfrak{m}(r, g, b)$, where r, g and b are variables that stand for the colour properties of certain quarks. Suppose that we want to implicitly define r, g and b as whatever triple of properties jointly satisfies $\mathfrak{m}$. Furthermore, suppose that $\mathfrak{m}$ is satisfied if r denotes R, g denotes G, and b denotes B. In other words, $\mathfrak{m}(R, G, B)$ is a dynamically possible model. It then seems that the implicit definition of r picks out R as its unique referent, and similarly for the other terms. But since our theory is colour permutation invariant, $\mathfrak{m}(G, B, R)$ is *also* a dynamically possible model. This means that $\mathfrak{m}(r, g, b)$ is *also* satisfied if r refers to G, and g refers to B, and b refers to R. Therefore, an implicit definition in terms of $\mathfrak{m}(r, g, b)$ leaves it indeterminate whether r refers to R or G (or even B!), and likewise for g and b. This is a consequence of the theory's colour permutation symmetry.

Healey concludes from this that variant quantities such as quark colour are not physically real. Similarly, Caulton (2015, 161) writes that this "prompts appropriate reform towards a *new* formalism, in which the physical properties and relations—including the unobservable ones—are transparently represented without redundancy". In other words, Caulton advocates reduction. However, the demand for a reduced theory is too strong: assuming Leibniz Equivalence is enough to avoid non-unique reference. Recall that the problem that causes the indeterminacy of reference of the terms r, g and b is that $\mathfrak{m}(r, g, b)$ is realised multiply by $\mathfrak{m}(R, G, B)$, $\mathfrak{m}(G, B, R)$ and $\mathfrak{m}(B, R, G)$. But if we accept Leibniz Equivalence, symmetry-related models are distinct representations of the *same* state of affairs, and hence the differences between these permuted models are merely representational. If this is the case, $\mathfrak{m}(r, g, b)$ is *not* multiply realised: it only has one physical realisation, which is redundantly represented by an equivalence class of symmetry-related models.[21] Leibniz Equivalence alone suffices to avoid multiple realisation, of which the indeterminacy of theoretical terms is a consequence.

On this view, what *do* the terms r, b and g refer to? Note that these terms are *values* of the colour quantity, rather than quantities themselves. According to sophistication, their referent is not fixed across models: there is no unique colour property to which we can consistently refer across theoretical models with r, b or g. Rather, whether we call a colour property 'red', 'green' or 'blue' is a conventional choice that may vary across models. Nevertheless, it *is* possible to identify theoretical terms qualitatively. For example, we can (arbitrarily) stipulate that r refers to the colour of the quark which is located at the centre of mass of a particular proton. This is entirely analogous to the stipulation that we will use the term 'kilogram' to refer to the mass of the standard kilogram in Paris. On this account, there is no sense in which 'r' refers to the *same* magnitude across models. Instead, each class of symmetry-related models requires a distinct convention on how the terms 'r', 'g' and

[21] The advocate of reduction may respond that its aim is exactly to avoid such representational redundancy. I agree that that is a legitimate motivation for a reduced formalism—but note that in many cases, finding such a formalism has proved difficult if not impossible. For further discussion, see Dewar (2019, §2).

'b' are used. But in each case, these terms have a unique referent.

2.6 Conclusion

I have analysed three arguments in favour of the Invariance Principle. These arguments all start from the claim that without the Invariance Principle, several woes will befall us: indeterminism, undetectability or a failure of reference. But as I have shown, all that is necessary to avoid these woes is Leibniz Equivalence. Leibniz Equivalence and the Invariance Principle are closely related, but they are not equivalent. In particular, sophistication is a view which simultaneously accepts Leibniz Equivalence and rejects the Invariance Principle. Reduction, on the other hand, accepts both Leibniz Equivalence and the Invariance Principle. But since arguments in favour of the Invariance Principle fail, the latter cannot support reduction over sophistication. This does not mean that reduction is for some reason undesirable: as far as the theoretical vices of indeterminism, undetectability and failure of reference are concerned, reduction and sophistication are exactly on a par.

What, then, could decide in favour of one or the other strategy for interpreting symmetry-related models? This is a significant question that I discuss in the next chapter. The clue lies in the 'gap' between reduction and sophistication. As I have argued, sophistication is committed to the existence of variant quantities, even though it does not entail that such quantities have determinate values. For example, the claim "the Eiffel Tower has an intrinsic mass" is (super)true in Newtonian Gravitation, despite the fact that intrinsic mass might be a variant quantity. The benefit of sophistication is that such quantities may serve an *explanatory* purpose. For example, I mentioned earlier that the electrostatic potential ϕ can explain why the electric force E is conservative. Similarly, Martens and Read (2020) argue that the existence of intrinsic mass can explain the 'transitivity of mass ratios', that is, the fact that the mass ratio between two bodies a and c is equal to the product of the mass ratios between a and b and b and c, for any body b.

We can put the point differently. As we saw at the end of the previous section, theories with symmetries exhibit a certain representational redundancy: different models represent the same state of affairs.[22] This redundancy provides one motivation for reduction, which eliminates the offending symmetries from the theory. But if the above is correct, there is a sense in which this additional structure is *not* redundant. Instead, variant quantities such as the electrostatic potential provide the theory with explanatory resources. Far from 'fluff', then, such quantities are necessary for theories to discharge their explanatory duties. But in order to 'access' these quantities, we require the semantics of supervaluation as discussed in §2.4. Sophistication offers a different way of semantically evaluating our theories which allows its 'redundant' structures to play an explanatory role.

I therefore contend that sophistication may strike the perfect balance between literalism's excess of physical content, which leads to underdetermination, and reduction's metaphysical sparseness, which hinders its explanatory aspirations. We ought to accept Leibniz Equivalence but reject the Invariance Principle.[23]

[22] Ismael and van Fraassen (2003) argue that symmetries are "a guide to superfluous theoretical structure"; Earman (2004) calls this 'descriptive fluff'. But Bradley and Weatherall (2020) challenge this view, arguing that theories whose symmetry-related models are isomorphic possess no surplus structure.

[23] This chapter was previously published as 'Invariance or Equivalence: a Tale of Two Principles' in *Synthese* (2021). DOI: https://doi.org/10.1007/s11229-021-03205-5.

Chapter 3

Comparativist Theories and Conspiracy Theories

Many of our most successful theories posit absolute quantities. For instance, Newtonian mechanics is an empirically adequate theory of medium-size low-velocity phenomena that seems to posit both absolute velocities and absolute masses. However, such absolute quantities are often said to be undetectable because they vary under the theory's symmetries.[1] For instance, absolute position varies under boosts; absolute mass (arguably) varies under uniform scalings.[2] In response, *comparativism* aims to reformulate our theories in terms of relational quantities that do not vary under symmetries, such as spatial distances or mass ratios. The absolutism-comparativism debate (or, in the context of discussions around space and time, the substantivalism-relationism debate) goes back to the Leibniz-Clarke correspondence, in which Leibniz argued that space could not be a real substance in which bodies were located since this implies the possibility of worlds that are qualitatively indiscernible yet which differ by a uniform shift of all material bodies. More recent times have seen signif-

[1]See Roberts (2008); Dasgupta (2016); Wallace (2019b); Middleton and Murgueitio Ramírez (2020); Jacobs (2020).

[2]Whether uniform mass scalings are symmetries of Newtonian mechanics is controversial: see Dasgupta (2013); Baker (2014); Roberts (2016); Dasgupta (2018); Martens (2019b).

icant debates over whether mass properties or mass relations are fundamental (see references in fn. 2), and over whether quantities that vary under gauge symmetries are fundamental (Healey, 2007; Arntzenius, 2012; Rovelli, 2014; Dewar, 2019). In both debates, absolutists are pitted against comparativists.[3]

If the comparativist is to succeed, her theory must 'save the phenomena', for even if there are *in fact* only relative velocities and mass ratios, the success of absolutist theories means that the world looks just *as if* there also are absolute velocities and masses. I will argue that if comparativism is true, there is a sense in which it is a *cosmic conspiracy* that the world looks just as if there are absolute quantities. The comparativist cannot satisfactorily explain why it is that our absolutist theories are empirically adequate if in fact there are no absolute quantities. The argument is structurally similar to the well-known No Miracles Argument for realism, which states that the success of any particular scientific theory is a cosmic coincidence if that theory is not in fact (approximately) true.[4] Just as anti-realism cannot explain the empirical adequacy of our theories in general, so comparativism cannot explain the empirical adequacy of absolutist theories in particular. For this reason, I call the argument presented here *the No Miracles Argument against comparativism*.

In §3.1, I explain in more detail the difference between absolutism and comparativism, connecting these positions to a debate over the interpretation of symmetries. In §3.2, I analyse a particular aspect of what it means for some quantity to be *fundamental*, namely that it satisfies a principle of free recombination. §3.3.1 presents a series of examples of *comparativist conspiracies*. These examples are instances of the No Miracles Argument against comparativism, which I outline in §3.3.2. I consider an objection against the argument in §3.4, namely that such conspiracies are simply part of or follow from the laws of any comparativist theory. I argue that such laws

[3]In the case of gauge quantities, Healey's holonomy approach is understood as a form of 'gauge relationism' (Arntzenius, 2012, 194). See §3.3.1 below, and further discussion in Chapter 7.

[4]The *locus classicus* is Putnam (1975); see also Boyd (1991); Psillos (1999).

would imply a suspicious form of communication-at-a-distance. §3.5 concludes with a discussion of absolutism vis à vis underdetermination.

3.1 Absolutism, Comparativism and Symmetries

Quantities are the fundamental building blocks of physical theories. They are similar to ordinary properties and relations, except that quantities can have a range of *values*. The quantity 'mass', for example, is not just the property of being massive, which a particle either has or does not have. Rather, the mass of a particle has a particular *magnitude*, such as 5 kilograms. This feature is often expressed in terms of the distinction between *determinable* and *determinate* properties. In the case of mass, the property of having a mass of (say) 5 kilograms is a determinate of the determinable property of being massive.[5] In addition to the determinable/determinate distinction, quantities possess a number of other well-known features: (i) they are expressed in terms of units; (ii) each object can instantiate at most one value of any quantity; (iii) and the values of quantities bear second-order relations to each other, for instance the relation of one mass being larger than another. I will say little here about what quantities *are*, metaphysically speaking; that is, whether they are universals, or tropes, or whether nominalism applies to them.[6] The answer to this question does not matter for my argument. But I will assume that quantities are represented by functions from the theory's domain into some *value space*. For our purposes, a value space is a structure in the sense of Bourbaki (1957): a set over which certain functions and relations are defined.[7] The value space itself represents a determinable quantity, whereas the elements of the value space represent the determinate magnitudes of this

[5]For more on the distinction between determinables and determinates, see Wilson (2017).

[6]For more on the metaphysics of quantities, see Eddon (2013b) and references therein. I take up the question in more detail in Chapter 5.

[7]The idea of a structured 'value space' for physical quantities is found, in various forms, in van Fraassen (1967); Stalnaker (1979); Gärdenfors (2000); Denby (2001); Funkhouser (2006); Caulton (2015).

quantity. For example, the mass quantity $m(i)$ in Newtonian mechanics is a function from the domain of particles into a value space whose elements represent particular mass magnitudes. The structure of this value space encodes relations between mass values, such as the fact that 6 kg is three times as much as 2 kg. Similarly, we can consider 'position' as a quantity of particles whose value space is physical space; the structure of this value space is then the structure of physical space, for instance Euclidean.[8] Finally, a value space usually has a *representation* in terms of the real numbers. For example, we can assign each mass value a positive real number. This in effect amounts to a choice of unit. But the value space *itself* does not have the structure of the real numbers (considered as a field), since many different maps from the mass values into the real numbers represent the former equally well.[9]

Both absolutism and comparativism are views about what type of quantities are fundamental within a particular class, such as the class of mass quantities (I will say more about what 'fundamental' means in the next section). Absolutism is the view that the fundamental quantities within some class are *monadic*, while comparativism is the view that the fundamental quantities within the same class are *dyadic*. The debate is thus over the *adicity* of the fundamental quantities of nature.[10] I thus agree with Martens (2019b, fn. 3) that the absolutism-comparativism debate does not in the first place concern whether nature's fundamental quantities are intrinsic or extrinsic, where a quantity is intrinsic just in case the value one or more object instantiate is purely a matter of how the objects are in themselves and (in the case of relational quantities) amongst each other.[11] However, we may still want to rule out some extrinsic monadic properties as not genuinely absolute. For example, the

[8]I leave it open whether this means that space or spacetime is really a physical quantity, as argued in Teller (1987).

[9]For more on the numerical representation of quantities, see Wolff (2020) and references therein.

[10]As Dasgupta (2013) notes, one can also imagine 'higher-order' forms of comparativism. For instance, higher-order relationism may claim that the ratios of distances are fundamental. I will not consider such forms of comparativism here.

[11]Cf. Lewis (1983). On this definitions, both properties and relations can count as either intrinsic or extrinsic, so comparativism is not automatically committed to extrinsic quantities.

property of being twice as massive as the Eiffel Tower is monadic in letter, but to admit such properties as fundamental is clearly not in the spirit of absolutism. I am inclined to assert that *any* physically fundamental quantity is intrinsic, but I will not defend this claim here. Instead, I will assume that we have a sufficient grasp on which quantities to count as absolute or comparative for the debate to proceed. For instance, paradigmatic examples of absolute quantities are: position, velocity, mass, local field values; paradigmatic examples of comparative quantities are: distance, relative velocity, mass ratio, field holonomies (cf. §3.3.1). The debate we are interested in is over whether quantities of the former type or quantities of the latter type are (more) fundamental.[12]

Finally, there is a deep connection between the absolutism-comparativism debate and the occurrence of *symmetries* in physics. Comparativism is often motivated by the claim that some absolute quantities are variant under a theory's group of symmetry transformations, whereas the relevant comparative quantities are not. For instance, distances but not positions vary under so-called static shifts. The variance of absolute quantities under symmetries is said to entail that they are undetectable, which leads to a particularly problematic form of underdetermination. For claims to this effect, see the discussion in §1.2.1 and references therein. The informal argument goes as follows: suppose that one were to move the entire material universe three metres north. Since this is a symmetry transformation which preserves all distances, the difference is indiscernible. In particular, any measurement device will display the exact same value before and after such a shift. But that means that the measurement device does not accurately measure absolute position, which varies under shifts. Therefore, absolute position is undetectable. The argument holds *mutatis mutandis* for other variant quantities. For this reason comparativism goes hand in hand with a view called *reductionism*: the demand for a 'reduced' theory formulated solely in terms of

[12]This means that the argument below applies just as much to 'substantivalist' accounts of quantities as to nominalist accounts in the style of Field (1980).

new, invariant quantities.[13] The comparativist follows the following recipe to obtain a reduced theory: (i) start with the theory's absolute quantities, m, which vary under some symmetry transformations; (ii) construct new, comparative quantities r from the old quantities m that are invariant under the same symmetry transformations; (iii) formulate a reduced theory in terms of these new quantities. For an example, consider the monadic function $m(i)$, which represents the mass of some particle i. Suppose that the $m(i)$ vary under the theory's symmetries, but that their ratio $m(i)/m(j)$ does not. The comparativist can then postulate a new, relational quantity $r(i,j)$ as fundamental, where by definition $r(i,j) \equiv m(i)/m(j)$. We can think of $r(i,j)$ as the mass ratio between i and j, but since the comparativist renounces the (fundamental) reality of absolute masses this is true only metaphorically; as Roberts (2016, 5) has it, "there is nothing for them to be ratios *of*." The comparativist claim is rather that ratios between absolute quantities merely represents the fundamental relational quantity in a redundant way.

In principle, reductionism and comparativism are orthogonal views: one can construct reduced theories with invariant but non-relational quantities, and conversely one can formulate a theory in terms of some variant relational quantities.[14] But the combination of comparativism and reduction is a natural one, since the same worry about superfluous symmetry-variant structure motivates both views. Both historical and contemporary defences of comparativism have drawn the connection to symmetries, hence reduction has often taken the form of a move from some set of variant quantities to another set of invariant quantities of some higher adicity. In the remainder of this chapter I will therefore focus just on forms of comparativism which posit *invariant* comparative quantities as fundamental.

[13]For more on reductionism, see Dewar (2019); Martens and Read (2020); Jacobs (2021), as well as the discussion in §1.2.

[14]For example, Newton-Cartan Theory is an example of reduction without relationism; a theory in which all positions are expressed in terms of their distance to a privileged point of the manifold would count as an example of relationism without reduction.

3.2 Fundamentality and Common Ground

I have defined absolutism and comparativism in terms of what sort of quantities are considered fundamental—but what does it mean for a quantity to *be* fundamental? Contemporary discussions of absolutism-versus-comparativism put the issue in terms of *ground* (Dasgupta, 2013; North, 2018; Martens, 2019b). We can think of ground as a 'vertical' relation between fundamental and non-fundamental entities. I will discuss this account of fundamentality in §3.2.1. But one can also consider 'horizontal' relations *between* fundamental entities. Here, the idea that fundamental quantities are 'modally free' has proven attractive. I discuss this idea in §3.2.2.[15] Finally, §3.2.3 puts the vertical and horizontal accounts of fundamentality together to formulate a criterion for what counts as *evidence* of fundamentality, namely the so-called *common ground inference* (Ismael and Schaffer, 2020).

3.2.1 Vertical Fundamentality

I follow Dasgupta (2013) in assuming that some class of quantities is less fundamental than another class of quantities iff any fact about the former class of quantities is (partially) grounded in some facts about the latter class of quantities. Here, *grounding* is a relation (between either facts or objects—I remain neutral on the exact details) which expresses a concept of "metaphysical *because*" (North, 2018).[16] For example, the fact that my bicycle is solid is grounded in the dynamical behaviour of the particles which constitute it. Similarly, the fact that it is currently raining is grounded in the fact that small drops of water produced by clouds are currently falling from the sky. Crucially, grounding is *not* a causal relation. The fact that drops of water are falling from the sky is not what causes it to rain, but what makes it the case *that* it is

[15] In terms of Dorr and Hawthorne's (2013) discussion of Lewisian naturalness, the vertical relation concerns (something close to) 'Supervenience', while the horizontal relation concerns (something close to) 'Independence'.

[16] For more on the metaphysics of grounding, see Fine (2012); Rosen (2010); Schaffer (2009) and references therein.

raining. Instead, grounding is a relation of constitutive explanation.

The debate between absolutism and comparativism is a debate about ground. In particular, the absolutist claims that comparative quantities of some class (distances, mass relations) are less fundamental than comparative quantities of the same class (positions, masses): any fact about the comparative quantities is (partially) grounded in some facts about the absolute quantities. The comparativist claims the contrary: some facts about the comparative quantities are not grounded in any facts about the absolute quantities.[17]

Put in these terms, then, the debate is about *relative* fundamentality: even if the facts about mass magnitudes ground the facts about mass relations, it might still be the case that facts about mass magnitudes are themselves grounded in some even more fundamental facts. But there is good reason for this, since we know that the theories in question are not fundamental. For example, Newtonian mechanics is a highly accurate theory within the domain of medium-sized low-velocity phenomena, but it does not describe nature at its most fundamental level. Therefore, there is little reason to believe that the fundamental quantities of Newtonian mechanics are fundamental quantities, period. In what follows I will usually discuss these positions *as if* they concern the most fundamental quantities in nature. For this allows us to appeal to the fact that a quantity is fundamental iff facts about this quantity are *ungrounded*, that is, iff nothing else grounds them. In what follows, I will simply say that absolutism claims that monadic but not comparative quantities are fundamental, whereas comparativism claims that comparative quantities are (also) fundamental.[18] But it is always understood that 'fundamental' here at most means: most fundamental with respect to the quantities posited by the theory in question. I will remain silent on

[17]On this definition, comparativism is consistent with the claim that both absolute and comparative quantities are equally fundamental. However, most (if not all) contemporary comparativists have made the stronger claim that comparative quantities are more fundamental than absolute quantities. The difference between these positions does not matter for our purposes.

[18]Martens (2019b) calls these positions 'Strong Absolutism' and 'Strong Comparativism' respectively.

the further question of whether the most fundamental quantities within a particular theory are also the fundamental quantities of nature.

3.2.2 Horizontal Fundamentality

So much for the 'vertical' relations between fundamental and non-fundamental quantities. What are the 'horizontal' relations *between* fundamental quantities? The guiding idea here is that fundamental quantities are 'modally free': since facts about fundamental quantities are ungrounded, any way those quantities could logically hang together is a way in which they possibly *can* hang together. This idea is often seen as a consequence of Hume's Dictum, which Wilson (2010) expresses as follows:

> *Hume's Dictum*: There are no metaphysically necessary connections between wholly distinct entities.

There are more than a few intricacies involved in this principle, which we need not unpack here. The relevant point for our purposes is this: Hume's Dictum implies that if x and y are wholly distinct entities and m is a certain quantity, then the values of $m(x)$ and $m(y)$ are mutually independent. Here I assume that entities are wholly distinct iff the entities neither ground each other nor possess a common ground.[19] So, if there is some possible world in which $m(x) = a$, and some (distinct) possible world in which $m(y) = b$, then there is a possible world in which both $m(x) = a$ *and* $m(y) = b$. If such a world is not possible, then there is a necessary connection between the value of $m(x)$ and $m(y)$. Wang (2016) calls this consequence of Hume's Dictum the 'Fundamentality Entails Modal Freedom' (FEMF) principle, which she glosses as follows: "[The set of fundamental properties and relations] is modally free iff any pattern of instantiation of the properties or relations in [this set] is possible" (451). Put differently, Hume's Dictum implies that a necessary condition on fundamentality is that fundamental quantities are *freely recombinable*.

[19]Ismael and Schaffer (2020, fn. 6) suggest that this is the notion of 'wholly distinct' Hume himself may have had in mind.

The above statement of FEMF explicitly mentions both properties and relations. I intend to use the principle in the same way. However, Hume's Dictum (and, by extension, FEMF) is usually taken to apply only to monadic quantities. For instance, David Lewis (1986) famously synthesised a principle of recombination with the thesis of *Humean supervenience*, which sees the world as "a vast mosaic of local matters of particular fact [...] all else supervenes on that" (ix). On Lewis' picture, Hume's Dictum only applies to monadic quantities because only monadic quantities are fundamental. But the restriction of Hume's Dictum to non-relational quantities is an historical accident. Some decades before Lewis, Carnap (1950, §118) formulated a syntactic principle of free recombination which applied to relations as much as to properties. For Carnap, possible worlds are represented by *state descriptions*: classes of sentences of a language L which, for each atomic sentence ϕ of L, contain either ϕ or $\neg\phi$—and nothing else. If the set of primitive predicates of L contains relations, then there is a state description for any of their possible patterns of instantiation. Carnap thus advocated a version of free recombination closer to mine than to Lewis.

Like Carnap, I believe that the intuition behind Hume's Dictum applies to relations as much as to properties. In what follows, I will therefore assume that any fundamental quantity is modally free, whether monadic or polyadic. We will see in §3.3 that it is exactly the failure of comparative quantities to freely recombine which suggests that they are non-fundamental. Before that I will first discuss the inference from a quantity's apparent failure to recombine to its lack of fundamentality in more detail.

3.2.3 Common Ground

I now put the vertical and horizontal accounts of fundamentality together. From the claim that fundamental quantities freely recombine, it follows via contraposition that quantities which do not recombine are not fundamental. Moreover, from the claim that quantities are fundamental iff facts about such quantities are not grounded in any further facts, it follows that facts about non-fundamental quantities *are* grounded in some further facts. If some quantity does not seem to freely recombine, then we have reason to believe that that quantity is grounded in some more fundamental quantity. This is the principle Ismael and Schaffer (2020) call *Grounding Inference*:[20]

> *Grounding Inference*: If non-identical entities a and b are modally connected, then either (i) a grounds b, or (ii) b grounds a, or (iii) a and b are joint results of some common ground c.

The 'entities' in question can include facts about the pattern of instantiation of certain quantities. We are interested specifically in cases of type (iii), which Ismael and Schaffer call *common ground inferences*. The idea of a common ground inference is analogous to Reichenbach's (1956) common cause principle, by which we infer a common cause from correlations between simultaneous events. Such principles are widespread in the history of science. Janssen (2002) coins the term 'common origin stories' for a form of inference to the best explanation which mirrors common ground inferences. Janssen (2009) uses this inference to conclude that spacetime structure underlies the symmetries of the dynamics of relativistic theories. In a similar spirit, Caulton (2013) has recently formulated the related idea of "Hume's Dictum as a guide to what there is", which he traces back to Hume, Carnap and the early Wittgenstein.

In order to clarify the idea of common ground inferences consider the following example, borrowed from Ismael and Schaffer (2020). The ideal gas law states that $pV = nRT$. For a given volume V, this means that the pressure of a gas is proportional to its temperature. This is a modal connection between facts about pressure and facts about temperature. For example, the ideal gas law tells us that if the pressure of a gas were to increase, then its temperature would also increase. From

[20]Note that Ismael and Schaffer conceive of 'Hume's Inference' as a *consequence* of Grounding Inference, whereas for us the reverse is true. The latter order seems more natural to me, since grounding inferences are chiefly justified by principles such as FEMF.

Grounding Inference it follows that these facts must have a common ground. This is indeed the case: as kinetic theory tells us, both the pressure and the temperature of a gas supervene on the movement of the molecules which compose it. Therefore, facts about molecular motion are the common ground of facts about both pressure and temperature. This example thus illustrates a successful common ground inference.

What is our *evidence* for the existence of such modal connections? The way we gain evidence for modal connections is just the way we gain evidence for counterfactual behaviour more broadly: a balance of introspection and experiment. For example, we know that there is a modal connection between the pressure and temperature of a gas since we have repeatedly observed that whenever we increase the pressure of a gas, its temperature also increases. Of course, such evidence is defeasible: it is always possible that what *seems* a modal connection is in fact just an accidental coincidence. But this is no problem: the same situation occurs for the common cause principle, since it is always possible (indeed, expected!) that some correlations are spurious. Expecting a perfect epistemic situation would be unreasonable, so we'll have to admit that common ground inferences are sometimes faulty. Ismael and Schaffer (2020, 4139) put this as follows: "Grounding Inference simply says that all else being equal, in the kind of epistemic setting in which we have no direct access to the grounding substructure of a collection of objects, a theory that explains constraints on their modal covariation by reference to a common ground is better than one that regards it as a brute modal connection between distinct existences. [...] Grounding Inference expresses a preference for theories that trace modal connections to common grounds over ones that don't." Grounding Inference is not intended as an *a priori* claim, but rather as a regulative principle.

Finally, there is a deep connection between Grounding Inference and laws of nature. In particular, we can view the laws of nature as modal connections which *resist* explanation in terms of common ground. Once we have explained away as many modal connections as possible, the ones we are left with are genuine laws of nature. Thus, our maxim is not to explain away absolutely every modal connection, but rather to *minimise the modal correlations posited by a theory (subject to empirical adequacy)*. This maxim reveals a preference for common ground inferences over scientific explanations. But this preference is well-justified, since common ground inferences are more modally robust. It is usually assumed that the grounding relation is stronger than necessitation (Rosen, 2010). It follows that if some modal connection is due to a common ground, then said modal connection holds with metaphysical necessity. On the other hand, a modal connection which results from some law only holds with physical necessity. Therefore, there is a sense in which explanations in terms of common ground are better than explanations in terms of laws of nature. The claim that the modal connections which resist common ground explanations are candidate laws is then justified via an appeal to a principle of simplicity: a theory which reduces modal connections to a common ground is more economical than a theory which postulates the same connections as novel laws. Recall that on the Best Systems Account of lawhood, the laws are the axioms and theorems of the deductive system which best balances strength and simplicity (Lewis, 1973, 73). If systems with fewer brute modal connections are simpler, then the modal connections which remain in such a system are the laws of nature (absent any reduction in strength). In this way, Grounding Inference acts as a *guide* to the laws.[21] I will return to these issues in §3.4.

[21]This connection between Hume's Dictum and the Best Systems Account deserves further exploration. These ideas are discussed in somewhat more detail by Ismael (MS).

3.3 Comparativist Conspiracies

In this section, I will first (§3.3.1) provide a number of examples of 'cosmic conspiracies' in comparativist theories.[22] None of these examples are new, but as far as I am aware no one has yet suggested that such conspiracies are a pervasive feature of comparativist theories.[23] I will go on to (§3.3.2) present a novel theorem which shows that comparativist conspiracies occur whenever some reasonable set of assumptions is satisfied. This constitutes the No Miracles Argument against comparativism.

Interestingly, a version of this objection was already levelled by Yehoshua Bar-Hillel (1951) against Carnap's syntactic principle of free recombination (see §3.2.2). Carnap required that all primitive predicates and relations are logically independent from each other. For example, the properties of being blue and of not being blue cannot both be primitive, since this allows for a state description in which some object is simultaneously blue and not blue. But as Bar-Hillel pointed out, it is not enough to require that all primitive relations are logically independent *from each other*, since some relations also seem to have an *intrinsic* logical structure. For example, the relation 'warmer than' (Wxy) is transitive, and so there is no system in which Wab, Wbc and Wca are all true. This is so even if W is the language's sole predicate, so the requirement of logical independence is trivially satisfied. In other words, Bar-Hillel's objection to Carnap's theory of state descriptions is that it implies that if the relation 'warmer than' is primitive, it must be possible for a to be warmer than b, b warmer than c yet c warmer than a—which is clearly not possible. In terms of the previous section, Bar-Hillel points out that the modal connections in the pattern of

[22]There are further examples, for which I do not have space to discuss here. For example, Dewar (2019) argues that comparativism cannot explain that *handedness* partitions the domain of bodies into equivalence classes; both Roberts (2016) and Dewar (2019) show that comparativism about potentials fails to account for the fact that certain forces are conservative (see also the discussion in §2.4); and Leeds (1999) believes that only absolutism about the vector potential can explain the exact form of the momentum operator in electrodynamics.

[23]Dewar (2019) mentions some of the examples below in the context of his discussion of reduction. However, Dewar does not relate the conspiracies to comparativism specifically.

instantiation of W suggests that W is not a fundamental relation. Carnap (1951) responded that in a sufficiently strong language, qualitative relations such as 'warmer than' are defined in terms of the numerical values of monadic quantities. For instance, let $T(x)$ denote the temperature of x with a real number, and stipulate that Wxy iff $T(x) > T(y)$. In that case, 'warmer than' must be transitive, since there is no triple of real numbers such that $T_1 > T_2$ and $T_2 > T_3$ yet $T_3 > T_1$. Carnap realised that the transitivity of 'warmer than' can only be guaranteed if 'warmer than' is not fundamental, but logically depends on some primitive monadic quantity. We can recognise a common ground inference in Carnap's response: the modal connection between temperature comparisons is explained by the fact that such comparisons are grounded in facts about monadic temperatures.

This same pattern will repeat itself over and over in the examples below: comparativist quantities seem to bear certain modal connections to each other, but if they are truly fundamental then these connections are cosmic coincidence. On the other hand, if the comparativist quantities are grounded in absolute quantities then the modal connections between the former arise as a necessary consequence: absolutism explains the structural patterns in the instantiation of comparative quantities. Therefore, our principle of Grounding Inference favours absolutism.

3.3.1 Comparativist Conspiracies: Some Examples

Here are some examples of comparativist conspiracies:

The Triangle Inequality

The first example concerns relationism in the traditional sense: the view that bodies have no absolute locations in space, but that instead the distances between them are fundamental. Maudlin (2007) argues that relationism cannot guarantee that the *Triangle Inequality* is satisfied. The Triangle Inequality states that, for any three particles in space, the distance between i and j added to the distance between j and k must equal or exceed the distance between i and k. It is essential for any empirically adequate relationist theory to satisfy the Triangle Inequality, since it holds true in the actual world.

Maudlin's argument concerns a spacetime of any arbitrary (Riemannian) geometry, but we can simplify the example by considering a three-dimensional Euclidean space. Let's introduce the fundamental dyadic quantity $\vec{d}(i,j)$ to represent the displacement vector between a pair of particles i and j (of course, just as fundamental mass relations aren't *really* ratios between masses, fundamental displacements aren't *really* vectors between locations). In terms of this quantity, the Triangle Inequality states that:

$$|\vec{d}(i,j)| + |\vec{d}(j,k)| \geq |\vec{d}(i,k)| \tag{3.1}$$

What explains the truth of (3.1)? When distances are fundamental, it follows from Hume's Dictum that $\vec{d}(i,j)$ and $\vec{d}(j,k)$ cannot constrain $\vec{d}(i,k)$; any pair of bodies could, in principle, have any distance between them—*even if this violates the Triangle Inequality.* Therefore, Maudlin concludes, the relationist has to *postulate* (3.1) in order account for the fact that it holds true in the actual world.[24] But this, of course, is not an explanation of *why* the Triangle Inequality holds.

The principle of Grounding Inference suggests that the modal connections between distances exemplified in the Triangle Inequality must be explained in terms of a grounding relation. In particular, we can explain the Triangle Inequality via a common ground inference to absolute positions. On a substantivalist view, distances between bodies are grounded in their absolute positions. If $x(i)$ denotes the position

[24] It has also been argued that it is *conceptually necessary* for distances to satisfy the Triangle Inequality (Bricker, 1993; Forrest, 1995; Maudlin, 2007; Belot, 2011). But if that is true, then it is the case whether or not distances are fundamental, so this issue does not directly bear on the matter at hand.

of i in a Euclidean (affine) space, then displacements are three-dimensional vectors over this space, i.e. $\vec{d}(i,j) := \overrightarrow{ij}$. The Triangle Inequality then reduces to:

$$|\overrightarrow{ij}| + |\overrightarrow{jk}| \geq |\overrightarrow{ij} + \overrightarrow{jk}| \tag{3.2}$$

Unlike (3.1), (3.2) is a mathematical theorem: the fact that distances are (represented by) displacement vectors between points on an affine space *explains* the Triangle Inequality. The Triangle Inequality is necessarily true on an absolutist picture of space. There is a sense in which it could not have failed to hold. Therefore, facts about absolute positions are the common ground for facts about distances which can explain (away) the modal connections between them.[25] The fact that distances obey the Triangle Inequality surely makes it seem *as if* they depend on positions. This conclusion is supported by a common ground inference.

Transitivity of Mass Ratios

The second example concerns particle mass in Newtonian mechanics. Recall that if mass is absolute, each particle i is assigned a fundamental mass magnitude which is represented by a monadic function $m(i)$. The mass relations supervene on these monadic quantities. But according to comparativism about mass, mass *relations* are fundamental: mass relations are not grounded in mass magnitudes. These mass relations are represented by functions $r(i,j)$ from *pairs* of particles into some value space. Now, masses vary under so-called *Active Leibniz Scalings*: uniform scalings of all masses by some positive constant. Dasgupta (2013) argues that such scalings are symmetries of Newtonian mechanics, which motivates the search for a reduced theory in terms of invariant mass relations. Whether mass scalings are indeed symmetries of Newtonian mechanics is a thorny issue which I will not address here; instead, I focus

[25]You may object that the substantivalist's explanation of the Triangle Inequality *also* requires a cosmic coincidence, namely, the fact that space in fact *has* an Euclidean structure (or any Riemannian geometry). Hold that thought—I will discuss this objection in detail in §3.4.

on what Martens (2019a) calls the *conspiracy of mass relations*.[26]

The comparativist's mass relations must obey the following constraint:

$$r(i,k) = r(i,j) \cdot r(j,k), \tag{3.3}$$

I will call (3.3) the *transitivity of mass relations* (TMR). TMR states that the mass ratio between i and j times the mass ratio between j and k is equal to the mass ratio between i and k. For example, if John is twice as massive as Jane, and Jane is three times as massive as Jules, then John must be six times as massive as Jules. Any comparativist theory must satisfy (3.3) in order to remain empirically adequate. But if mass relations are fundamental, then nothing guarantees the validity of (3.3). After all, it follows from Hume's Dictum that any pair of bodies can bear any mass relation to each other. The transitivity of mass relations could easily have been false.

Therefore, just like the Triangle Inequality, TMR is the sort of cosmic conspiracy that demands explanation in terms of a common ground.[27] If mass relations are grounded in monadic masses, then TMR is a mathematical theorem (Martens and Read, 2020). For in that case we can write $r(i,j) = m(i)/m(j)$, and hence:

$$m(i,k) = \frac{m(i)}{m(k)} = \frac{m(i)}{m(j)} \cdot \frac{m(j)}{m(k)} = r(i,j) \cdot r(j,k) \tag{3.4}$$

Clearly, (3.4) holds for whatever values we assign to $m(i)$, $m(j)$ and $m(k)$. Therefore, the existence of monadic masses in which mass relations are grounded allows us to explain the transitivity of mass ratios. Put differently, the TMR surely makes it seem *as if* mass relations are really ratios between mass properties. The common ground inference supports the conclusion that mass properties are indeed more fundamental than mass relations.

[26] See Baker (2014) or Martens (2019a) for a response to Dasgupta's arguments.
[27] This point was already made by Russell (1903); see also Armstrong (1988).

Composite Loop Multiplication

The final example concerns the ontology of electrodynamics. Since the example is somewhat technically involved, I will be brief here: I discuss electrodynamics in more detail in Chapter 7.

On a literal reading of electrodynamics, the vector potential $\mathbf{A}$ is a real vector field: it assigns a three-dimensional vector to each spacetime point. For any open path γ, this field defines a line integral $\mathbf{A}(\gamma) = \int_\gamma \mathbf{A}d\mathbf{x}$. The vector potential is a gauge quantity, which means that it varies under the theory's (local) gauge symmetries. For this reason, $\mathbf{A}$ is traditionally seen as a mathematical artefact which does not represent a real field. But the discovery of the Aharonov-Bohm effect, in which $\mathbf{A}$ seems to play a causal role, has led many physicists—most notably Feynman et al. (1964)—to consider $\mathbf{A}$ as physically real. Call this view *gauge absolutism.*

However, that $\mathbf{A}$ is variant means that its values are unmeasurable. In response to this objection, Healey (2007) proposes that not the field itself but its *holonomies* are fundamental: functions of integrals of $\mathbf{A}$ around closed paths, or 'loops', in spacetime. The holonomy of a loop l is defined as $H(l) = e^{iq \oint_l \mathbf{A}d\mathbf{x}}$, where q is a fixed unit of charge. Unlike the vector field itself, holonomies *are* invariant under the gauge symmetries of electrodynamics. Since any loop l is composed of a (non-unique) pair of open paths γ_1 and γ_2, we can consider holonomies as relations between those pairs of open paths which jointly form a closed loop. For any pair of open paths that jointly form a closed loop, $H(\gamma_1, \gamma_2)$ is their 'holonomy relation'. For this reason Arntzenius (2012) dubs Healey's view 'gauge relationism'.

Arntzenius also points out that, in order for gauge relationism to remain empirically adequate, holonomies must satisfy the relation of *Composite Loop Multiplication*:

$$H(l_1 \circ l_2) = H(l_1)\, H(l_2), \tag{3.5}$$

where $l_1 \circ l_2$ denotes the concatenation of l_1 and l_2, that is, the path that results from first going around l_1 and then going around l_2. It will come as no surprise that in this case, too, the relationist cannot explain why (3.5) holds: if holonomies are fundamental, then $H(l_1 \circ l_2)$ could have had any value no matter what the values of $H(l_1)$ and $H(l_2)$ are. The fact that this relation seems to hold in the actual world is another comparativist conspiracy. Once more, Grounding Inference suggests that holonomies possess a common ground.

If holonomies supervene on integrals of $\mathbf{A}$ along open paths, then (3.5) reduces to a mathematical theorem:

$$\begin{aligned} H(l_1)\, H(l_2) &= \exp\left[iq \oint_{l_1} \mathbf{A} d\mathbf{x}\right] \cdot \exp\left[iq \oint_{l_2} \mathbf{A} d\mathbf{x}\right] \\ &= \exp\left[iq \oint_{l_1 \circ l_2} \mathbf{A} d\mathbf{x}\right] \\ &= H(l_1 \circ l_2) \end{aligned} \tag{3.6}$$

As Arntzenius (2012, 195) puts it, "a fairly obvious explanation of why [CLM] hold[s] is that the map H is, roughly speaking, the integration of [$\mathbf{A}$] around a loop". Put differently, (3.5) surely makes it look *as if* holonomies are simply loop integrals of a real vector field $\mathbf{A}$. The claim that the vector field acts as the common ground of the holonomies is then simply the result of a common ground inference.

3.3.2 Comparativist Conspiracies: Proof Outline

It is no coincidence that comparativist theories are full of cosmic conspiracies. Such conspiracies are a pervasive feature of comparativism. In Appendix A, I substantiate this claim by means of a formal proof that, under a fairly general set of assumptions, *any* reductionist form of comparativism implies cosmic conspiracies. In this section I give an informal outline of the proof.

The essential point is that relational quantities are usually not postulated *de novo*, but rather defined in terms of the old, absolute quantities. As we have seen, this is slightly awkward since (for example) fundamental mass ratios are not really ratios *of* anything. But it also entails that relational quantities are *constrained*, since it is not the case that any arbitrary pattern of instantiation of (for example) mass relations is consistent with some assignment of masses. If the pattern of mass relations fails to satisfy the transitivity condition, it is not consistent with any assignment of monadic mass properties. In more detail, we can prove that if the relational quantity r is invariant under a symmetry group G which acts on the absolute quantity's value space $\mathcal{A}$, then the value space $\mathcal{R}$ of the relational quantity r 'inherits' the mathematical structure of G. Therefore, *each value of the relational quantity itself represents a symmetry transformation.* For example, each displacement vector between a pair of bodies can itself act on Euclidean space to effect a uniform translation. Similarly, each mass ratio—which is expressed by a positive real number α—corresponds to a uniform scaling of all masses by the same factor α. This connection between symmetries and relations is expressed in the following theorem (see Appendix A for the proof):

Theorem. *If $r : \mathcal{A}\times\mathcal{A} \to \mathcal{R}$ is invariant under a regular group action of G on $\mathcal{A}$, then $\mathcal{A}$ is a principal homogeneous space for $\mathcal{R}$, i.e. there exists a function $r^{-1} : \mathcal{A}\times\mathcal{R} \to \mathcal{A}$ which represents a regular group action of $\mathcal{R}$ on $\mathcal{A}$.*

From this theorem the existence of cosmic conspiracies follows as a corollary. The intuitive idea is that symmetry transformations must compose in certain ways. For example, the result of one symmetry transformation after another is itself a symmetry transformation. In the case of mass scalings, this means that the result of the successive application of one scaling after another is itself also a scaling. Since there is a correspondence between symmetry transformations and values of r, these composition rules also apply to the latter. Put differently, we can always define the 'composite' of a pair of relational values. In the case of mass scalings the composite of the mass relation represented by the real number α and the mass relation represented by the real number β is the mass relation represented by the product $\alpha \cdot \beta$ (since mass ratios are dimensionless, this result is independent from our choice of units). Generally, we have that:

Corollary (Cosmic Conspiracy). *For any $i, j, k \in \mathcal{D}$,*

$$r(i,j) \circ r(j,k) \circ r(k,i) = I \tag{3.7}$$

where $\circ$ denotes the composition operation and I is the identity element, i.e. the unique relation that each object bears to itself. For example, in the case of mass ratios the identity element is 1, since any object is equally massive as itself—so (3.7) says that the product of the mass ratios of any three objects equals 1, which is just the transitivity of mass ratios from before. Similarly, the sum of the vectors between any triple of points in a Euclidean space is the zero vector: the vector that points from any location to itself. The corollary says that the values for r of any triple of objects have to 'add up' to the identity. The result immediately implies that relational quantities violate the principle of free recombination.

3.4 Conspiracies or Laws?

I anticipate the following objection to the No Miracles Argument against comparativism, using the substantivalism-relationism debate as a representative example:[28]

> The claim that positions are fundamental does *not* entail the Triangle Inequality any more than the claim that distances are fundamental does. After all, the Triangle Inequality is only satisfied if physical space has the structure of a *metric space*, such as Euclidean space. But this is a substantive assumption about the distance relation between points of space themselves. If the absolutist is allowed such an assumption about distances between points, then why isn't comparativist allowed a similar

[28]Compare Dees' (2015) response to Maudlin on the Triangle Inequality.

> assumption about distances between *particles*? The absolutist simply replaces one miracle with another! But we cannot explain miracles with miracles, so comparativism is no worse off than absolutism.

The same objection can be made *mutatis mutandis* for quantities such as mass. In the case of mass, for instance, the transitivity of mass ratios is guaranteed by the fact that we can represent mass magnitudes with positive real numbers, which in turn follows from the fact that mass value space has an *additive structure*.[29] But, the objection goes, mass value space could have had a different structure for which mass multiplication is not transitive. If this objection succeeds, then comparativist 'conspiracies' are in fact just laws, in the same way that it is a law of Newtonian mechanics that space is Euclidean.

I concede the central contention: absolutism *does* require an additional posit, for instance that space is Euclidean or that the structure of mass value space is additive. But I do not believe that this fact poses a challenge to the No Miracles Argument against comparativism. In particular, I will argue that comparativism is still not allowed to postulate conspiracies as laws. This restores the explanatory asymmetry between absolutism and comparativism implicit in the common ground inferences from above.[30]

Recall from §3.2.3 that we can never explain away all modal connections, since the laws themselves are modal connections of a sort. If it were a consequence of Grounding Inference that the laws stand in need of explanation, then we cannot take it seriously as a scientific principle. For that would in effect entail a version of the Principle of Sufficient Reason, which states that *any* truth—even the truth of the laws—must have an explanation. That seems too ambitious.[31] Instead, I advanced the idea that common ground inferences 'bottom out' at the laws of nature. Once we

[29]Roughly, this means that order and addition are well-defined. For more on additive structures, see Krantz et al. (1971, Ch. 3) or (Wolff, 2020, §5.2). See also the discussion in Chapter 6.

[30]I thank Jenann Ismael for an insightful exchange of ideas on this point.

[31]Dasgupta (2016) presents a contemporary version of such a principle. But although Dasgupta argues that the principle is not as absurd as it seems, he does not actually defend it.

have explained away the modal connections we *can* explain away, the laws are what we are left with. Put differently, we have a decision procedure: can we explain (away) some modal connection between quantities via a common ground inference? If the answer is 'Yes', then we conclude that these quantities are not fundamental. If the answer is 'No', then the modal connection is a law-like relation between fundamental quantities. Since a theory with as few bare modal claims as possible is, all else being equal, simpler, such a theory is preferred on broadly empiricist grounds.

Let us now apply these ideas to the case of substantivalism versus relationism. The main claim is that the Triangle Inequality for *particles* admits of a common ground inference, whereas the Triangle Inequality for *points* does not. From the above, it then follows that the latter counts as a genuine law, whereas the former is a spurious modal connection between non-fundamental quantities. The first half of this claim—that we can explain the Triangle Inequality for particles in terms of their absolute position within a Euclidean space—was defended in §3.3.1. My aim here is to defend the converse, namely that no such explanation exists for the pattern of distance relations between points.[32] In particular, we cannot ground the structure of space itself in facts about the distances between particles which occupy space. This restores the asymmetry between substantivalism and relationism.

The reason one cannot ground the structure of space in facts about distances between particles is that particles in different worlds occupy different points of space. Imagine a class of models of Newtonian mechanics, each of which represents a finite number of particles positioned within a Euclidean affine space. The relationist may argue that in any particular world, the distance between a pair of *occupied* points supervenes on the distance between the particles which occupy these points. But this leaves us without any ground for the distances between *unoccupied* points. Moreover,

[32]Of course, one may be able to explain the (appearance of) Euclidean structure in terms of some more fundamental theory, so that this structure is only fundamental with respect to the posits of Newtonian mechanics itself—cf. the remark at the end of §3.2.1.

it is wholly unclear how one *could* ground these distances solely in facts about the distances between actual particles. On the other hand, the particles across these worlds are all positioned within the same Euclidean space (up to isomorphism). Therefore, we can reduce facts about the distances between particles in different physically possible worlds to the same set of facts about the actual structure of space. In this sense, absolutism offers a unificatory account of the Triangle Inequality in the sense of Friedman (1974) and Kitcher (1981).[33]

These obstacles to relationism are of course well-known.[34] In response, relationists have attempted to supplement their accounts of the structure of space with further facts. I lack the space to discuss these responses in detail, so I will barely scratch the surface of this debate. But let me briefly mention two approaches, and my reasons for scepticism. The first approach is to include, in addition to facts about the distances between *actual* particles, also facts about the distances between *possible particles* (or, perhaps, facts about the possible distances between actual particles). But this approach—called *modal relationism*—has come in for substantial criticism; see, for instance, Field (1984); Butterfield (1984); Belot (2011). In particular, it seems that such modal relationism has little appeal for empiricists about laws, whose aim it is to reduce facts about what is physically possible to facts about what is actually the case. I will not further discuss this approach here. The second approach is to include, in addition to facts about the distances between particles, facts about the form of the laws; in particular, facts about whether there exists a coordinate system in which the laws have a 'simple' form. Huggett's (2006) *regularity relationism* is an example of

[33]Of course, if matter forms a plenum *in every physically possible world*, then one can offer an account of the structure of space in terms of the distances between matter points. But as Field (1984) notes, this move essentially gives up the spirit of relationism. Moreover, the assumption of a plenum is hardly tenable for non-spatiotemporal quantities, such as mass; for the analogue assumption is that in *each* physically possible world, *every* mass value is instantiated by some matter point. But this is not the case even for the field theories which seem most amenable to a treatment of matter as a plenum.

[34]See, for instance, Field (1980, Ch. 6), Belot (2011, Ch. 2), Arntzenius and Dorr (2012) and references therein.

such an approach. Martens (2017) sketches an analogous approach for comparativism about mass, which he calls ‘regularity comparativism’. If such approaches succeed, then the symmetry between absolutism and comparativism is restored. However, I do not believe that regularity approaches succeed. In brief, regularity comparativism runs the risk of collapsing into regularity *eliminativism*, the view which holds that there are *no* fundamental quantities, whether absolute or comparative. Pooley (2013) expresses this worry clearly:

> Once the strict requirement that primitive vocabulary should express primitive, perfectly natural properties and relations is relaxed, what governs which quantities are part of the supervenience base and which quantities are supervenient? Why stop at a reduction of genidentity and inertial structure? Why not seek to offer a reductive account of mass and charge too? Why not even seek a reductive account of the temporal metric and instantaneous spatial distances? Once the reduction via the dynamical laws of some apparently natural properties to the others is on the table, we need some principles to determine which properties are ripe for reduction and which are to be part of the basic ideology. (568)

So regularity approaches overshoot their aim. For a similar criticism of of regularity approaches for mass quantities, see Martens (2017). Nevertheless, it remains the case that such approaches are at present the most promising response for comparativism to the No Miracles Argument. I intend to address these issues in more detail in future research.

Finally, note that these considerations apply *mutatis mutandis* to quantities such as mass or field strength. For instance, while one can ground facts about mass relations in facts about monadic mass properties, the reverse is not the case. This is a consequence of the fact that each particle (in any world) instantiates a mass value from the same value space, whereas different mass values are instantiated by different particles. A single mass value space unifies facts about particles across space, time and possible worlds. Therefore, the modal connections between mass values too are not cosmic conspiracies, but genuine laws.

3.5 Conclusion

The No Miracles Argument against comparativism is *ipso facto* an argument in favour of absolutism, since the latter can explain comparativist conspiracies via common ground inferences. But note that I have used a 'thin' definition of absolutism, namely as a commitment to monadic quantities. The term 'absolutism' is often used elsewhere in a much stronger sense. For instance, a belief in absolute positions is normally understood to entail that worlds related by static shifts—that is, uniform translations of the entire material universe—are physically distinct. This leads to the well-known worry that absolutist theories with non-trivial symmetries are committed to a harmful form of underdetermination. But my definition of absolutism does not have this consequence. For example, absolutism in the thin sense is perfectly consistent with anti-haecceitism, the thesis that there are no numerically distinct yet qualitatively identical worlds. Since shift-related worlds *are* qualitatively identical, anti-haecceitism implies that models related by a shift represent the same possibility: shifts are 'distinctions without a difference'. The conjunction of substantivalism and anti-haecceitism is called *sophisticated substantivalism* (Pooley, 2006). Importantly, sophisticated substantivalism is still a form of absolutism in the thin sense, since it is committed to the fact that objects possess some position in space: that's what makes it a form of substantivalism. Therefore, sophisticated substantivalism can explain the Triangle Inequality, which is simply a consequence of the fact that objects are located on a Riemannian manifold, without underdetermination.

A similar stratagem exists for for non-spatiotemporal quantities, such as mass. Here, the analogue of anti-haecceitism is *anti-quidditism*, the thesis that determinate values are qualitatively individuated. For example, suppose that Active Leibniz Scalings leave all qualitative mass facts the same (that is, all mass facts except for facts about *which* mass values are instantiated): then anti-quidditism implies that objects in mass-scaled worlds in fact have the *same* mass. Just like static shifts, scalings are revealed to be distinctions without a difference. Dewar (2019) has recently explored the application of sophistication to internal quantities in much detail; in Chapter 6 and 7 of this thesis, I discuss sophistication for mass and for gauge quantities respectively.

In conclusion, sophistication offers a 'third way' position. Of course, there is much more to be said here. But for our purposes, the crucial point is that sophisticated absolutism does not result in underdetermination. Since this is the chief argument against absolutism, we can conclude that the No Miracles Argument shows that we should prefer absolutism over comparativism. For absolutism can explain the cosmic coincidences which are conspiracies for the comparativist, so all else being equal the former is preferable to the latter. Moreover, on a 'sophisticated' interpretation absolutism does not entail underdetermination, so all else *is* equal. To paraphrase Putnam: "absolutism is the only philosophy that doesn't make the success of comparativist theories a miracle".

3.A Proof of Theorem

Suppose that $m : \mathcal{D} \to \mathcal{A}$ is a function from a certain domain $\mathcal{D}$ into a value space $\mathcal{A}$. Call m the *absolute quantity*.[35] Let $\bar{r} : \mathcal{D}^2 \to \mathcal{R}$ be a function from ordered pairs of objects in $\mathcal{D}$ into a different value space $\mathcal{R}$. Call $\bar{r}$ the *relational quantity*. We assume that $\bar{r}(i, j) \equiv r(m(i), m(j))$, where r is a surjective function from $\mathcal{A} \times \mathcal{A}$ onto $\mathcal{R}$. We require that r is surjective, i.e. that each relational value corresponds to some pair of absolute values. If this were not the case, then the relational value space would contain superfluous elements, given that it is meant to recover the empirical content of the absolute value space (that is, the superfluous relational values would be empirically inaccessible given the theory's dynamics).

In order to account for the invariance of r under some relevant symmetry group G, we first suppose that G has an action ψ on $\mathcal{A}$. For ease of expression, instead of $\psi_g(x)$ we will simply write gx for the result of acting with g on x. We will use e to denote G's identity, and assume that the action of G is *regular*, i.e. it is both *free* (if $gx = x$ then $g = e$) and *transitive* (for any x, y, there exists a g such that $y = gx$). To be sure, this is not the case for all symmetries. For instance, rotations of an affine space are not free since they have a fixed point. Conversely, charge conjugation is not transitive, since it is not the case that any pair of charges is related by a charge conjugation. Whether a similar result holds for these symmetries is an open question.

We can then define the invariance of r as follows:

> *Invariance*. The relational quantity r is invariant under G's action on $\mathcal{A}$ iff for any $x, y, x', y' \in \mathcal{A}$,
>
> $$r(x, y) = r(x', y') \text{ iff } y' = gy \text{ and } x' = gx \text{ for some } g \in G.$$

Invariance says that symmetry transformations—and *only* symmetry transformations—

[35] $\mathcal{D}$ could also consist of tuples of objects to allow for cases of higher-order comparativism, as mentioned in §3.1.

leave the values of r invariant. This means that a pair of models agrees on the values of r if and only if they are symmetry-related.

We will prove that when r is an invariant relation, the values of $r(x, y)$ and $r(y, z)$ constrain the value of $r(x, z)$. If that is the case, the relational quantities are not mutually independent: a cosmic conspiracy. In order to show this, we first prove a few lemmas:

Lemma 3.A.1. *For any $x \in \mathcal{A}$, there exists a bijection f_x between $\mathcal{A}$ and G and a function $r_x : G \times G \to \mathcal{R}$ such that $r_x(f_x(y), f_x(z)) = r(y, z)$ for all $y, z \in \mathcal{A}$.*

Proof. Let x denote an arbitrary 'base point' of $\mathcal{A}$, and let $f_x(y)$ denote the unique $g \in G$ such that $y = gx$. Since G's action on $\mathcal{A}$ is regular, such a unique g is guaranteed to exist. Define $r_x(g, h) \coloneqq r(gx, hx)$; it then immediately follows that $r_x(f_x(y), f_x(z)) = r(y, z)$. □

Lemma 3.A.2. *The function $k(r_x(g, h)) \coloneqq g^{-1}h$ is a bijection between G and $\mathcal{R}$.*[36]

Proof. We first show that k is well-defined. Since r_x is surjective, any element of $\mathcal{R}$ has an expression of the form $r_x(g, h)$. From Invariance it follows that for any element of $\mathcal{R}$ the product $g^{-1}h$ is unique. For $r_x(g, h) = r_x(g', h')$ iff $r(gx, hx) = r(g'x, h'x)$, iff there is a $k \in G$ such that $g'x = kgx$ and $h'x = khx$, and so $g' = kg$ and $h' = kh$; thus $g'^{-1}h' = (kg)^{-1}kh = g^{-1}h$. So, k is indeed a function from $\mathcal{R}$ to G. It also follows from Invariance that k is injective: if $g^{-1}h = g'^{-1}h'$ then $r_x(g, h) = r_x(g', h')$. For let $g' = jg$ and $h' = kh$; then $g'^{-1}h' = g^{-1}j^{-1}kh$, which is equal to $g^{-1}h$ iff $j = k$. But then $r_x(g', h') = r_x(jg, jh) = r_x(g, h)$ from Invariance, as required. Finally, k is clearly surjective, since for any $g \in G$, $k(r_x(e, g)) = g$. Therefore, k is a bijection. For ease of expression, we will write v_g for the (unique) element of $\mathcal{R}$ such that $k(v_g) = g$. □

[36]This is not a *canonical* bijection: unless G is Abelian, k will depend on our choice of base point of $\mathcal{A}$. But this will not implicate the main result below, which hold for any choice of base point; see Corollary 3.A.3.1.

Lemma 3.A.3. *The function r_x has an inverse $r_x^{-1} : G \times \mathcal{R} \to G$ such that $r_x^{-1}(g, v)$ yields the unique $h \in G$ for which $r_x(g, h) = v$.*

Proof. For ease of exposition, we will write $r_x^{-1}(g, v) =: g + v$. Define $g + v := gk(v)$; then it immediately follows that $g + r_x(g, h) = h$. $\square$

Corollary 3.A.3.1. *Define the inverse of $r(y, z)$ as $r^{-1}(y, v) := f_x^{-1}\big(r_x^{-1}(f_x(y), v)\big)$: this function is independent of the choice of base point x.*

Proof. Let $x = gy$, so $f_y(z) = f_x(z)g$. Suppose $v = r(p, q)$ for some $p, q \in \mathcal{A}$ such that $p = g_p x$ and $q = g_q x$. Then $v = r_x(g_p, g_q) = r_y(g_p g, g_q g)$ and so $r_y^{-1}(f_y(z), v)\big) = \big(f_x(z)gg^{-1}g_p g_q\big)g = \big(r_x^{-1}(f_x(z), v)\big)g$. Therefore, $f_y^{-1}\big(r_y^{-1}(f_y(z), v)\big) = f_x^{-1}\big(r_x^{-1}(f_x(z), v)\big)$. $\square$

With these results in hand, we can prove the following theorem of interest:

Theorem. *If r is invariant under a regular group action of G on $\mathcal{A}$, then r^{-1} represents a regular group action of $\mathcal{R}$ on $\mathcal{A}$ (i.e. $\mathcal{A}$ is a* principal homogeneous space *for $\mathcal{R}$).*

Proof. We will use the fact that for any y there exists a unique g such that $y = gx$ for some base point x, so it is sufficient to show that $\mathcal{R}$ has a regular action on G. It follows from Corollary 3.A.3.1 that the particular choice of base point x does not matter. Therefore, we will write drop the indices and simply write $r(g, h)$ and $r^{-1}(g, v)$. Let $v_g := k^{-1}(g)$, and define $v_g \circ v_h := v_{hg}$. Since k is a bijection between $\mathcal{R}$ and G, it is clear that $\langle \mathcal{R}, \circ \rangle$ forms a group. We will now show that r^{-1} represents a (regular) group action of $\mathcal{R}$ on $\mathcal{A}$.

1. *Identity*: for all $g \in G$, $g + v_e = g$. *Proof*: by definition, $v_e = r(e, e)$, so $g + v_e = g + r(e, e) = g$

2. *Associativity*: for any $g \in G$, $v_h, v_j \in \mathcal{R}$, $(g + v_h) + v_j = g + (v_h \circ v_j)$. *Proof*: by definition, $v_h = r(e, h)$ and $v_j = r(e, j)$, and so

$$
\begin{aligned}
(g + r(e,h)) + r(e,j) = hg + r(e,j) = jhg \\
= g + r(e, jh) = g + v_{jh} = g + (v_h \circ v_j)
\end{aligned}
$$

3. *Regularity*: for any $g, h \in G$, there exists a unique $v \in \mathcal{R}$ such that $g + v = h$. *Proof*: let $v := r(g, h)$; then by definition, $g + r(g, h) = h$.

It follows that $\langle \mathcal{A}, + \rangle$ is a principal homogeneous space for $\mathcal{R}$. □

From this theorem it follows that relational quantities are involved in cosmic conspiracies. We will abbreviate $\bar{r}(i, j) =: v_{ij}$. The claim then is that:

Corollary (Cosmic Conspiracy). *For any $i, j, k \in \mathcal{D}$,*

$$v_{ij} \circ v_{jk} \circ v_{ki} = v_e \tag{A.1}$$

Proof. From the definition of $\bar{r}(i, j)$, it follows that $v_{ij} = r(m(i), m(j))$. Let $m(i) = g_i x_0$, so $v_{ij} = r(g_i, g_j)$. Therefore,

$$
\begin{aligned}
g_i + (v_{ij} \circ v_{jk} \circ v_{ki}) &= (g_i + v_{ij}) + (v_{jk} \circ v_{ki}) && \text{(A.2)} \\
&= g_j + (v_{jk} \circ v_{ki}) && \text{(A.3)} \\
&= (g_j + v_{jk}) + v_{ki} && \text{(A.4)} \\
&= g_k + v_{ki} && \text{(A.5)} \\
&= g_i && \text{(A.6)}
\end{aligned}
$$

where we have used the associativity of $+$ and $\circ$. Since $g_i + (v_{ij} \circ v_{jk} \circ v_{ki}) = m(i)$, it follows that $v_{ij} \circ v_{jk} \circ v_{ki}$ is the identity, v_e. Hence, $v_{ik} = v_{ij} \circ v_{jk}$. The values of v_{ij} and v_{jk} thus uniquely constrain the value of v_{ki}; this proves the corollary. □

3.B Worked Examples

3.B.1 Vector Relationism

We assume that:

1. $x(i) : \mathcal{D} \to A$, where $\mathbb{A}$ is a set of points;
2. $\vec{d}(i,j) : \mathcal{D}^2 \to \mathbb{V}$, where $\mathbb{V}$ is a vector space;
3. $\vec{d}(i,j) \equiv x(j) - x(i)$, for some binary function '$-$';
4. $\vec{d}(i,j)$ is invariant under T, the group of translations of $\mathbb{A}$.

From the Theorem, it follows that the inverse of $\vec{d}$ is a free and transitive action of $\mathbb{V}$ on $\mathbb{A}$. Since $\mathbb{V}$ is a vector space, this means that $\mathbb{A}$ is *affine space*: $y - x$ yields the displacement vector from x to y. The Cosmic Conspiracy corollary follows straightforwardly (where we abbreviate $\vec{d}(i,j)$ to $\overrightarrow{ij}$):

$$
\begin{aligned}
x(i) + (\overrightarrow{ij} + \overrightarrow{jk} + \overrightarrow{ki}) &= (x(i) + \overrightarrow{ij}) + (\overrightarrow{jk} + \overrightarrow{ki}) && \text{(A.7)} \\
&= x(j) + (\overrightarrow{jk} + \overrightarrow{ki}) && \text{(A.8)} \\
&= (x(j) + \overrightarrow{jk}) + \overrightarrow{ki} && \text{(A.9)} \\
&= x(k) + \overrightarrow{ki} && \text{(A.10)} \\
&= x(i) && \text{(A.11)}
\end{aligned}
$$

and hence $\overrightarrow{ij} + \overrightarrow{jk} + \overrightarrow{ki} = 0$. In physical terms, what this means is that for any triple of objects, the displacement vector from i to j plus the displacement vector from j

to k is identical to the displacement vector from i to k.

3.B.2 Mass Ratios

We assume that:

1. $m(i) : \mathcal{D} \to \mathcal{V}_m$, where $\mathcal{V}_m$ is our mass value space;
2. $r(i,j) : \mathcal{D}^2 \to \mathbb{R}^+$, where $\mathbb{R}^+$ is the group of positive real numbers under multiplication;
3. $r(i,j) \equiv \bar{r}(m(i), m(j))$, for some binary function $\bar{r}$;
4. $r(i,j)$ is invariant under S, the group of scalings of $\mathcal{V}_m$.

From the theorem, it follows that $\mathcal{V}_m$ is a principal homogeneous space for $\mathbb{R}^+$, such that $\bar{r}(m(i), m(j)) \coloneqq m(i)/m(j)$. We can then easily prove the existence of Cosmic Conspiracies:

$$r(i,j) \cdot r(j,k) \cdot r(k,i) = \frac{m(i)}{m(j)} \cdot \frac{m(j)}{m(k)} \cdot \frac{m(k)}{m(i)} \tag{A.12}$$

$$= 1 \tag{A.13}$$

In physical terms, this means that the mass ratio between i and j times the mass ratio between j and k is identical to the mass ratio between i and k.

3.B.3 Holonomies

Let P_G be a principal G-bundle over a manifold M. We assume that:

1. $A(\gamma) : \Gamma \to h$, where Γ is the set of open paths $\gamma[a,b]$ on M, and h is the set of all homomorphisms between the fibre above a and the fibre above b;

2. $H(l_a) : L \to G$, where L is the set of all closed paths l_a with base point a on M, and G is the set of functions from the fibre above a to itself (which is isomorphic to G);

3. $H(\gamma_1 \circ \gamma_2) \equiv r(A(\gamma_1), A(\gamma_2))$, for some binary function r, whenever the concatenation of γ_1 and γ_2 forms a loop;

4. $H(\gamma_1 \circ \gamma_2)$ is invariant under G, the structure group of P_G.

From the theorem, it follows that h is a principal homogeneous space for G such that $H(\gamma_1 \circ \gamma_2) = A(\gamma_1) \circ A(\gamma_2)$, where the rhs denotes a composition of homomorphisms. We can then prove the cosmic conspiracy as follows. Let $l_1 = \gamma_1 \circ \gamma_2$, and $l_2 = \gamma_2^{-1} \circ \gamma_3$.

Then:

$$H(l_1)H(l_2) = (A(\gamma_1) \circ A(\gamma_2))(A(\gamma_2^{-1}) \circ A(\gamma_2)) \tag{A.14}$$

$$= A(\gamma_1) \circ A(\gamma_3) \tag{A.15}$$

$$= H(l_1 \circ l_2) \tag{A.16}$$

Chapter 4

Sophistication: Symmetry versus Structure

In the previous chapter we have seen that certain forms of reduction often entail cosmic conspiracies. Of the arguments against reduction surveyed in the introduction, this is perhaps the most serious. We also saw that sophistication can explain (or perhaps explain *away*) such coincidences because it is committed to the reality of variant quantities. The remainder of this dissertation consists of a defence of sophistication as a better alternative to reduction. I will return to the explanation of cosmic coincidences in the case studies in Chapter 6 and 7. In this chapter I will first offer a more detailed account of the formalism of sophisticated theories. In the introduction, I briefly mentioned the existence of two distinct approaches: an 'intrinsic' approach on which a theory's models are defined axiomatically, and an 'extrinsic' approach on which a theory's models are defined in terms of the isomorphisms between them. Dewar (2019) has recently defended the latter approach, which was criticised by Martens and Read (2020) as 'sophistry'. I will here develop the distinction between the intrinsic and extrinsic approaches in more detail. I will conclude that the extrinsic approach to sophistication is indeed ultimately unsatisfactory, but not (only) for the reasons given by Martens and Read.

Dewar (2019) offers a sustained defence of sophistication as an approach to symmetry-related models (SRMs). The defence focuses on theories whose SRMs are not already isomorphic, in which case sophistication seems *prima facie* difficult. The aim of sophistication in such cases is to *remove* structure from the theory's models until their symmetry-related instances are isomorphic. Once symmetry-related models are isomorphic, we can interpret them anti-haecceitistically or anti-quidditistically as physically equivalent. Dewar draws an important distinction between what he calls the *internal* and the *external* approach to sophistication. Since the terms 'internal' and 'external' are also used to distinguish types of symmetries, I will instead call these the *structure-first* and the *symmetry-first* approaches to sophistication. Both approaches aim to redefine a theory's models in such a way that its SRMs become isomorphic. But on the structure-first approach one does so by intrinsically specifying a particular Tarski-style model, i.e. a set of relations, functions and operators defined over some domain, while on the symmetry-first approach one gives an 'extrinsic' characterisation of the same structure in terms of a class of allowed symmetry transformations. I describe the difference between these approaches in more detail below.[1]

Although sophistication is expressed here in purely formal terms—what is the appropriate mathematical structure for our models?—the approach is clearly meant to provide us with an *interpretation* of symmetry-related models. In other words, both structure-first and symmetry-first sophistication aim to provide us with a picture of 'what the world is like'. Therefore, we should evaluate both approaches *as* interpretations of SRMs. A purely formal criterion is unlikely to suffice for this purpose, since one also has to be able to say what the formalism represents. While both

[1]A similar distinction was in fact already drawn in Suppes (1967). Martens and Read (2020) draw the internal/external distinction in much the same way as I. I contrast my analysis with theirs in §4.3. The symmetry-first approach is closely aligned to the category-theoretic view of theories (Weatherall, 2015; Halvorson and Tsementzis, 2018), on which isomorphisms between models are stipulated rather than derived.

approaches succeed in the formal aspects to sophistication, my claim is that only the structure-first approach provides a 'perspicuous metaphysical picture' of what the world is like. Only the structure-first approach successfully offers an *interpretation* of symmetry-related models.

4.1 Structure-First Sophistication

On the structure-first approach, the aim is to give an 'intrinsic' characterisation of a structure in terms of relations defined over its domain, such that this structure is invariant under the dynamical symmetries of the theory. In more detail, suppose that the general form of a theory's models is $\langle \mathcal{D}_1, \mathcal{D}_2, ..., \mathcal{Q}_1, \mathcal{Q}_2, ... \rangle$, where the $\mathcal{D}$s (the sub-domains) are structured sets and the $\mathcal{Q}$s (the quantities) are functions from some $\mathcal{D}_i$ into some $\mathcal{D}_j$. This form for a theory's models is rather more abstract than is usual in discussions of symmetries, but the models of more familiar theories fit easily within this scheme.[2] For example, the models of Newtonian Gravitation are of the form $\langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \rho \rangle$, where $\mathcal{M} = \langle M, t_{ab}, h^{ab}, \sigma^a \rangle$ is a Newtonian spacetime; $\mathcal{G}$ is a one-dimensional affine space in which the gravitational potential field has its values; $\mathcal{V}_m$ is a structured set whose elements represent mass density values; ϕ, the gravitational potential, is a function from $\mathcal{M}$ into $\mathcal{G}$; and ρ, the mass distribution, is a function from $\mathcal{M}$ into $\mathcal{V}_m$. We can write each of the structured sub-domains $\mathcal{D}_i$ as $\langle D_i, \mathbf{R}_i \rangle$, where D_i is the (unstructured) base set of $\mathcal{D}_i$ and $\mathbf{R}_i$ is a set of functions and relations defined over D_i. So the sub-domains are structures in the sense of Bourbaki (1957). For example, $\mathcal{M}$ consists of a base manifold M over which we have defined some (geometrical) structure in the form of the tensor fields t_{ab}, h^{ab} and σ^a (of course, M itself has a non-trivial topological structure).

Within this framework we can characterise structure-first sophistication as follows. Suppose that a theory's symmetries are induced by bijections of some base set D_i (we

[2] I will say more about this form for a theory's models in the next chapter.

will see an example of this below). On the structure-first approach, the aim is to define a structure $\mathcal{D}_i$ such that the relations $\mathbf{R}_i$ are invariant under these bijections. If this is the case, the induced symmetries partition the space of models into equivalence classes of isomorphic SRMs. As a result there is a 'match' between the structure of the theory's models and the dynamics of the theory in question. This approach is in fact a generalisation of Earman's (1989) symmetry principles, which say that a theory's spacetime symmetries ought to coincide with its dynamical symmetries. For example, if uniform boosts of the universe's matter content are among the symmetries of Newtonian Gravitation, then Earman's symmetry principles require that the theory's spacetime structure is invariant under such boosts. This means that this spacetime cannot contain a standard of absolute rest. But spacetime is only one of the theory's sub-domains. On the present approach, we require that the structure of *each* of the theory's sub-domains is invariant under the relevant dynamical symmetries. We thereby extend Earman's symmetry principles to internal symmetries.[3]

The mass symmetries of Newtonian Gravitation provide a simple example of structure-first sophistication for internal symmetries.[4] We can first define a *mass density value space* $\mathcal{V}_m := \langle D_m, \leqslant, \circ \rangle$, where D_m is a set of cardinality $2^{\aleph_0}$. The relation $\leqslant$ imposes a total order on D_m, and $\circ$ is an associative binary function. In physical terms, $\leqslant$ is interpreted as the relation of one mass per unit volume being less than or equal to another, and $\circ$ as a function from two values of mass per unit volume to their sum. When these relations satisfy certain axioms, one can prove a *representation theorem* which states that there exists a function $f_r : D_m \to \mathbb{R}^+$ such that $m_i \leqslant m_j$ iff $f_r(m_i) \leq fr(m_j)$ and $m_i \circ m_j = m_k$ iff $f_r(m_i) + f_r(m_j) = f_r(m_k)$.[5]

[3] I say more about this extension in the next chapter.

[4] Whether mass scalings *are* symmetries of NG is a matter of debate: see Dasgupta (2013); Martens (2019a,b). I will argue in Chapter 6 that uniform mass scalings *are* symmetries of NG *when we also transform the gravitational constant G*. But the outcome of this debate does not matter here, since I only intend to use the example of mass scalings for illustrative purposes.

[5] The axioms are listed in Hölder (1901); Krantz et al. (1971); see also Chapter 6 of this dissertation.

In other words, we can represent $\mathcal{V}_m$ on the positive real numbers $\mathbb{R}^+$. Furthermore, one can prove a *uniqueness theorem* which states that if f_r is such a function then so is αf_r for any positive constant α. We can think of each f_r as a system of units, namely an assignment of a (positive) real number to each mass value.

Suppose that *uniform mass scalings* are symmetries of Newtonian Gravitation. Specifically, a uniform mass scaling is a bijection ψ_α on D_m such that $f_r(\psi_\alpha \circ m)= \alpha f_r(m)$ for all $m \in D_m$. Put differently, if a mass m is associated in some representation f_r with some real number x, then ψ_α maps m onto the mass m' which is represented by f_r with the real number αx. It follows from the representation and uniqueness theorems above that uniform mass scalings must preserve the relations $\leqslant$ and $\circ$, so ψ_α is an automorphism of $\mathcal{V}_m$ for any α. Furthermore, the ψ_α induce transformations on the theory's models defined as follows:

$$\langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \rho \rangle \xrightarrow{g_\alpha} \langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \psi_\alpha \circ \rho \rangle \tag{4.1}$$

The claim that uniform mass scalings are symmetries of NG means that the transformations g_α preserve solution-hood. As mentioned above, I will assume that this is the case. Crucial for the project of sophistication is the fact that the models related by (4.1) are isomorphic when the bijections ψ_α are automorphisms of $\mathcal{V}_m$. For example, it is straightforward to see that $\rho(x) \leqslant \rho(y)$ iff $(\psi_\alpha \circ \rho)(x) \leqslant (\psi_\alpha \circ \rho)(y)$; likewise for the binary operator $\circ$. Therefore, we say that Newtonian Gravitation is sophisticated with respect to mass symmetries when $\mathcal{V}_m$ has the structure defined above.

Because models related by mass scalings are isomorphic, we are licensed to interpret them as physically equivalent via an appeal to *anti-quidditism* about masses. Black (2000) defined quidditism in terms of determinable quantities: anti-quidditism then implies that there are no distinct world in which (for instance) mass and charge are swapped. But I will use the term in a slightly different sense, namely to cover the determinate *values* of physical quantities. The idea is that (for instance) mass magnitudes have no primitive identities, but are qualitatively identified via their pattern of instantiation. In this sense, anti-quidditism is simply the analogue of anti-haecceitism for determinates.[6] The differences between isomorphic SRMs are representationally irrelevant, since they concern quidditistic facts about *which* particular mass values are instantiated. Because structural relations between mass values are the same in isomorphic models, anti-quidditism implies that the *same* masses are instantiated across those models.

We have thus satisfied the aim of sophistication, namely to structure the models of our theories such that we can interpret symmetry-related models as physically equivalent. The example illustrates what structure-first sophistication consists of: the stipulation of a set of relations over the theory's sub-domains, such that the bijections which induce dynamical symmetries of the theory's models leave these relations invariant. If we agree that an interpretation of a theory provides a metaphysically perspicuous picture if it tells us plainly and clearly which entities the theory posits (ontology) and what the fundamental relations between these entities are (ideology), then structure-first sophistication is perspicuous in this sense. Moreover, there are clear differences between structure-first sophistication and reduction. Firstly, the symmetry-related models of the sophisticated theory are distinct (albeit isomorphic), since $\psi \circ \rho \neq \rho$ whenever ψ is a non-trivial automorphism. In contrast, any putative reduced theory will have a unique model corresponding to an equivalence class of SRMs of the old theory. Secondly, the sophisticated theory is committed to the existence of a monadic mass quantity represented by the unary function ρ. But ρ varies under the mass symmetries of NG, whereas reduced theories only use invariant quantities. The reduced theory will have to posit a different ontology, for example one

[6]See Dewar (2019) for more detail; Martens and Read (2020) also discuss anti-quidditism in the context of sophistication. Cf. Esfeld and Lam's (2008) 'moderate structuralism' for a similar proposal. I discuss anti-quidditism in the context of mass symmetries in more detail in Chapter 6.

that consists of a relational quantity which represents the mass density ratios *between* points (of course, whether one can reformulate the laws of NG in terms of such a relational quantity is far from trivial). Reduction and sophistication thus present us with both a distinct formalism and distinct metaphysics.

4.2 Symmetry-First Sophistication

Let us now turn towards a characterisation of *symmetry-first* sophistication. In discussing a related issue, Møller-Nielsen (2017, 1262) alleges that we cannot assume that "there will always be [a sophisticated theory] waiting in logical space to be discovered". This sounds true enough: the moves from Newtonian to Neo-Newtonian spacetime or from fields to fibre bundles are certainly non-trivial. But Wallace (2019c, 16) points out that this "underestimates the powerful, general resources by which structure can be subtracted from mathematical theories". The symmetry-first approach to sophistication aims to make full use of these resources and thereby 'brute force' the isomorphism requirement. The idea behind symmetry-first sophistication is to define structures *in terms of their isomorphisms*, such that they are invariant under the theory's symmetries by definition. Dewar has put this somewhat puzzlingly as "declaring, by fiat, that the symmetry transformations are now going to 'count' as isomorphisms" (Dewar, 2019, 502-3). But this obfuscates what is really going on, namely that new structures are defined in terms of the relations of isomorphism between them.[7]

Wallace (2019a) contains a clear exposition of what these structures are like. Here I adapt his treatment to the example of mass scalings. We start with an arbitrary

[7] The symmetry-first approach is thus closely associated with the category-theoretic approach to theoretical equivalence. On this approach, one defines a theory by specifying both a class of models and a set of arrows between these models. When there is an arrow between a pair of models, we can treat them 'as if' they are isomorphic. On this view, theories are physically equivalent iff they are equivalent as categories. I lack the space to consider this approach in detail, but I believe that many of my arguments against symmetry-first sophistication carry over to the category-theoretic approach.

representation function f_r of the mass elements D_m on $\mathbb{R}^+$. Then, we construct an equivalence class $[f_r]$ from f_r, such that $f'_r \in [f_r]$ iff $f'_r = \alpha f_r$ for some positive constant α. In other words, $[f_r]$ specifies which transformations on D_m (construed as *active* transformations) are 'postulated' as isomorphisms. This results in a structure $\mathcal{V}_m \coloneqq \langle D_m, [f_r]\rangle$. In general, one starts with a domain D, a group of symmetry transformations Ψ and a representation function f_r, from which one constructs the structure $\langle D, [f_r]_\Psi\rangle$ where the second item is the equivalence class of functions f'_r such that $f'_r = f_r \circ \psi$ for some $\psi \in \Psi$. These types of structures have some important features. Firstly, they are invariant under the transformations ψ which we have used to define the equivalence class $[f_r]$.[8] For example, $\langle D_m, [f_r]\rangle$ as defined above is invariant under uniform mass scalings. This means that symmetry-first sophistication satisfies the formal aim of sophistication, namely to define structures which are invariant under the theory's symmetries. Secondly, $\langle D_m, [f_r]\rangle$ is unique up to isomorphism (Wallace, 2019a, 128). One could have taken any arbitrary representation function f_r as one's starting point in order to obtain an isomorphic structure. Therefore, the symmetry-first approach defines a more-or-less unique structure which is invariant under any allowed symmetry transformation.

But there are also some crucial differences between the structure-first and the symmetry-first approaches. Firstly, notice that the individual functions f_r are *not* invariant under mass scalings, unlike the relations $\leqslant$ and $\circ$. We cannot consider the individual f_r to represent any physical feature of mass value space, since we require that such features are invariant under the theory's symmetries. Therefore, the f_r's can only represent the structure of $\mathcal{V}_m$ jointly, if at all. Secondly, the structure $\langle D_m, [f_r]\rangle$ is defined in terms of representation functions, which we conceive of as coordinates/unit systems. But coordinates themselves have no physical import, so $\langle D_m, [f_r]\rangle$ does not

[8] *Proof*: For any $f'_r = f_r\psi$ and $\psi' \in \Psi$, $(f'_r\psi')(m) = (f_r\psi\psi')(m)$. But since Ψ is a group, $\psi\psi' \in \Psi$. Therefore, $f'_r \in [f_r]_\Psi]$. Furthermore, ψ has an inverse, hence there is some f''_r such that $f'_r = \psi' f''_r$. Therefore, the set $[f_r]$ is invariant under the action of Ψ.

describe mass space in 'intrinsic' terms. In fact, the symmetry-first structure defined above amounts to a definition of mass value space as 'whatever structure is invariant under uniform mass scalings'. While this may capture the structure of mass value space in the sense that it picks out the correct symmetries, it is hard to see how this delivers any insight into what that structure actually *is*. Compare this to the models of structure-first sophistication, which explicitly lay down relations which characterise mass as a determinable quantity. I will consider this contrast in more detail in §4.5.

4.3 Symmetries and Supervaluation

Call the above feature of symmetry-first sophistication its *metaphysical opaqueness* (contrast with: a 'perspicuous metaphysical picture'). The symmetry-first approach has been attacked on these grounds by 'motivationalists' such as Møller-Nielsen (2017) and Martens and Read (2020):[9]

> [I]t is simply opaque what, according to [the symmetry-first approach], the world is really like. (Møller-Nielsen, 2017, 1264)

> We take it that a complete and honest form of realism should not only take these questions [about metaphysical commitment] on board, but consider them to be crucial. Leaving them out is at best a dishonest form of realism, and at worst a form of anti-realism. (Martens and Read, 2020, 27)

I agree with the above commentators that there is a sense in which symmetry-first sophistication is metaphysically opaque. However, it is an overstatement to call the approach "a form of anti-realism"; as I will show in this section, symmetry-first sophistication *does* have definite metaphysical commitments. Neither is it fair to claim that symmetry-first sophistication is a 'dishonest' form of realism: the symmetry-first

[9] See Martens and Read (2020) for further discussion of how the distinction between structure-first and symmetry-first sophistication ties into the debate between motivationalism and interpretationalism.

approach does not wear its metaphysical commitments on its sleeve, but it does not misrepresent them either. I claim that (a) the symmetry-first approach has an effective *decision procedure* for which fundamental relations it is ontologically committed to, and (b) from this decision procedure, it follows that symmetry-first sophistication has the *same* ontological commitments as structure-first sophistication.

First, consider the *semantics* of symmetry-first sophistication: when does a model satisfy the laws? There is a puzzle here, because the laws are certainly not stated in terms of equivalence classes of representation relations. Dewar (2019, 499) suggests that we can decide the truth-value of theoretical sentences with a *supervaluationist semantics*.[10] The idea is as follows. Consider a sentence *expressed in terms of the numerical representations* of some quantities; for example, the law of universal gravitation:

$$f_r(G)\frac{f_r(m_1)f_r(m_2)}{r^2} = f_r(m_1)a_1 \tag{4.2}$$

Now, consider a putative model of Newtonian Gravitation (a 'kinematically possible model' or KPM). How, on the symmetry-first approach, do we assess whether this is also a 'dynamically possible model' (or DPM), i.e. whether it satisfies (4.2)? According to Dewar's supervaluationist semantics, (4.2) is (super)true iff

$$f_r(G)\frac{f_r(m_1)f_r(m_2)}{r^2} = f_r(m_1)a_1 \text{ for all } f_r \in [f_r] \tag{4.3}$$

Of course, since $[f_r]$ is invariant under the dynamical symmetries, it follows that any KPM either obeys (4.3) for all f_r or for none. Therefore, the laws are always either supertrue (true for all f_r) or superfalse (false for all f_r) on the symmetry-first approach. The dynamics partition the space of KPMs into dynamically possible and dynamically *im*possible models.

[10]Cf. the discussion of supervaluation in Chapter 2; Jacobs (2021).

But we can also use supervaluation to assess the status of specific relations between quantities. For example, the sentence

$$\frac{m_1}{m_2} = c \text{ (where } c \text{ is a constant)} \tag{4.4}$$

is invariant under mass scalings, and hence for any pair of particles it is either (super)true or (super)false that their mass ratio is equal to some constant c. It is often thought that only invariant statements are physically meaningful or 'objective'.[11] For instance, absolute simultaneity is not objective in Special Relativity because it is not Lorentz-invariant. On the symmetry-first approach, this is simply a consequence of the theory's supervaluationist semantics. Thus, assertions of mass ratios are meaningful on a symmetry-first account, whereas assertions of mass *differences*—which are neither supertrue nor superfalse—are physically meaningless. This forms the basis for symmetry-first approach's decision procedure for ontological commitment.

In order to formulate this decision procedure, I will assume the following criterion of ontological commitment: if one posits a certain structure, then one is committed to whatever one can define in terms of that structure. In a sense, all definable structure 'comes for free'. We can call this 'derivative' or 'conditional' commitment. This criterion is in the spirit of Butterfield's (2011) identification of supervenience with (implicit) definability. The criterion clearly departs from Quine's view that one is committed only to whatever entities one quantifies over. There are many properties and relations which are definable in terms of a theory's fundamental posits, but which are not themselves amongst those posits. Some of these properties and relations are famously problematic, for instance the property of being grue or that of not being black. The criterion of commitment which I propose implies that such properties are real, albeit in a derivative sense. It is consistent with this claim that blue, green and black are the more fundamental colour properties. The non-Quinean criterion of

[11] Cf. Mundy (1986); Debs and Redhead (1996); Nozick (2001); Saunders (2003b); Baker (2014).

derivative commitment is motivated by the idea that (implicitly) definable structure is *already there* in the theory: to use a surely worn-out metaphor, when God created the world's most basic structure, it didn't require any further act of creation to bring into being any structure which is definable from this basic structure. Therefore, the departure from Quine here is no surprise, since the symmetry-first approach explicitly does *not* quantify over the intrinsic structure of physical quantities. Instead, it defines this structure extrinsically in terms of the invariance under certain symmetries. Therefore, a non-Quinean criterion is apposite in order to draw out the symmetry-first approach's ontological commitments.

In terms of this criterion of conditional commitment, the question is this: which relations are definable in terms of the structure $\langle D, [f_r]\rangle$? I will now prove that (a) any symmetry-invariant piece of structure is (implicitly) definable from $\langle D, [f_r]\rangle$, and therefore (b) the ontological commitments of symmetry-first and structure-first sophistication are the same. I will rely on the following theorem due to Barrett (2017, Theorem 2).[12] Here, $\mathfrak{m}$ is a model of a theory T and R is some particular relation. We define $\Sigma = \mathbf{R} \setminus R$, where $\mathbf{R}$ is the set of relations R_i on the domain D_i of $\mathfrak{m}$. In other words, $\mathfrak{m}|_\Sigma$ is the *reduct* of $\mathfrak{m}$ obtained by removing R. Barrett's theorem then states:

> **If**, for any models $\mathfrak{m}, \mathfrak{n}$ of T, if $h : \mathfrak{m}|_\Sigma \to \mathfrak{n}|_\Sigma$ is an isomorphism then $h[R^\mathfrak{m}] = R^\mathfrak{n}$ **then**, for any models $\mathfrak{m}, \mathfrak{n}$ of T, if $\mathfrak{m}|_\Sigma = \mathfrak{n}|_\Sigma$ then $R^\mathfrak{m} = R^\mathfrak{n}$.

The if-then condition on the left-hand side captures the idea that some piece of structure R is *invariant* under the symmetries of a class of models. Suppose that we start with a model $\mathfrak{m}|_\Sigma$ and append some piece of structure R—for example, a new relation on the model's domain. We then consider the isomorphisms of the *original* model. If, under these isomorphisms, the extension of $R^\mathfrak{m}$ carries over to the extension of $R^\mathfrak{n}$, then R is invariant under the symmetries of $\mathfrak{m}|_\Sigma$. The if-then

[12] For a precursor of this theorem, see Earman (1989, 58-60).

condition on the right-hand side captures the idea that R is *implicitly definable* from $\mathfrak{m}|_\Sigma$, in the sense that any models which agree on the structure of $\mathfrak{m}|_\Sigma$ must also agree on the extension of R. Barrett's theorem states that a piece of structure is implicitly definable from the theory's solutions if it is invariant under the isomorphisms of these models. The theorem does *not* depend on the assumption that T is expressed in first-order logic (Barrett's proof of the reverse theorem *does* depend on that assumption, but see Dewar (2022, Prop. 1.16) for a purely model-theoretic proof). We can use the theorem even if we adopt the so-called semantic view of theories on which theories are specified directly by their class of models, as we have done here.

Here is an example to clarify the theorem. Consider once more models of NG of the form $\langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \rho \rangle$, where $\mathcal{V}_m = \langle D_m, [f_r] \rangle$. Denote these models $\mathfrak{m}|_\Sigma$. Now, consider the extra piece of structure $\circ$, which is a binary operator on D_m. The 'expanded' models $\mathfrak{m}$ are of the form $\langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \rho, \circ \rangle$. For the proof, suppose that if it is the case that $\langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \rho \rangle \to \langle \mathcal{M}, \mathcal{G}, \mathcal{V}_m, \phi, \psi \circ \rho \rangle$ is an isomorphism, then $\psi[\circ^{\mathfrak{m}}] = \circ^{\mathfrak{n}}$ (i.e. the LHS of Barrett's theorem). This is indeed the case for $\circ$, since $\circ$ is invariant under the automorphisms in Ψ. Furthermore, suppose that $\mathfrak{m}|_\Sigma = \mathfrak{n}|_\Sigma$ (i.e. the antecedent of the RHS). In that case, the identity map $I : \mathfrak{m}|_\Sigma \to \mathfrak{n}|_\Sigma$ is an isomorphism. So, from the LHS of Barrett's theorem, $I[\circ^{\mathfrak{m}}] = \circ^{\mathfrak{m}} = \circ^{\mathfrak{n}}$. Therefore, if $\mathfrak{m}|_\Sigma = \mathfrak{n}|_\Sigma$, then $\circ^{\mathfrak{m}} = \circ^{\mathfrak{n}}$, so $\circ$ is implicitly definable from $\mathfrak{m}$. The same holds for any other Ψ-invariant relations, such as the order $\leqslant$ defined above.

Models on either the structure-first or the symmetry-first approach are generally constructed such that the symmetries of the theory are isomorphisms between its models. This means that *exactly the same structure is (implicitly) definable on both approaches*. For example, we have just seen that the binary function $\circ$ is invariant under mass scalings, and hence implicitly definable in terms of $[f_r]$. If definability is our criterion for conditional ontological commitment, then the structure-first and the symmetry-first approach to sophistication have the same commitments in virtue of their invariance under the same group of symmetry transformations. Furthermore, we can now easily prove that given a class of *non-isomorphic* models with the same domain, symmetry-first sophistication is committed to *exactly that structure which is invariant under the theory's symmetry transformations.* After all, it is this structure which is definable from the theory's models. This explains the sense in which we can act 'as if' SRMs as isomorphic: we simply affirm commitment to their symmetry-invariant content. For an example, consider the shift from Newtonian to Galilean spacetime. We can prove that trans-temporal identities of spacetime points are variant under the boost symmetries of NG. It follows that a commitment to only that part of the structure of Newtonian spacetime which is invariant under Galilean transformations rules out the existence of trans-temporal identities for spacetime points. Importantly, this is the case even if no (structure-first) account of NG set on Galilean spacetime is offered.

As we can see, the symmetry-first approach is not quite a metaphysical black box. Nevertheless, there remains an important difference between the structure-first and the symmetry-first approach. For the structure-first approach presents us with a set of invariant relations which characterise the theory's value space in terms of the physical relations which hold between the relevant quantity's values. The symmetry-first approach, on the other hand, only provides us with a decision procedure to determine if it is committed to some given piece of structure. It then declares a blanket commitment to any such structure. But we don't know in advance which functions and relations *are* invariant under the theory's symmetries. The symmetry-first approach can tell us whether some theoretical claim is supertrue or superfalse, but it does not specify a list of fundamental posits. The symmetry-first approach is a little like a Freedom of Information request: in order to get the answers you are after, you first have to know the right questions to ask.

4.4 Aside: Implicit and Explicit Definability

Barrett's theorem shows that the symmetry-first approach's ontology is implicitly definable from the theory's invariant structure. The structure-first approach, on the other hand, *explicitly* defines its commitments, for example as a pair of relations $\leqslant$ and $\circ$ which obey certain axioms. The distinction between implicit and explicit definability is therefore helpful in analysing the difference between these approaches. Specifically, *Beth's theorem* states that structure is implicitly definable if and only if it is explicitly definable, where a relation R is explicitly definable iff there is a formula σ in the signature Σ of the theory T such that $T \models \forall x, y, ...(R(x, y, ...) \leftrightarrow \sigma(x, y, ...))$. Since structure-first sophistication demands an explicitly defined structure whereas symmetry-first sophistication only offers an implicit definition of the same structure, it may seem that the latter can use Beth's theorem to her advantage. After all, if implicit definability entails explicit definability, then there is not much of a difference between these approaches.

In this section I argue that this line of defence does not succeed. The structure-first approach does not just require that structure is (explicitly) defin*able*, but that it is actually *defined*. We do not just want to know that there *exists* a Σ-formula which defines R—we want to know the formula itself! This leaves the advocate of symmetry-first sophistication with a dilemma: either she does not explicitly define the theory's intrinsic structure, falling short of the structure-first demands for a perspicuous metaphysical picture; or she *does* define this structure, in which case the symmetry-first approach simply collapses into the structure-first approach. After all, the *raison d'être* of the former is that it is often difficult to explicitly define structures which are invariant under the appropriate symmetries: "The advantage of defining the new semantics externally [i.e. symmetry-first] is that it offers a relatively easy means of characterizing the objects of the semantics" (Dewar, 2019, 503). But if it is *always* possible to construct an explicit definition of these objects, then there is no reason to prefer a symmetry-first approach over a structure-first approach.

In the first place, it is questionable whether Beth's theorem even applies here. The theorem applies only to models of a first-order theory T. But the theories we have considered are not first-order theories. For example, (4.3) is clearly not a sentence of first-order logic, since it quantifies over various types of functions. Beth's theorem is unlikely to apply to any physical theories of interest. But suppose that we can set these concerns aside. Even then, the fact that some piece of structure is definable is not enough: what we are after is the actual definition of said structure. It turns out that Beth's theorem is constructive (Fitting, 1996), so there exists a procedure for constructing an explicit definition of any implicitly definable piece of structure (of course, Fitting's result is also limited to first-order logic). But there is a hitch: while Fitting proves that there exists a procedure for constructing the definition of a *given* invariant relation, this does not amount to a procedure for *finding* the invariant relations of a theory. For example, the constructive version of Beth's theorem (were it to apply to our version Newtonian Gravitation) implies that we can construct an explicit definition of $\leqslant$ and $\circ$ in terms of the f_r functions, but only once we already know the form of those relations. Unfortunately we cannot extract this information from a symmetry-first model.

This is the first horn of the dilemma: even if Beth's theorem applies in the context of real physical theories, this does not allow the advocate of the structure-first approach to explicate their implicit commitments. The second horn is no better. For if there *were* an algorithm by which one could find all the relations which are invariant under a given group of symmetry transformations, this would only count as a Pyrrhic victory for symmetry-first sophistication. For such an algorithm *ipso facto* yields a structure-first theory by allowing us to explicitly define a set of invariant relations which characterise a value space. If such an algorithm were to exist, then, structure-first sophistication is guaranteed to succeed. But the main motivation behind the symmetry-first approach is that it is easier to carry out: even Dewar admits that a structure-first theory, if available, is more perspicuous than a symmetry-first one (see §4.5). If a structure-first characterisation is not only always possible in principle but if there also exists an algorithm which allows us to define structure-first models, then there is simply no need for symmetry-first sophistication.

4.5 Perspicuous Metaphysics

We can thus say that ontological commitments are *definable* (either implicitly or explicitly) on the symmetry-first approach, but not always *defined.* The approach is simply committed to "whatever it is that is invariant under these transformations". Dewar is well aware of this 'feature' of his proposal:

> [A structure-first] presentation lets us see what it is we are committed to by our (qualified) realism about the theory; if we want to know the answers to specific questions about the nature of a theory's ontology and ideology, then this is invaluable. (Dewar, 2015, 326)

In response, Dewar has advocated a view he calls 'qualified realism', according to which we commit ourselves to symmetry-invariant structure without exactly knowing what it is that we are thereby committed to. On the one hand, the previous sections have shown that such qualified realism has more to it than its critics suggest; it is not, in Martens and Read's words, 'sophistry'. On the other hand, the symmetry-first approach *is* metaphysically opaque. From a realist point of view this is a problem, since the reason we engage in philosophical interpretation of scientific theories in the first place is to find out their metaphysical implications. For an empiricist it does not matter much whether spacetime is Newtonian or Galilean—the equations are the same, after all. But one of the philosophical achievements of the shift from the former to the latter is the metaphysically weighty conclusion that there is no trans-temporal identification of spacetime points.

In this section I explain in more detail why it is so important to have a 'perspicuous metaphysical picture', and hence why the symmetry-first approach fails as an interpretation of SRMs. I will present three criticisms: that symmetry-first sophistication does not provide us with causal explanations of physical phenomena (§4.5.1); that it assumes the representational equivalence of SRMs as a brute fact (§4.5.2); and that its account of reality involves physically inert structure (§4.5.3). I claim that these issues arise because symmetry-first sophistication does not characterise a theory's models *intrinsically*, in the sense of Field (1980). These arguments thus support a 'motivationalist' view of symmetries according to which we are only allowed to identify SRMs once we have a perspicuous explanation of their equivalence (Møller-Nielsen, 2017). The symmetry-first approach is unsatisfactory exactly because it fails the motivationalist's test. The structure-first sophistication *does* offer a perspicuous characterisation of the physical content of SRMs, and so details one way in which one can satisfy the motivationalist's demands.

4.5.1 Causal Explanations

We are not just interested in a theory's ontology for its own sake. We also use a theory's ontological posits to offer explanations of physical events. As Møller-Nielsen argues, "it is simply not clear what causal-explanatory, realistic picture of the world is being propounded by the defender of the interpretational view" (Møller-Nielsen, 2017, 1264). (The 'interpretational view' Møller-Nielsen refers to assumes, in line with symmetry-first sophistication, that one can interpret SRMs as physically equivalent *before* one has a perspicuous metaphysical picture on hand). The first worry then is that without clear ontological commitments, attempts at scientific explanation are obstructed.

It is unclear, for example, how symmetry-first sophistication can explain the Aharonov-Bohm effect, in which a charged particle that moves around an impenetrable solenoid picks up a phase which is proportional to the flux through the solenoid. The proponent of symmetry-first sophistication cannot without further ado appeal to the fact that the *holonomies* of the vector potential field are invariant under gauge transformations, since this first requires an expression of holonomies in terms of a representation function from spacetime points to complex numbers. Of course, once one has such an expression it follows from Barrett's theorem that the symmetry-first approach is committed to holonomy-structure. But this is a non-trivial conclusion which goes beyond the minimal commitments of Dewar's qualified realism.

A second example concerns the structure of Maxwellian spacetime, which includes so-called dynamic shifts as spacetime symmetries.[13] Earman (1989) originally defined the structure of Maxwellian spacetime in terms of an equivalence class $[\nabla]$ of covariant derivative operators. But none of the elements of this class are supposed to represent geometrical structure (since each individual ∇ varies under dynamic shifts), so it is unclear what the physical explanation of a particle's trajectory in Maxwellian spacetime consists of. In essence, Earman defines the standard of rotation as 'whatever it is that is common to all $\nabla \in [\nabla]$', but this is not a closed definition of any geometrical object. In contrast, Weatherall (2017) offers an intrinsic characterisation of the standard of rotation. This allows us to define the structure of Maxwellian spacetime in a structure-first way. Although no one has yet expressed the *laws* of NG in terms of this object, Weatherall shows that one can use his standard of rotation to define a notion of relative acceleration which could enter in explanations of physical phenomena such as Newton's famous bucket experiment.

4.5.2 Unexplained Equivalences

The second criticism of the symmetry-first approach is discussed by Martens and Read (2020): "it is often unclear what grounds or explains or justifies the physical

[13] I thank James Read for suggesting this example.

equivalence of models that, on a natural interpretation, represent distinct possible worlds" (26).[14] We can put the criticism as follows. *Prima facie*, SRMs seem to represent distinct possible worlds; for example, worlds in which particles have different absolute velocities. Sophistication nevertheless claims that SRMs are physically equivalent. We can explain this equivalence in terms of an underlying structure, such as Galilean spacetime: since any model of Newtonian Gravitation set on Galilean spacetime is embeddable within an equivalence class of SRMs of Newtonian Gravitation set on Newtonian spacetime, each model within such a class represent the fundamental Galilean structure of the world equally well.

But on the symmetry-first approach it is simply *assumed* that SRMs are physically equivalent; symmetry-first structures are then *defined* as structures which are invariant under certain symmetries. This means that the symmetry-first approach does nothing to *explain* these equivalences. Specifically, it does not come with a metaphysical picture which explains the fact that the world is insensitive to certain transformations, such as velocity boosts or mass scalings. The symmetry-first approach can at best claim that the invariant structure of our models, *whatever it is*, explains this equivalence. That is not an explanation so much as a promissory note.

4.5.3 Intrinsic Structure

Finally, symmetry-first sophistication fails to give an *intrinsic* characterisation of physical structure. This Fieldian objection has not been as prominent in the literature so far, but captures an important source of dissatisfaction with the symmetry-first approach.[15] It seems to me that Field was concerned with exactly the kind of questions that are at issue in debates around symmetries, but this connection has largely gone unnoticed. The core objection is that the symmetry-first approach characterises the physical content of models in terms of the mathematical relations *between* those

[14]See Earman (1989, 171) for a similar point.
[15]Sider's (2020) discussion of what he calls 'quotienting' is an exception.

models; for instance, their isomorphisms. But these relations themselves are unphysical, as they merely relate abstract mathematical structures. The worry is that a definition of the physical content of a theory's models in terms of the non-physical relations between these models is in some sense 'impure'. It is as if someone were to characterise the referent of the name 'Kripke' as "the man whom we call 'Kripke'". Perhaps this fixes the correct extension for the name 'Kripke', but it doesn't seem to tell us anything about who Kripke really is.

Field's intrinsicalist programme requires instead that we characterise a theory's structure in purely physical terms. We are required to lay down a set of relations, functions and operators which explicitly represent the world's physical structure. Let me stress that my aim here is not to resurrect Field's programme in its entirety. Instead, I want to emphasise his insight into the importance of intrinsically-defined theories, which is under-appreciated in the contemporary literature. Why accept Field's requirement of intrinsicality? We can extract a number of motivations from his seminal book *Science Without Numbers* (Field, 1980). The first is a commitment to nominalism about mathematical entities. However, as Chen (2018, fn. 7) has pointed out, "the intrinsicalist and nominalistic visions can also come apart. For example, we can, in the case of mass, adopt an intrinsic yet platonistic theory of mass ratios." Since mass ratios, which are identified with positive real numbers, are invariant under mass scalings, they do not depend on arbitrary conventions. They intrinsically characterise the second-order relations between determinate mass values. It is not the appeal to numerical structure *per se* that poses a problem for symmetry-first sophistication.

The second motivation is that there is something arbitrary about the coordinate functions f_r, in the sense that *any* $f_r \in [f_r]$ represents a quantity's structure equally well. (That, of course, is just the content of the uniqueness theorems discussed in §4.1). The intrinsicalist's aim, on the other hand, is "to explain, in terms of intrinsic facts [...] which are statable without such arbitrary choices, why the choice of functions to be invoked in the extrinsic theory will be arbitrary to precisely the extent that it is" (Field, 1980, 46). But although each individual f_r is arbitrarily chosen, it is not the case that the equivalence class $[f_r]$ is arbitrary as a whole. After all, $[f_r]$ is the unique equivalence class of representation functions that is invariant under the dynamical symmetries of the theory, as we saw in §4.2. It is therefore not obvious that symmetry-first structures as a whole suffer from arbitrariness, although it is still plausible that a collection of individually arbitrary functions is unsatisfactory from a realist's point of view.

There is a third Fieldian concern which I believe to be more basic, namely the claim that the functions in $[f_r]$ are not *physically* relevant. The f_r's are functions from (for example) mass value space into some system of units, such as the positive real numbers. But units themselves play no causal role; they are physically inert. Field expresses this worry most clearly when he writes about the gravitational constant G, the numerical value of which arbitrarily depends on a choice of units:

> The role [G] plays is as an entity extrinsic to the process to be explained, an entity related to the process to be explained only by a function (a rather arbitrarily chosen function at that). Surely then it would be illuminating if we could show that a purely intrinsic explanation of the process was possible, an explanation that did not invoke functions to extrinsic and causally irrelevant entities. (Field, 1980, 44)

I believe that this is the crucial defect of the symmetry-first approach: it appeals to physically irrelevant functions in order to characterise the structure of physical quantities such as mass or spacetime.[16] The symmetry-first approach is metaphysically opaque because we cannot interpret the f_r's themselves as physically meaningful, since they only encode facts about the representational equivalence of the theory's

[16] In focusing on physical irrelevance instead of arbitrariness, I differ from Wolff (2020) and Sider (2020); but we all agree that what is problematic about views like the symmetry-first approach is their appeal to representation functions.

models.

It is exactly because the symmetry-first approach fails the requirement of intrinsicality that it cannot explain the physical equivalence of SRMs. After all, the symmetry-first approach aims to extract physical content from the *assumption* that SRMs are physically equivalent, and hence an explanation of that equivalence in physical terms becomes viciously circular. The failure of intrinsicality likewise means that a symmetry-first account is unable to offer causal explanations: only intrinsic structure is causally efficacious, but symmetry-first sophistication does not specify this intrinsic structure. The relations between mathematical models which the symmetry-first approach uses to extract physical content have no causal effects. The issue that lies at the foundation of symmetry-first sophistication, then, is that it tries to characterise the physical world in terms of formal relations between representations of the world. This yields only an indirect picture of the world itself; and it is one whose content is not readily surveyable.

4.6 Conclusion

For the reasons above, only structure-first sophistication provides a satisfactory interpretation of symmetry-related models. Nevertheless, symmetry-first sophistication presents us us with a powerful mathematical apparatus. I conclude with the suggestion that the latter approach may be useful as an *interim* interpretation of SRMs. We can use the symmetry-first approach as a vantage point from which to find an intrinsically-defined structure. The strategy is to assume that there exist invariant structures in terms of which one can specify the laws, and use educated guesses to find them. Indeed, Dewar mentions a similar proposal:

> Assuming that one accepts the [symmetry-first] method of definition as mathematically legitimate, then its application gives us a way of defining a sophisticated semantics for the theory, by brute force. It then means

> that we do have a precise target for a sophisticated semantics that is internally defined: we are looking for some [structure-first] construction which delivers an equivalent class of structures. (Dewar, 2019, 504)

As a methodological half-way house, symmetry-first sophistication is unobjectionable. But the interpretation of symmetry-related models is only really finished once we have a perspicuous picture of their physical content in purely intrinsic terms. To achieve that aim, structure-first sophistication is required.

Chapter 5

Quantities and Symmetries

In the previous chapter, I argued that we ought to define a theory's models explicitly in terms of those functions and relations which represent physical structure. I called this approach *structure-first sophistication*. In the following chapters I will present a pair of detailed case studies of structure-first sophistication for internal symmetries (both global and local). The main aim of this chapter is to lay down the formal groundwork for these case studies. I will define a few central concepts, namely: (i) *quantities*; (ii) *structure*; and (iii) *symmetries*. I will also discuss the distinction between *internal* and *external* symmetries. Earman (1989) famously laid down a pair of principles for external (or spacetime) symmetries; I will extend these principles to cover the class of internal symmetries. The relevance of these symmetry principles for our purposes is that, as I will argue, any sophisticated theory must satisfy them. The principles thus play an important role in what will follow.

This chapter also has a subsidiary, more metaphysical aim. The claim I wish to defend is that the traditional distinction between a theory's ontology and ideology is superseded by the distinction between internal and external degrees of freedom. This latter distinction is not absolute, but relative to the quantities a theory posits. It is not a distinction of kind: the same quantity may count as internal in one theory, but as external in another. In this way, an investigation into the structure of our best physical theories leads to a reconfiguration of our traditional metaphysical categories. The secondary aim of this chapter is thus to engage in what Ladyman et al. (2007) call *naturalised metaphysics*, in particular the metaphysics of quantities.

5.1 Quantities and Structure

5.1.1 Ontology and Ideology

Let's start with Quine's (1951) distinction between ontology and ideology. The former concerns what exists; the latter what we can say about that which exists. For instance, the Eiffel Tower exists; we can say *of* the Eiffel Tower that it is made of iron, or that it is impressive. Although Quine's terminology seems to suggest that ideology is concerned with ideas in the mind, Sider (2011) has recently emphasised that it is as objective a subject matter as ontology. The external world itself dictates which concepts are needed to express the physical facts.

Quine argued that we can read off a theory's ontology from its formalism: "the ontology to which an (interpreted) theory is committed comprises all and only the objects over which the bound variables of the theory have to be construed as ranging in order that the statements affirmed in the theory be true" (Quine, 1951, 11). Of course, this criterion of ontological commitment is controversial, but that is not the point I wish to address here. The question is whether we can formulate a similar criterion for a theory's ideology. Quine criticises a proposal due to Gustav Bergmann, namely that the ideology to which a theory is committed comprises all and only the properties over which the bound *predicate variables* of the theory have to be construed as ranging in order that the statements affirmed in the theory be true. Quine's response is that the theory need not quantify into predicate position at all; as is well-known, Quine had a predilection for first-order theories. But it remains a question of interest which concepts are necessary to formulate an empirically adequate theory, so we still need a criterion for membership of the theory's ideology. Quine's own suggestion is that a theory's ideology simply consists of the concepts or ideas expressed by the theory's predicates. Moreover, Quine points out that one can define said concepts either extensionally or intensionally; he himself of course prefers the former. Below, I will investigate whether this suggestion still makes sense for modern physical theories.

5.1.2 The Problem of Quantities

In the language of model theory, a *model* (or interpretation) of a theory T in a signature Σ consists of a domain of objects and a function which assigns an extension to each of Σ's names and predicate symbols. We will assume without loss of generality that the theory does not contain names, since we can always replace each name with a unary predicate (Quine, 1960; Lewis, 1970). The models of T are then of the form:

$$\langle D, P_1, P_2, ..., P_n \rangle$$

where the P_i range over the theory's predicates, broadly understood to cover both monadic and polyadic predicates (i.e. relations). The elements of D over which T's quantifiers range are the ontology; the concepts expressed by the predicate symbols P_i are the ideology.

For illustrative purposes, consider a theory of classical particles which move under the force of gravitation. The ontology of this theory consists of

(i) a collection $\mathcal{B}$ of massive bodies; and

(ii) a collection P of spacetime points.

The ideology of the theory consists of

(i) a collection $\mathbf{D}$ of relations between spacetime points;

(ii) a collection of mass properties $\mathbf{M}$ (such as the property of being 1 kg); and

(iii) a position function $x(i)$ from particles to points.

The models of this theory are of the form:[1]

$$\langle \mathcal{B}, \mathrm{P}, \mathbf{D}, \mathbf{M}, x \rangle$$

I will leave the relations contained in $\mathbf{D}$ aside, except to note that they must satisfy certain axioms in order to represent a classical spacetime. I will focus instead on a problem of a different kind.[2] Consider a universe which contains two particles: one's mass is 1 kg, and the other's mass is 10 kg. We would like to say that the second particle is more massive than the first. However, none of the theory's resources allow us to say this: the relation 'more massive than' is not within the theory's ideology. In particular, note that although we have associated mass properties with real numbers, this association is merely conventional: the elements of $\mathbf{M}$ are primitive predicates. Therefore, we cannot say that the 1 kg particle is less massive than the 10 kg particle simply because 1 is less than 10. Intuitively, it seems that the first particle is less massive than the second particle in virtue of the monadic mass property each particle has, but the theory as it stands does not reflect this fact.

It would seem as if there is an easy fix: introduce a new symbol $\leqslant$ which expresses the relation 'less than or equally massive as'. But the introduction of a new relation by itself does not suffice, since we must also introduce new axioms to ensure that $\leqslant$ is well-behaved. For instance, we need to stipulate that if particle a is 1 kg and particle b is 2 kg, then $a \leqslant b$; and likewise that if particle a is 1 kg and particle b is 3 kg, then $a \leqslant b$; and so on. Clearly, there is no end to this: since there are

[1] It is convention that the first element of a model is the domain. But since our theory consists, in a sense, of multiple domains, I will list them separately.

[2] For a clear statement of this problem, see Bigelow and Pargetter (1988). I should note that much of the below debate actually took place in the context of spacetime, but exactly for that reason I prefer the less muddled case of a quantity such as mass.

infinitely many mass properties, we will also need infinitely many axioms for $\leqslant$. But this means that our theory, once it includes $\leqslant$, is not recursively axiomatisable! It is a basic requirement of a successful theory that we can contain it within the pages of a textbook, but this is patently not the case for the toy theory of classical mechanics we have formulated here. Field (1984) calls this the 'problem of quantities'. Field also notes that we can solve the problem via an appeal to 'heavy duty Platonism': identify the mass properties with the positive real numbers. We can then simply say that one mass is smaller than another iff it is associated to a smaller real number. But this solution is unsatisfactory. While I do not share Field's commitment to nominalism, I agree that physical properties are structurally different from the real numbers, for reasons explained in the next section. We need a different solution to the problem of quantities.

Nominalism

We have hit upon a junction with two exits. On the one hand, we can start again from scratch and see if we can present our theory in terms of a different ideology. This is the nominalist project of Field (1980). The term 'nominalism' here refers not just to a denunciation of numbers and other abstracta, but also of universals; that is, we are not allowed to quantify over properties (whether the quantification is first- or second-order). Instead of the set $\mathbf{M}$ of mass properties, Field introduces a pair of mass *relations*: a ternary *betweenness* relation, $zBxy$, and a quaternary *congruence* relation, $xyCvw$. Field proves a *representation theorem* which shows that it is possible to present a finite set of axioms for these relations such that we can associate each particle with a positive real number via an intuitively well-behaved function f. For instance: if $f(a) = 1$, $f(b) = 2$ and $f(c) = 3$, then b is between a and c. The numbers here seem to play the role of monadic mass properties, but unlike the predicates in $\mathbf{M}$ they are not part of the theory's fundamental ideology, which consists only of the relations B and C. This is reflected in the fact that the association between particles and numbers is not unique. In particular, Field also proves a *uniqueness theorem*, which shows that if f is a well-behaved representation function then so is αf, where α is any positive constant.

However, Field's axiomatisation is successful only when a certain 'richness' assumption is satisfied, namely that for any pair of particles a and c there exists another particle b whose mass is between a and c. In effect, Field defines quantitative relations—twice as massive as, equally massive to—via quantification over particles. For instance, the mass ratio between any pair of particles a and b is twice that between any (distinct) pair of particles c and d iff there exists a particle z such that (i) $zBab$, (ii) $azCzb$ and (iii) $zbCcd$. If there is no such particle z, then this definition is trivially false. But the assumption that there always exist particles which satisfy these constraints is unreasonable, since it rules out worlds which contain only a finite number of particles, or worlds in which quantities such as mass or charge come in discrete quanta. For this reason, I will not follow Field's nominalist project, but consider the alternative route.[3]

Platonism

The second route is to relinquish our commitment to nominalism. If we are allowed to quantify over the elements of $\mathbf{M}$, then we *can* finitely axiomatise structural relations such as $\leqslant$. For our requirement is just that $\leqslant$ defines a (total) order on the set $\mathbf{m}$, which means that $\mathbf{m}$ must satisfy a total of three axioms: antisymmetry, transitivity and connexity. For instance, antisymmetry says that

$$\forall m_1, m_{2 \in \mathbf{M}} (m_1 \leqslant m_2 \rightarrow m_2 \nleqslant m_1)$$

[3]Field's own response is that for physically realistic field theories the richness assumption is satisfied, since fields are continuous functions over the manifold M. But this response hasn't convinced many. For further discussion, see Arntzenius and Dorr (2012) and Eddon (2013a).

On this view, the explanation of the fact that a 1 kg particle is less massive than a 10 kg particle is that the relation 'less than' relation holds between the respective mass properties in question. But of course, as soon as we quantify over the elements of $\mathbf{M}$, by Quine's own criterion these elements become part of the theory's ontology, rather than its ideology. In addition to massive particles and spacetime points, our theory is now committed to the existence of an infinitude of mass magnitudes. This inflation of the theory's ontology may seem unparsimonious, but I suggest that we embrace it. For the inflation is not gratuitous, since it allows us to express certain facts about relations between mass values which are necessary for an adequate formulation of classical particle mechanics. I say this in a Quinean spirit: if our best theory quantifies over mass magnitudes, then mass magnitudes exist. This sort of view has been defended by Mundy (1987) and more recently by Eddon (2013a); Arntzenius and Dorr's (2012) 'locationism' is a closely related position. The latter stress the similarity between physical magnitudes and spacetime locations: in each case, particles are said to 'occupy' or 'instantiate' a point or magnitude respectively, and both points and magnitudes stand in certain relations. Crucially, both spacetime points and physical magnitudes are now considered ontological posits of the theory.

We have arrived at a theory whose ontology consists of

(i) a collection $\mathcal{B}$ of massive particles;

(ii) a collection P of spacetime points; and

(iii) a collection $\mathbf{M}$ of mass magnitudes;

and whose ideology consists of

(i) a collection $\mathbf{D}$ of relations between spacetime points;

(ii) a collection $\mathbf{R}$ of relations between mass magnitudes (e.g. $\leqslant$);

(iii) a position function $x(i)$ from particles into points; and

(iv) a mass function $m(i)$ from particles into mass magnitudes.

The introduction of a mass function in place of mass predication is a matter of convenience which incurs no loss of generality; see Lewis (1970, 429) for a similar move. To further clean up our notation, we introduce the shorthand $\mathcal{M} \coloneqq \langle \mathrm{P}, \mathbf{D} \rangle$ and $\mathcal{V} \coloneqq \langle \mathbf{M}, \mathbf{R} \rangle$. With this notation in place we can express the models of our theory as follows:

$$\langle \mathcal{B}, \mathcal{M}, \mathcal{V}, x(i), m(i) \rangle$$

I will refer to the theory with models of this form as classical particle mechanics.

The desire to draw mass comparisons has led us to a formalism which blurs the traditional distinction between ontology and ideology in classical particle mechanics. In particular, mass properties are usually thought of as part of the theory's ideology. Martens (2019b), for instance, defines classical particle mechanics as a theory with "an ontology of n particles in a flat Newtonian space and time, with an ideology of absolute distances, velocities and finite, positive, non-zero masses" (2511, reference and footnote suppressed). But in our formalism, mass properties are instead part of the theory's ontology. Sider (2011, Ch. 2) mentions the shift from distinct electric and magnetic fields to a unified electromagnetic field as a matter of ideology. But on the assumption that fields are represented by functions from spacetime to the space of field values, the move from a pair of distinct value spaces to a single value space also indicates a novel ontology. Instead of physical properties, the ideology of classical particle mechanics consists of two types of concepts: firstly, the structural relations between spacetime points and mass properties respectively; and secondly, the position and mass functions. I will call the former *structure* and the latter *quantities*. Thus, within the category of ideology we have structure, which consists of relations between collections of the theory's ontology, and quantities, which are functions between different such collections. This means that Sider's (2011) second example of ideological change—the shift from classical to relativistic spacetime—*is* correct, since it illustrates a change in spacetime structure, which by the above is part of a theory's ideology.

5.1.3 Local Spacetime Theories

Of course, classical particle mechanics in the way I have formulated it here is only one theory—and not a very realistic one at that! But we can draw similar conclusions for the quantities of different theories, such as colour or charge. For instance, in order to say that one particle is oppositely charged from another particle we need to introduce the relation 'oppositely charged'. But this relation is well-behaved only if it partitions the charge properties into disjoint sets: the positive and the negative ones. We can finitely axiomatise this relation only if we are allowed to quantify over charge properties themselves. Therefore, the conclusions drawn for particle mechanics generalise. Specifically, we will suppose that the models of any physical theory consist of:

1. a collection of structured sub-domains $\mathcal{D}_i := \langle \mathrm{D}_i, \mathbf{R}_i \rangle$; and
2. a collection of functions ϕ_i between these sub-domains.

We will then define *structures* as sets of relations $\mathbf{R}_i$ on the theory's sub-domains; and *quantities* as functions ϕ_i between them. Furthermore, the co-domain of each quantity is called a *value space*.[4] We can draw a connection to the distinction between determinables (mass, charge, colour) and determinates (5 kg, $1.67 \cdot 10^{-19}$ C, scarlet red).[5] In particular: each value space represents a determinable, and each element of a value space represents a determinate of that determinable.

[4]The idea of a structured value space for physical quantities is found, in different forms, in Stalnaker (1979); Gärdenfors (2000); Denby (2001); Funkhouser (2006); Caulton (2015).

[5]For more on the determinable/determinate distinction, see Wilson (2017).

I claim that this framework is sufficiently general to encompass most contemporary physical theories. In particular, it encompasses what Earman and Norton (1987) call *local spacetime theories*: theories

(a) whose models are of the form $\langle M, O_1, ..., O_n \rangle$, where M is a differentiable manifold and $O_1, ..., O_n$ are geometric objects defined over M; and

(b) whose equations of motion are of the form $\sum_i c_i O_i = 0$; and

(c) which satisfy the 'completeness condition' that *any* model of the form $\langle M, O_1, ..., O_n \rangle$ that satisfies the theory's equations is a solution of the theory.

For instance, classical *field* mechanics is a local spacetime theory whose models are of the form $\langle M, t_{ab}, h^{ab}, \nabla, \varphi, \rho, \xi^a \rangle$. The dynamical equations of this theory are listed in Chapter 1. Note, however, that local spacetime theories do not include theories whose fields are represented as sections of fibre bundles, such as (a certain formulation of) scalar electrodynamics. The reason is that such sections are not geometric objects: their fibre bundles are not 'natural' bundles (cf. Dewar (2020a)). I discuss such theories in Chapter 7.

In order to finesse this theory into our preferred format, a few interventions are required. Firstly, note that the geometric objects O_i fall into two classes: absolute objects A_i which represent spacetime structure, and dynamic objects P_i which represent physical fields. In the case of classical field mechanics, the absolute objects are t_{ab}, h^{ab} and ∇, while the dynamic objects are φ, ρ and ξ^a. We hence let $\mathcal{M} := \langle M, A_1, ..., A_n \rangle$, so $\mathcal{M}$ denotes a spacetime manifold (ontology) *with* associated spacetime structure (ideology). Secondly, the dynamic objects P_i are functions from the manifold M into some value space. For instance, the matter field ρ is a function from M into some mass density value space $\mathcal{V}$. We can imagine theories whose dynamic objects are not functions. But this is not the case for local spacetime theories in physics, from classical mechanics to general relativity.[6] Finally, we ask: functions into *what*? For Earman and Norton, dynamic objects are functions into n-tuples of real numbers, i.e. elements of $\mathbb{R}^n$. Put differently, Earman and Norton equate the structure of a value space with that of the space of real numbers. As I will argue in the next section, I believe that this is a mistake: the space of real numbers has *more* structure than the space of dimensionful values. But we can remain neutral on this issue for now by assigning the label $\mathcal{V}_i$ to the value space of quantity ϕ_i. Note that Earman and Norton do not list value spaces as separate entries in a theory's models. I will include them for completeness' sake. This suffices to show that the framework we have developed encompasses local spacetime theories in the sense of Earman and Norton.

We started with the suggestion that the values of physical quantities are part of a theory's ideology, in the sense of Quine (1951). But we have seen that quantities must instantiate relations between them in order for us to draw comparisons between their values. Therefore, values are part of the theory's ontology rather than ideology. Although we can still draw a distinction between ontology and ideology—where a theory's ideology consists of (i) structure and (ii) quantities—the boundaries have become blurred. I will argue in §5.4 that the distinction between a theory's *internal* and *external* degrees of freedom (DOFs) will prove an adequate substitute for the classical ontology/ideology dichotomy. But first I will address a question raised in the previous paragraph—whether value space structure is isomorphic to $\mathbb{R}$—in more detail.

5.2 Interlude: Measurement Theory

The previous section argued that quantities' value spaces possess structure. Before moving on to the metaphysical claim of this chapter—that the distinction between

[6]Whether quantum theories fit this mould is a different issue. Perhaps Wallace and Timpson's (2010) 'Spacetime State Realism' does the job, but that is a question for another time.

internal and external DOFs in some sense supersedes the one between ontology and ideology—I want to dispel the following myth: that the structure of a quantity's value space is just that of the real numbers. If that were the case, the question *which* structure a value space has would become theory-independent. But this is not the situation we are in: although we can use numbers to *represent* the values of physical quantities, the latter usually have less structure than the former, in a sense to be made precise below. In this section, I will use the resources of the field of *measurement theory* to substantiate these claims.[7]

It is not uncommon to see quantities defined as functions from some domain into the real numbers $\mathbb{R}$ (or some generalisation thereof, such as $\mathbb{R}^n$ or $\mathbb{C}$).[8] For instance, van Fraassen (1989, 278) writes that "a (physical) quantity is (represented by) a function of a special sort: its arguments are the 'bearers' of that quantity, which are in general not mathematical entities at all. [...] *The values generally are mathematical entities: a temperature is a real number, a position or velocity is a vector (which is a finite sequence of numbers)*" (emphasis mine). Similarly, foundational treatments of local spacetime theories standardly define ρ and ϕ as scalar fields, that is, continuously differentiable functions from M into $\mathbb{R}$. But while numbers play a crucial role in calculations, they are not well-suited to represent the *physical* structure of quantities. The problem is that $\mathbb{R}$ is too highly structured; not all of this structure represents physical relations. For an example, consider the coordinatisation of time on the real number line. As Arntzenius and Dorr (2012) point out, while one can add up real numbers, it doesn't seem to make much sense to add up moments in time. So some of the structure of the real numbers—in this case, their additive structure—does not represent a genuine physical relation between moments of time. Likewise, it is usually

[7]Krantz et al.'s (1971) three-volume *Foundations of Measurement* is the *locus classicus*; see Wolff (2020) for a recent overview.

[8]Here and in what follows, by 'the real numbers' I will understand the Dedekind-complete ordered field $\langle \mathrm{R}, +, \times, < \rangle$; the positive reals are the complete ordered semi-field $\langle \mathrm{R}^+, +, \times, < \rangle$. For more detail, see Chapter 6.

thought that there is no meaningful notion of absolute difference between two masses, even though the difference between positive real numbers is well-defined (Martens, 2019b; Wolff, 2020).

The fact quantities have less structure than the real numbers (in a sense that will become more precise below) is a consequence of the fact that they are *dimensionful*: they are always expressed in terms of a system of units. The fact that different unit systems represent a certain quantity equally well implies that the differences between these systems are physically irrelevant.[9] In the case of mass, for instance, the numerical value of the difference $m_1 - m_2$ between masses depends on a choice of units: if the difference is 1 in kilograms, then is it approximately 2.2 in pounds. If mass were to have the structure of $\mathbb{R}^+$, then there would be a fact of the matter as to which number *correctly* expresses the difference between any pair of masses. In effect, the structure of $\mathbb{R}$ encodes a choice of units. Intuitively, then, dimensionful quantities have less structure than the real numbers in the sense that many different assignments of real numbers to the values of these quantities are equally good for the purposes of physics.

The central ideas of measurement theory are familiar from standard treatments of differentiable manifolds. Specifically, an n-dimensional differentiable manifold $\mathcal{M}$ is usually understood to have less structure than $\mathbb{R}^n$, in the sense that there exist many equally good mappings from $\mathcal{M}$ into $\mathbb{R}^n$. Nevertheless, it is possible to numerically represent $\mathcal{M}$ via an atlas of coordinate functions into $\mathbb{R}^n$. Furthermore, if our spacetime has some additional structure, then we can require that the coordinate functions captures this structure. For example, Newtonian spacetime has a standard of rest in the form of σ^a. We can numerically represent this structure by stipulating that $\vec{x}(x) = \vec{x}(y)$ if and only if there exists an integral curve of σ^a from x to y, where $\vec{x}(p)$ is a three-dimensional vector of the spatial coordinates of a point $p \in \mathcal{M}$. For

[9]Cf. Baker's (2014) discussion of scale-independence; see also Martens (2019b).

a given Newtonian spacetime, there are many coordinate systems which satisfy this requirement: if $\vec{x}$ is one such set of coordinates, then so is $\vec{x}' = \vec{x} + \vec{a}$ where $\vec{a}$ is a constant vector. So, there exists a numerical representation of σ^a on $\mathbb{R}^3$; and this representation is unique up to the addition of a constant vector. The spacetime $\mathcal{M}$ has less structure than $\mathbb{R}^n$ in the sense that there exist many coordinate charts which represent $\mathcal{M}$ equally well; formally, we say that there is no unique homomorphism between $\mathcal{M}$ and $\mathbb{R}^n$.[10]

The field of measurement theory is interested in these sorts of results. Consider, for example, so-called *additive structures* of the form $\mathcal{V} := \langle D, \leqslant, \circ \rangle$, where D is a set of cardinality $2^{\aleph_0}$, $\leqslant$ imposes an order on D and $\circ$ is a binary operator on D. These relations must satisfy certain axioms, for instance that $\leqslant$ is a complete linear order.[11] When the axioms are satisfied, we can prove that there exists a function $f_r : D \to \mathbb{R}^+$ such that (i) $m_1 \leqslant m_2$ iff $f_r(m_1) \leq f_r(m_2)$ and (ii) $f_r(m_1 \circ m_2) = f_r(m_1) + f_r(m_2)$. This is called a *representation theorem*, since it shows that we can represent the structure of $\mathcal{V}$ on the (positive) real numbers. Put differently, there exists a systematic mapping from $\mathcal{V}$ to $\mathbb{R}^+$. This is analogous to the existence of a coordinate system adapted to a particular spacetime structure. However, this representation is not unique: one can also prove that f_r represents $\mathcal{V}$ iff any function f_r' such that $f_r'(m) = \alpha f_r(m)$ also represents $\mathcal{V}$, where α is a positive real constant. This is called a *uniqueness theorem*: it shows that our representation functions are unique up to certain transformations. The transformations between representation functions

[10]This conception of structure is closely related to Klein's Erlangen project; for more on this connection, see Wallace (2019c). I do not claim that this is the only way to 'count' structure. For instance, Barrett (2020) proposes that one space has more structure than another iff the former space's automorphism group is a subset of the latter's. On this view, some spacetimes—for instance, a Euclidean spacetime with a preferred origin and orientation—have the same structure as the real numbers, since both have trivial automorphisms. Nevertheless, one can still coordinatise such spacetimes in different ways; in particular, there is no requirement to label the preferred origin with the zero vector. But my claim does not stand in tension with Barrett's: I emphasise one way in which dimensionful value spaces have less structure than the real numbers, but there may be other standards by which they have the same amount of structure.

[11]There is no need to list these axioms here; see Hölder (1901) or Krantz et al. (1971, 73) for a complete exposition. I discuss these axioms in Chapter 6.

correspond to changes in units—for instance from grams to kilograms. The uniqueness theorem thus proves that unit systems for mass are related by uniform scalings. This is analogous to the result that coordinate systems adapted to Newtonian spacetime are related by uniform translations and scalings. Uniqueness theorems offer an explication of what it means to say that $\mathbb{R}^+$ has more structure than $\mathcal{V}$, namely that there is no unique homomorphisms between them.

We can summarise these results in the form of a *three-level view of scientific representation.* This is in contrast with a two-level view on which real numbers directly represent physical values. But since $\mathbb{R}$ has more structure than a dimensionful value space, the two-level view fails to represent the latter structure accurately. On the three-level view, on the other hand, *non-numerical* structures represent the structure of physical quantities. We will call such structures *theoretical structures.* We can then use representation and uniqueness theorems to show that theoretical structures are in turn represented by certain numerical structures. In more detail, at the bottom level we posit the structure of the world, $\mathcal{W}$, which it is our aim to represent. (There are well-known problems with the idea that the world has a structure that our theories aim to represent, viz. Putnam's Paradox. But we can safely put this issue aside here, since our main interest is in the relation between theoretical and numerical structures, rather than the relation between theoretical and worldly structures.) In our example, $\mathcal{W}$ consists of all physical mass values and the relations between them. The middle level then consists of a theoretical structure such as $\mathcal{V}$. When this structure accurately represents $\mathcal{W}$, there exists an *interpretation*, f_i, between $\mathcal{V}$ and $\mathcal{W}$, which is analogous to a structure-preserving map between mathematical structures. The real numbers reappear at the top level, which contains richer numerical structures. These are related to the theoretical structure via *representations*, f_r. The uniqueness theorems show that the functions f_r form an equivalence class of representations.

The three-level view is illustrated in the following diagram:

$$\begin{tikzcd}
\langle \mathbb{R}^+, \leq, + \rangle \\
\mathcal{V} = \langle D, \leqslant, \circ \rangle \arrow[u, "f_r : D \to \mathbb{R}"'] \arrow[d, "f_i : D \to \mathcal{W}"] \\
\mathcal{W}
\end{tikzcd}$$

The benefit of this account is that it clearly separates the theoretical structures which represent physical relations from the numerical structures which have no representational significance. The theoretical structures aim to offer a minimal picture of the world without superfluous structure. The fact that there exist many representations of a theoretical structure onto the same numerical structure shows that the latter does contain superfluous elements.

Finally, and to connect the present discussion to the topic of symmetries, we can use the three-level view to draw a distinction between active and passive transformations.[12] Suppose that a theoretical model consists of a domain of objects, D, and a function $m : D \to \mathcal{V}$ which assigns a value from the value space $\mathcal{V}$ to each object. The three-level view posits a pair of distinct functions. On the one hand, we have a 'map' f_i between elements of $\mathcal{V}$ and the world—for example, *this* element of $\mathcal{V}$ represents *that* mass. On the other hand, we have a map r from the same value space into $\mathbb{R}$: *this* number represents *that* element of $\mathcal{V}$. I will define a *passive transformation* as a 'relabelling' of the middle-level theoretical structure. We thus keep fixed the interpretation f_i between $\mathcal{V}$ and $\mathcal{W}$ while varying the representation function f_r from $\mathcal{V}$ into $\mathbb{R}^+$:

Passive Transformation. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a bijection of $\mathbb{R}^+$; then a

[12]A similar account is found in Wolff (2020, §8.2.1). But my definitions depart from Wolff's, who defines active transformations as automorphisms of the relevant structures. Instead, I call such transformations kinematical symmetries; see the next section.

> passive transformation is a map $\psi^* : f_r \to \psi \circ f_r$.

Passive transformations act on the numerical representation r of quantities, but leave alone their physical interpretation f_i. For example, a transformation from grams to kilograms changes the numerical label associated with each mass value, but leaves the actual mass assigned to particles the same.

Active transformations, on the other hand, leave the units fixed but assigns objects a different actual mass value. Active transformations act on theoretical models, rather than on their numerical representation. Despite what the name may suggest, then, an active transformation does *not* necessarily enact a physical transformation: it is a transformation of the theory's models, but these models may stand in a many-to-one relation to the physical possibilities. With this caveat in mind, I define active transformations as follows (where m is a quantity with value space $\mathcal{V}$):

> *Active Transformation.* Let $\psi : \mathcal{V} \to \mathcal{V}$ be a bijection of $\mathcal{V}$; then an active transformation is a map $\psi^* : m \to \psi \circ m$.

Since both f_i and f_r are kept fixed, a particle is assigned both a different physical mass *and* a different numerical label after an active transformation. For example, if a particle was previously assigned a mass of 5 kilograms, then after an active transformation it is assigned a mass of (say) 10 kilograms. In particular, note that the active transformation on $\mathcal{V}$ 'drags along' the representation function f_r in this way. As we will see below, the symmetries we are interested in are active transformations of this sort.

5.3 Symmetries: Internal and External

We now return to Quine's distinction between ontology and ideology. In this section, I will argue that this distinction is superseded by a different one: the distinction between *internal* and *external* degrees of freedom (DOFs). This latter distinction differs from Quine's in that it is not absolute, but relative to the quantities a theory posits. This reconfiguration of our metaphysical categories on the basis of our best scientific theories is in the spirit of naturalised metaphysics.

5.3.1 The Internal/External Distinction

The internal/external distinction is most familiar from the literature on symmetries. I will define symmetries in more detail below, so for now it is only relevant to note that symmetries are often said to 'act' on some part of the theory's ontology (in our expanded sense). For instance, spacetime symmetries act on $\mathcal{M}$, the manifold of spacetime points. Similarly, mass symmetries act on the mass value space $\mathcal{V}$. We can connect this sense of 'acting on' to the notion of active transformations defined above. The idea is that spacetime symmetries are, or are induced by, active transformations of $\mathcal{M}$, and *mutatis mutandis* for mass symmetries. The distinction between internal and external symmetries is then drawn in terms of this action. Here the literature diverges. On the one hand, one often sees external symmetries defined as exactly those symmetries which act on spacetime—all other symmetries are internal. For instance, Kosso (2000, 84) writes that "[a]n *external* symmetry is one in which the transformation is a change of a spatial or temporal variable. An *internal* symmetry is one in which the transformation is not a change of spatial or temporal parameters". On the other hand, one also comes across definitions of external symmetries as those which act on a theory's *independent* DOFs, while internal symmetries act on the theory's *dependent* DOFs. For instance, Belot (2013, 330) defines internal symmetries as symmetries "that involve transformations of dependent variables of the theory that do not affect the independent variables". It seems that the first definition is slightly more common, especially in the philosophical literature. But for reasons that will become apparent, I prefer the second definition. For local spacetime theories (as defined in §5.1.3) the second definition is equivalent to the first. I will now explain the second definition in more detail.

Recall that a theory's quantities are functions between (sub-)domains. For any function ϕ, we call its domain the independent DOF and its codomain the dependent DOF. For instance: if $\rho(x)$ is a function from spacetime into mass density value space, then the spacetime coordinates x are the independent DOFs and the mass density values $\rho(x)$ are the dependent DOFs. This definition accords with common usage in physics. Traditionally, the independent variables are identified with the x-axis of a graph, and the dependent variables with the y-axis; for example, an x-vs-t graph plots location as a function of time. It follows that the independent/dependent distinction is relative to a choice of function. It is possible, in principle, for some value space to be both domain of one function and codomain of another. For example, in classical particle mechanics the spacetime manifold $\mathcal{M}$ is the codomain of the particle position function $x(i)$, but also the domain of the gravitational potential field $\varphi(x)$. But as we will see below, in the case of local spacetime theories we can usually speak of *the* (in)dependent variables of a theory.

It may seem as if the distinction between independent and dependent variables is arbitrary. After all, couldn't we introduce an inverse function ρ^{-1} from mass density values into spacetime? In that case, spacetime points are the dependent DOFs. The answer is 'No': there is an asymmetry between independent and dependent variables. The reason is that physical quantities are typically not bijections. For instance, it is possible for distinct particles to have the same mass or colour, or for the gravitational potential to have the same value at distinct spacetime points. But it is not possible for the same particle to have two masses at once, or for the gravitational field to have distinct values at the same spacetime point. Therefore, the distinction between independent and dependent DOFs is justified by the formal features of the function between them. Carnap (1928, §158) already drew the same conclusion, calling the different classes of DOFs 'objects of the first type' and 'objects of the second type' respectively. However, there is no *a priori* reason to expect that physical quantities must be represented by non-injective functions. We can easily imagine a theory in which some quantities *are* represented by bijections. The distinction between independent and dependent DOFs then vanishes. We can think of this puzzle—why *are* all nature's quantities represented by non-injective functions?—as the converse of the famous *exclusion problem.* The latter problem refers to the fact that any object instantiates a unique value of a determinable quantity. For instance, as we saw above each particle has one unique mass. Put differently, the exclusion problem raises the question: why are all quantities represented by functions? The related puzzle we have encountered here instead asks: why are all quantities represented by *non-injective* functions? It would be far beyond the scope of this chapter to solve either of these puzzles, let alone both. Suffice it to say that the independent/dependent distinction depends on the fact that nature seems to prefer non-injective functions, for some reason or another.

We can now define external symmetries as those that act on a theory's independent DOFs, and internal symmetries as those which act on a theory's dependent DOFs. I will offer a more precise definition of both types of symmetries below. First, I will return to the claim that in the case of local spacetime theories, a theory's external symmetries just are its spacetime symmetries. Recall that a local spacetime theory consists of a manifold $\mathcal{M}$ over which certain geometric objects are defined. These objects fall into two classes: spacetime structure, and dynamical objects. The latter are the theory's quantities, namely functions from $\mathcal{M}$ into some value space. The fact that all dynamical objects are (non-invertible!) functions *from* $\mathcal{M}$ into some value space means that for each quantity, the independent DOFs just are the spacetime variables. Therefore, for local spacetime symmetries the second definition of external symmetries as those which act on the theory's independent DOFs collapses into the first definition, on which external symmetries are spacetime symmetries.[13]

5.3.2 Internal and External Symmetries

I will now present a more precise definition of symmetries. Dasgupta (2016) has given a useful tripartite classification of symmetry definitions. The first type, *formal* definitions, define symmetries in purely formal terms, absent any physical interpretation of the formalism. The second, *ontic* type define symmetries as those transformations that preserve some *physical* features, such as a distinguished set of quantities; this requires a physical interpretation of the theory's formalism. The third type of definition is *epistemic*: symmetries are defined as transformations that preserve empirical content (of course, this is a necessary but not a sufficient condition).

The definition of symmetries I will offer is a formal one. Formal definitions have recently come under attack from Belot (2013) and Dasgupta (2016), who both argue that neither formal nor ontic accounts can ensure that symmetry-related states are empirically equivalent. Since symmetry transformations are philosophically relevant because they seem to relate empirically equivalent states of affairs, this is a problem for formal definitions of symmetries. It is not my aim here to defend the claim that symmetries under some formal definition invariably relate empirically equivalent states of affairs.[14] Rather, my more limited aim is to offer a definition that encompasses a broad class of symmetries of philosophical interest—including both external symmetries and internal symmetries—but excludes less interesting transformations, such as arbitrary permutations of possibilities or symmetries of particular subsystems. This will suffice to establish the required connection between symmetries and

[13]Indeed, some recent developments in the philosophy of physics further collapse this distinction. In particular, both Wallace (2015) and Dewar (2020a) present mathematical frameworks in which spacetime symmetries act on the theory's *dependent* DOFs, and hence (on our definition) count as internal symmetries. These systems provide further support for our defence of the second definition, since on the first definition the claim that spacetime symmetries are internal is incoherent. The versatility of the second definition is thus more true to the range of metaphysical possibilities even for our current best physical theories.

[14]For a to my mind convincing response to Dasgupta's and Belot's criticisms, see Wallace (2019b).

structure. With this aim in mind, I can find no fault with a formal definition.

I define symmetries for local spacetime theories as follows. Recall that a local spacetime theory consists of a spacetime $\mathcal{M}$ and a set of functions ϕ_i into associated value spaces $\mathcal{V}_i$. Denote a model of such a theory $\langle \mathcal{M}, \mathcal{V}, \phi \rangle$. We then define external symmetries as follows:[15]

> *External Symmetry*: Let ψ be a diffeomorphism which acts on M; then ψ is an external symmetry iff the map $g_\psi : \langle M, \mathcal{V}, \phi \rangle \to \langle M, \mathcal{V}, \psi_* \phi \rangle$ maps the space of DPMs onto itself (where ψ_* is the pushforward of ψ).

This definition of external symmetries is entirely standard; cf. Earman (1989, 45). In effect, it says that if one has a permutation ψ of points of M, then this permutation induces a transformation of the fields defined over M. In particular, suppose that ϕ takes values in a value-space whose values are common to all points in the base space, and that ψ maps a point p to a distinct point q; then if the value of ϕ at q before the transformation was $\phi(q)$, its value after the transformation at the same point q is $\phi(p)$—that is, the value that ϕ previously had at p. If this transformation preserves the theory's dynamics (that is, if it maps solutions to solutions), then it counts as an external symmetry. In this way, the above definition encompasses all of the usual spacetime symmetries. For example, a static shift is induced by a diffeomorphism which maps each point to another point some fixed distance away from the first in some particular direction. On the other hand, the definition excludes arbitrary permutations of a theory's DPMs, since not all maps on the space of KPMs are expressible as the 'lift' of a diffeomorphism.[16]

We can define internal symmetries analogously. I am not aware of any previous definitions of internal symmetries along these lines, but the parallels with the

[15]But note that this definition fails in the context of theories set on fibre bundles, since for those theories the action of the push-forward map on a section of the bundle is ill-defined. As Dewar (2020a) argues, in that context it is easier to interpret the external symmetries as a subset of the internal symmetries.

[16]Cf. Dasgupta's (2016) notion of a 'generated' versus a 'bare' transformation.

definition of external symmetries above are clear:[17]

> *Internal Symmetry*: Let ψ be a bijection of $\mathcal{V}$; then ψ is an internal symmetry iff the map $g_\psi : \langle M, \mathcal{V}, \phi\rangle \to \langle M, \mathcal{V}, \psi \circ \phi\rangle$ maps the space of DPMs onto itself.

An internal symmetry is induced by a permutation of the values of some quantity. For example, a uniform mass scaling sends every mass value to another one which is a multiple of the first. Such a permutation induces a transformation of the fields defined over M, as follows: if before the transformation the value of ϕ at p is $\phi(p)$, then after the transformation the value of ϕ at p is $(\psi \circ \phi)(p)$. Again, such transformations count as symmetries iff they preserve the dynamics. As I will show in Chapter 8, this definition accounts for the gauge symmetries of classical field theories when formulated in terms of fibre bundles. On the other hand, as before this definition excludes arbitrary permutations of the theory's models, since not all maps on the space of KPMs are expressible as the 'lift' of an active transformation of one of the value spaces.

The similarities between our definitions of external and internal symmetries further support the thesis that the traditional view of entities as part of a theory's ontology and properties and relations as part of a theory's ideology requires revision. Rather, it seems that the internal and external variables of a theory are to some extent on a par: *both* range over the theory's ontology. The theory's ideology consists of functions between the theory's sub-domains on the one hand, and the structure of these sub-domains on the other. I will say more about the latter in the next section.

5.4 Symmetry Principles

Quine famously believed that we could read off a theory's ontological commitments from its formalism: the theory is committed to whatever one needs to quantify over

[17] I believe that my definition here is a special case of Dewar's (2020a).

in order for the theory's claims to come out as true. This is still the case for modern physical theories, since we can read off the theory's ontology from the quantities it posits. Since quantities themselves are part of the theory's ideology, we can also read off aspects of the latter from the formalism. But can we also discover a theory's *structure*—the second half of its ideology—in the same way? I believe the answer is 'No'. For history is familiar with instances of empirically successful theories which exhibit superfluous structure. For instance, Newton posited absolute space, but absolute positions and velocities play no role in the novel predictions or scientific explanations of classical mechanics. Therefore, a 'literal' interpretation of a theory does not always reveal the most appropriate physical structures. Fortunately, history is also familiar with an indirect method for the discovery of structure: consideration of symmetries.

Earman (1989) articulates a pair of symmetry principles which link a theory's (external) dynamical symmetries to its spacetime symmetries. Earman's spacetime symmetries are quite unlike the internal and external symmetries defined above—which we will collectively call *dynamical symmetries*—as they are independent from the theory's dynamics. Rather, a theory's spacetime symmetries are diffeomorphisms of $\mathcal{M}$ which preserve spacetime structure, for instance the metrical distances between points. In technical terms, spacetime symmetries are *automorphisms* of $\mathcal{M}$, where $\mathcal{M}$ includes all (absolute) spacetime structure. The reason that spacetime symmetries can reveal the structure of spacetime is that there exists a duality between structure and symmetries. On the one hand, the more structure a spacetime has, the fewer symmetries. For each additional piece of structure, a symmetry transformation will have to preserve that structure—so the more properties and relations we build into a spacetime, the more difficult it will become to preserve spacetime structure. On the other hand, the more symmetries a spacetime has, the less structure. In fact, on some views of structure we can *identify* a spacetime structure with its symmetries. The idea is that even if some property or relation is not explicitly included in the definition of a particular spacetime structure, it is implicitly definable from that structure iff it is invariant under the symmetries of the spacetime in question (Barrett, 2017, cf. chapter 4). Therefore, *to know a space's symmetries is to know its structure.*

According to Earman, we can determine the symmetries of spacetime from a theory's external symmetries.[18] In order to do so, Earman postulates the following pair of principles:

$\mathrm{SP1}_{\mathrm{ext}}$ Any external symmetry of T is a spacetime symmetry of T;

$\mathrm{SP2}_{\mathrm{ext}}$ Any spacetime symmetry of T is an external symmetry of T.

These principles are meant to ensure that a theory's spacetime has neither too much nor too little structure. If $\mathrm{SP1}_{\mathrm{ext}}$ fails, then the theory contains spatiotemporal structure which is dynamically irrelevant. For example, Newtonian spacetime contains a standard of absolute rest, which means that boosts are not symmetries of this spacetime. But boosts *are* dynamical symmetries of Newtonian Gravitation, which means that the standard of absolute rest is irrelevant to the theory's dynamics. Conversely, if $\mathrm{SP}_{\mathrm{ext}}$ fails then the theory has too little structure to sustain the dynamics. As a consequence the theory has to draw distinctions between regions of spacetime which are qualitatively identical, but which behave differently dynamically. Together $\mathrm{SP1}_{\mathrm{ext}}$ and $\mathrm{SP2}_{\mathrm{ext}}$ are useful tools in order to determine the appropriate amount of spacetime structure for local spacetime theories.

But Earman does not consider internal symmetries, nor the structure of internal value spaces (such as mass value space). I propose to extend Earman's principles to both types of symmetries.[19] Firstly, let the term *dynamical symmetries* cover both internal and external symmetries as defined in §5.3. Importantly, each dynamical symmetry is a bijection ψ of some sub-domain of the theory, whether that is spacetime

[18] Earman calls these 'dynamical symmetries', but we will reserve the term to include both external *and* internal symmetries.

[19] As far as I am aware, the first one to suggest an extension of Earman's principles to internal symmetries was Hetzroni (2019), but the idea hasn't surfaced in the literature since then.

or an internal value space. Secondly, let the term *kinematical symmetries* denote any automorphism of one of the theory's sub-domains. We can see that spacetime symmetries are a species of kinematical symmetries, but so are automorphisms of any internal value space. For instance, it is often thought that the automorphisms of mass value space are uniform scaling transformations (I will address this question fully in Chapter 6). In analogy with $\mathrm{SP1}_{\mathrm{ext}}$ and $\mathrm{SP2}_{\mathrm{ext}}$, then, I propose the following symmetry principles must hold for any dynamical or kinematical symmetry, whether internal or external:

SP1 Any dynamical symmetry of T is a kinematical symmetry of T;

SP2 Any kinematical symmetry of T is a dynamical symmetry of T.

The motivation behind these extended principles is analogous to Earman's original motivation for $\mathrm{SP1}_{\mathrm{ext}}$ and $\mathrm{SP2}_{\mathrm{ext}}$. For example, one way in which SP1 can fail for internal symmetries is when we identify the structure of the value space of some dimensionful quantity with that of the real numbers. As we have seen in §5.2, this in effect endows that quantity with a preferred system of units. But the dynamics are not sensitive to this structure, since the laws *also* hold in different units. Conversely, SP2 ensures that the theory has *enough* internal structure to sustain the dynamics. If SP2 fails, then values which are qualitatively identical play distinct dynamical roles. This violates the plausible assumption that the dynamics are only sensitive to qualitative structure. Together SP1 and SP2 act as guides towards the appropriate structure of our theory's value spaces, whether internal or external.

5.5 Conclusion

I have argued that a theory's quantities possess a certain internal structure, and that symmetries can tell us what this structure is. In conclusion, let me sketch how this view compares with the orthodoxy in philosophy of physics.

The orthodoxy is expressed in Ismael and van Fraassen's (2003) slogan that *symmetries are a guide towards superfluous theoretical structure*. Now, there is a sense in which this is correct: as we have seen, when a theory's kinematical symmetries *exceed* their dynamical symmetries, this is a sign that the theory contains superfluous kinematical structure. However, the converse is also true: if a theory's dynamical symmetries exceed its kinematical symmetries, then that theory does not posit *enough* kinematical structure. The orthodoxy only pays attention to one side of the same coin. Furthermore, the orthodox view often overshoots its goal by claiming that the variant quantities *themselves* are in some sense redundant. This leads to reductionism, the demand for theories formulated solely in terms of invariant quantities. But as our symmetry principles show, reduction is not required for their satisfaction. Rather, the lesson for quantities is similar to that for spacetime: the symmetries of our theory tell us what their structure is, not whether they exist. In the case of static shifts, for example, the lesson is *not* that particles have no position in space at all. After all, our theories are still expressed in terms of particles with trajectories on a manifold (or, in the case of local spacetime theories, fields defined over some manifold). Instead, the lesson of the shift-invariance of classical mechanics is one about spacetime *structure*, namely that there exists no privileged coordinate system. Put differently: whereas Ismael and van Fraassen believe that symmetries hint at superfluous *ontology*, on our view symmetries indicate superfluous *ideology*.

These remarks lead us to a different slogan: *symmetries are a guide towards the structure of physical quantities*. Rather than an essentially negative role in telling us which parts of our theories don't refer, on this view symmetries play a positive role in informing us about the structure of the world. I believe that this alternative conception of symmetries has fruitful applications, as I will now show in two detailed case studies in the remaining chapters.

Chapter 6

Case Study I: Mass Scalings

We have seen that a theory meets the formal requirements for sophistication when every kinematical symmetry is a dynamical symmetry and *vice versa*. Then—and only then—are we allowed to interpret the theory's symmetry-related models as physically equivalent. Some theories are 'born' sophisticated: their symmetries are already isomorphisms. This is the case for General Relativity, since its symmetries are diffeomorphisms of the spacetime manifold. But not all theories are so lucky. In particular, *internal* symmetries are usually not isomorphisms of the theory in question. Examples of internal symmetries are mass scalings, shifts of the electrostatic potential, and so-called gauge transformations. As we saw in the previous chapter, quantities such as mass or the electrostatic potential are usually held to take values in the real numbers. But the real numbers have no (non-trivial) automorphisms (a claim I will justify below), so on this flat-footed approach internal symmetries are never isomorphisms of the theory's models. The standard response to this puzzle is reduction: just remove the variant quantities from the theory. In Chapter 3, I presented the case against reduction in the form of the No Miracles Argument for comparativism. Here we will choose the alternative approach: sophistication.

In order to achieve sophistication, we must restructure the theory's internal value spaces such that the kinematical symmetries match the theory's dynamical symmetries. When this condition is met, we can interpret symmetry-related models anti-quidditistically as physically equivalent. In this chapter and the next, I will carry out this process of sophistication in two detailed case studies. In this chapter, I turn my attention to the internal symmetries of Newtonian Gravitation (NG). In particular, we will consider *mass scalings* as symmetries of NG. In the next chapter, I extend this approach to *local* (or 'gauge') symmetries.

The outline for the present chapter is as follows. In §6.1, I present the debate between absolutists and comparativists. The debate turns on whether mass scalings are symmetries of NG. In §6.2, I discuss the (kinematical) structure of mass value space. Following the previous chapter, I argue that this is not the structure of the real numbers, but a so-called *additive structure*. In §6.3, I consider mass scalings in more detail. We will see that in the first instance, mass scalings are not symmetries of NG. However, mass scalings which *include* the gravitational constant G *are* symmetries of NG, as I will argue in §6.4. This fact puts pressure both on absolutism and comparativism. In §6.5, I present an alternative: sophistication about mass. §6.6 discusses the implications of sophistication for Martens' (2019b) thesis of 'kinematic comparativism. §6.7 concludes.

6.1 Absolutism Versus Comparativism

The question whether mass scalings are true symmetries of Newtonian Gravitation is controversial. The result of a mass scaling is to 'multiply' all mass quantities by a uniform positive constant (I will define this transformation in more detail below). For instance, if before a scaling the mass of a litre of water is 1 kg, then after the scaling the mass of a litre of water is 2 kg. The contested claim that uniform mass scalings are symmetries of NG has generated significant debate over the status of absolute masses.

In this debate, the *comparativist* claims that mass scalings *are* symmetries of NG, and hence that intrinsic masses are not the theory's fundamental quantities.[1] Dasgupta (2013) argues that the existence of mass symmetries implies that particle masses are undetectable, since there are no observable differences between symmetry-related worlds. We can illustrate this claim with the example of weights on a balance scale. If the weight on the left of the scale is more massive than the one on the right, then this is still the case when both masses are doubled; and if a weight on the left is equal in mass to a pair of weights on the right, then this also is still the case when their respective masses are doubled. It seems to follow that we cannot measure the absolute mass of any object in this way. Therefore, the comparativist proposes that *mass ratios* are nature's fundamental mass quantities. Since mass ratios are invariant under mass scalings, their values *are* detectable. We can recognise mass comparativism as an instance of reduction: a reformulation of the theory in terms of quantities that are invariant under its symmetries.

On the other side of the debate, the *absolutist* responds that mass scalings are *not* symmetries of NG, because absolute masses are (more) fundamental (than mass ratios).[2] By 'absolute masses' I mean *monadic* mass properties of individual particles. Martens argues that the laws of NG dynamically distinguish mass magnitudes from each other, such that a uniform mass scaling would violate these laws. For example, the escape velocity of a projectile is proportional to $\sqrt{m}$, so a mass scaling can mean the difference between escaping or not. This means that there are observable differences between mass-scaled worlds, *pace* Dasgupta. If the dynamics of Newtonian Gravitation draw distinctions more finely than comparativism, mass ratios are an

[1]See, for example, Dasgupta (2013); Roberts (2016). Advocates of comparativism also include Russell (1903); Field (1980); Ellis (1966); Bigelow and Pargetter (1988).

[2]See, for example, Baker (2014); Martens (2019b). Other advocated of absolutism include Mundy (1987); Armstrong (1988); Lewis (1986). For a comprehensive overview of absolutist and comparativist positions, see Martens (2019b).

incomplete basis of fundamental quantities.

The point I wish to emphasise is that the absolutism-comparativism debate assumes that the status of absolute masses as fundamental quantities stands or falls with their invariance under the internal dynamical symmetries of NG. If masses vary under those symmetries, then comparativism is preferable, while absolutism carries the day if the theory possesses no non-trivial mass symmetries. This is a false dilemma: it is possible for mass quantities to vary under the theory's symmetries *and* represent elements of fundamental reality. For as I pointed out in Chapter 3, absolutism is neutral between literalism—the view that symmetry-related models represent distinct states of affairs—and sophistication, which says that symmetry-related models represent the *same* state of affairs. The latter view does not entail underdetermination, so the argument for comparativism loses its force. This dialectic is familiar from the debate around the Hole Argument: there it was initially claimed that the symmetries of General Relativity motivated a form of relationism, but later realised that a sophisticated form of substantivalism can avoid underdetermination without the loss of spacetime points. I will argue that in the case of mass, too, sophistication offers a 'third way' between absolutism and comparativism: one that agrees with the comparativist that NG possesses non-trivial mass symmetries, and with the absolutist that monadic mass quantities are fundamental.

I explained in the introduction to this chapter that sophistication requires that the structure of a quantity's value space matches the theory's internal symmetries. In order to apply sophistication to the case of mass, then, an investigation into the structure of mass value space is required. I will turn to this task now.

6.2 The Structure of Newtonian Mass

Let's first consider the form of the models of Newtonian Gravitation. In the following, I will focus on the classical dynamics of massive particles under the force of gravitation. I will assume that the models of NG have the following form:

$$\text{NG: } \langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathbf{g}, x(i,t), m(i), \mathcal{V}_m \rangle$$

where $\mathbb{A}$ is a three-dimensional Euclidean affine space; $\mathbb{T}$ is a one-dimensional temporal manifold; $\mathcal{B}$ is a set of massive particles; $\mathbf{g}$ is a spacelike vector field on $\mathbb{A} \times \mathbb{T}$ which represents the gravitational field; $x(i,t) : \mathcal{B} \times \mathbb{T} \to \mathbb{A}$ is a smooth function which represents the trajectories of massive bodies; and $m(i) : \mathcal{B} \to \mathcal{V}_m$ is a function from the set of particles into a 'mass value space' $\mathcal{V}_m$, which represents the mass of each particle. Let me emphasise that the mass value space $\mathcal{V}_m$ is meant to represent a *physical* structure. The elements of the underlying domain of $\mathcal{V}_m$ represent determinate mass properties, and the structure of $\mathcal{V}_m$ represents real second-order relations between them. For example, the mass property 'being 10 kg' stands in the second-order relation of 'being twice as much' to the mass property 'being 5 kg'. I defended the need for such structure in the previous chapter. The *appropriate* structure for $\mathcal{V}_m$ is that which accurately represents these physical relations. The question of this section is what that structure is, given the symmetries of NG.

This account of NG departs from standard modern formulations of the theory on a few points. Firstly, it is common to set NG on a Newtonian spacetime $\langle M, t_{ab}, h^{ab}, \nabla, \sigma^a \rangle$, where M is a differentiable manifold. However, the idea that NG is more appropriately set on a Euclidean affine space is familiar from Stachel (1993); see also Saunders (2013) and Dewar (2015). The particular reason I adopt this view here is that it allows for an easier definition of the gravitational constant G in §6.4. Secondly, standard formulations employ the gravitational potential ϕ, rather than the gravitational field $\mathbf{g} \coloneqq -\nabla\phi$. But since any transformation of ϕ that leaves $-\nabla\phi$ the same is a symmetry of NG, it is clearly the gravitational field $\mathbf{g}$ which represents physical structure (Knox, 2014).[3] Finally, I have presented a particle rather than a field formulation of NG. This choice is merely a matter of convenience: the arguments below remain valid if one replaces the function m with a scalar field ρ from spacetime into some mass *density* value space $\mathcal{V}_\rho$. However, note that ρ has different units from m, so $\mathcal{V}_\rho$ represents a different physical structure from $\mathcal{V}_m$. In this chapter, I will focus on the latter value space.

With the structure of NG in place, I will now discuss the particular structure of $\mathcal{V}_m$. In the first instance, I will consider *real number structures* (§6.2.1). But we will see that these structures are too rich for the representation of dimensionful quantities. This echoes some of the conclusions from the previous chapter. This leads to a second proposal, the so-called *additive extensive structures* (§6.2.2), which provides a more realistic representation of the relations between determinate masses.

6.2.1 Real Number Structures

The mass quantity $m(i)$ is a function from particle trajectories—but into what codomain? It is often assumed that the mass value space $\mathcal{V}_m$ is isomorphic to the positive real numbers. In foundational treatments of NG considered as a theory of fields, the mass density field ρ is similarly defined as a scalar field, i.e. a function from spacetime points into the (positive) real numbers (Friedman, 1983; Malament, 2012). But as I explain below, this 'received view' effectively assumes a preferred unit system. This implies that it endows $\mathcal{V}_m$ with *too much* structure.

[3]However, one might worry that one ought to be realist about the gravitational potential for the same reasons that I argued, following Dewar (2019), one ought to be realist about the electrostatic potential (cf. §2.4). I do not believe that the main argument in this chapter rests on this issue, so here I will adopt realism about the gravitational field for convenience.

In more detail,[4] the positive real numbers $\mathbb{R}^+$ form a structure $\langle R, \leq, +, \times \rangle$, where:

- R is a base set of cardinality $2^{\aleph_0}$;
- $\leq$ defines a total order on R;[5]
- $\langle R, < \rangle$ is Dedekind-complete;[6]
- $\langle \mathrm{R}, < \rangle$ is Archimedian;[7]
- $+$ and $\cdot$ form a commutative semi-field;[8]
- The order $\leq$ is compatible with addition;[9]
- The members of R are all positive.[10]

This structure is called a *complete ordered positive semifield*. The claim that this is just the structure of the positive real numbers is justified by the fact that for any complete ordered positive semifield, there exists a unique isomorphism to $\mathbb{R}^+$ (Spivak, 1994, Ch. 30).

But the problem with the suggestion that the above structure represents the structure of physical mass is that it, in effect, endows mass with a preferred unit system. For suppose that, for some body i and some positive real number x, $m(i) = x$. Since $x \in \mathbb{R}^+$, the mass of i is associated with a particular real number. Furthermore, since $\mathcal{V}_m$ is meant to represent *physical* structure, this association does not depend

[4]These details are expounded in even more detail in Spivak (1994). See Dewar (2020b, §A) for a lucid exposition of these results.

[5]An order is total iff it is transitive, anti-symmetric and connex, i.e. (i) $a \leq b$ and $b \leq c$ implies $a \leq c$, (ii) if $a \leq b$ and $b \leq a$ then $a = b$, and (iii) either $a \leq b$ or $b \leq a$.

[6]A set is Dedekind-complete iff every non-empty subset which has an upper bound has a least upper bound, where an upper bound of a subset A of R is an element $x \in R$ such that $a \leq x$ for all $a \in A$, and a least upper bound is an upper bound x such that $x \leq y$ for all upper bounds y of A.

[7]For any $x \leq y$, there is a unique z such that $x + z = y$.

[8]A commutative semifield consists of a set R with two binary operations $\cdot$ and $+$, such that $\cdot$ is a commutative group and $+$ is a commutative semigroup (i.e. a group with no inverse or identity), and $a \cdot (b + c) = a \cdot b + a \cdot c$.

[9]An order $\leq$ is compatible with addition iff if $x \leq y$, then $x + z \leq y + z$ for any $z \in R$.

[10]The members of R are positive iff $x \leq x + y$ for all $x, y \in R$.

on any particular choice of units. But this association between numbers and mass magnitudes is spurious. As Martens (2019b) notes: there is nothing 'five-ish' about an object that is 5 kg in mass. So, mass value space cannot have the structure of the real numbers. For a helpful analogy, consider the spacetime manifold M. Suppose that we assert that the manifold has the structure of $\mathbb{R}^4$ considered as a complete ordered semifield. Since $\mathbb{R}^4$ has a unique zero element, this means that the manifold M also has a preferred point zero. But this endows M with more structure than is necessary, since in fact *any* point in M could play the role of the origin of a coordinate system. Likewise, then, we can assign the number '5' to any point of mass value space via a particular choice of units. Therefore, to identify any particular mass with that number confuses fact with convention.

6.2.2 Additive Extensive Structures

Martens (2019b) similarly argues that the received view amounts to a "fail[ure] to distinguish between *physical magnitudes* and the *numerical quantities* used to represent them". Instead, he endorses a type of structure that is studied under the guise of 'measurement theory'.[11] These structures are known as *additive extensive structures*, and characterised in Hölder (1901) as models of the form $\langle D_m, \leqslant, \circ \rangle$,[12] where

- D_m is a base set of cardinality $2^{\aleph_0}$;
- $\leqslant$ defines a total order on D_m;[13]
- $\langle D_m, \leqslant \rangle$ is Dedekind-complete;
- $\langle D_m, \leqslant \rangle$ is Archimedian;

[11] The *locus classicus* is Krantz et al. (1971); see Wolff (2020, Ch. 5) for a recent treatment.

[12] There is a subtlety here, since the structures defined by Hölder are different from those discussed in Krantz et al. Specifically, the former map *onto* $\mathbb{R}^+$, whereas the latter map *into* $\mathbb{R}^+$, as is pointed out in Krantz et al. (1971, §2.2.6).

[13] In fact, the assumption of transitivity is unnecessary, as it follows from the axioms below (Hölder, 1901, §2.1).

- $\langle D_m, \circ\rangle$ forms a semigroup;[14]

- For any $x \in D_m$, there is a $y \in D_m$ such that $x \leqslant y$;

- The members of D_m are all positive.

In physical terms, $\leqslant$ is interpreted as the binary relation of one mass being less than or equal to than another, and $\circ$ as the tertiary relation of two masses equalling a third. As Martens (2019b) notes, the intuitive thought behind the fact that $\circ$ admits of no inverses or identity is that there are no zero or negative masses.

Hölder (1901) has proven that one can represent $\langle D_m, \leqslant, \circ\rangle$ on $\mathbb{R}^+$, in the following sense: there exists a function $f_r : D_m \to \mathbb{R}^+$ such that (i) $x \leqslant y$ iff $f(x) \leq f(y)$ and (ii) $x \circ y = z$ iff $f(x) + f(y) = f(z)$. Furthermore, this representation is unique up to multiplication by a positive constant α, so that if f_r and f'_r both represent $\langle D_m, \leqslant, \circ\rangle$, then there is some $\alpha > 0$ such that $f'_r = \alpha f_r$. Therefore, the representation of an additive extensive structure onto $\mathbb{R}^+$ is not *unique*, and hence there is a precise sense in which the former has less structure than the latter.[15] Unlike the real number structure, an additive extensive structure does not define a privileged unit system. Instead, the fact that f_r is defined up to multiplication of a positive constant means that any unit system related by such a transformation represents the structure of $\langle D_m, \leqslant, \circ\rangle$ equally well. Since different units for mass *are* related by such transformations, $\langle D_m, \leqslant, \circ\rangle$ is a plausible candidate for the correct structure of $\mathcal{V}_m$.

Martens (2019b) points out that under this structure, the elements of D_m are *absolutely indiscernible*: for no $m \in D_m$ does there exists a formula with one free variable that is true of m only.[16] In other words, all mass values 'look the same'

[14] A semigroup consists of a set D_m and a binary operator $\circ$ such that $\circ$ is associative, i.e. $(a \circ b) \circ c = a \circ (b \circ c)$, but does not admit of inverses or an identity.

[15] In Dewar's (2020b) terms, $\langle D_m, \leqslant, \circ\rangle$ is a *principal homogeneous space* for $\mathbb{R}^+$ which "forgets the origin" of multiplication.

[16] Martens writes that masses are 'qualitatively identical', but this is false in the sense that elements of m are at least *relatively discernible*, i.e. there are formulas with *more than one* free variable that apply to sets of elements in only one particular order (cf. Saunders (2003b)). For example, if $m_1 < m_2$, then m_1 is relatively discernible from m_2 as the smallest of the two.

intrinsically. For example, for any mass m there are infinitely many smaller masses, and infinitely many larger ones. However, Martens argues, the laws of nature treat different masses differently—the more massive an object is, the slower it accelerates in response to forces.[17] Therefore, Martens continues, we need to introduce intrinsic identities, or *quiddities*: non-qualitative properties which are uniquely possessed by individual mass magnitudes.[18] The quiddities tell forces how to 'latch onto' masses: they are "required for the forces to be well-defined, in the sense of uniquely matching up instances of initial conditions, including masses, with, say, accelerations" (Martens, 2019b, 2519). Since quiddities allow for trans-world identification, their existence implies that mass-scaled worlds are physically distinct. In §6.5, I will argue that quiddities are superfluous for exactly this reason. But first I will discuss the nature of mass scalings in more detail.

6.3 Active Leibniz Scalings

We can now define what Martens calls *Active Leibniz Scalings* (ALS). Unlike a passive transformation, which only alters which numbers we use to represent particular masses, an active transformation transforms the actual mass of particles themselves. The effect of an active mass *scaling* is that each particle's mass (as represented on $\mathbb{R}^+$) is multiplied by a constant α.

Suppose that we have a representation f_r of mass values on $\mathbb{R}^+$ (where f_r is the 'internal space' analogue of a coordinate chart).[19] We can carry out a passive

[17] This is at least true for 'standard' formulations of Newtonian Gravitation. Martens (2019a) considers alternative formulations for which this is not the case, but I will not discuss these here.

[18] It is common to refer to the intrinsic identities of properties as *quiddities*, but the term is used both for *determinables*, such as 'mass' and 'charge', and *determinates*, such as the property of being five kilograms. I will use the term in the second sense; cf. Martens (2019b, fn. 6).

[19] One may wish for a more 'intrinsic' characterisation of Active Leibniz Scalings, i.e. one that does not proceed via a particular choice of units. This is indeed possible, but somewhat involved. As Hölder (1901) shows, one can associate a unique positive real number to every *pair* of elements of $\mathcal{V}_m$, which denotes their ratio (this is possible because ratios are scale-invariant). In order to multiply an element m_1 by a constant α, then, one maps it to the unique element m_2 such that their

transformation on this representation from $f_r(m)$ to $f'_r(m) = \alpha f_r(m)$. In order to turn this passive transformation into an *active* one, we consider the 'pullback' of $f'_r(m)$. In other words, we keep fixed the representation relation f_r and consider an automorphism ϕ on M induced by α, such that $\phi(m) = (f_r^{-1} \circ \alpha f_r)(m)$. Intuitively, if a mass m is associated in some system of units with some number x, then ϕ maps m onto the unique mass that is associated with the number αx, *keeping the units fixed.*

The automorphism ϕ in turn acts on $m(x)$—the function from particles into $\mathcal{V}_m$—as $\phi \circ m(i) \equiv (\phi \circ m)_i$. An Active Leibniz Scaling then induces the following transformation on the models of NG:

$$\langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathcal{V}_m, \mathbf{g}, x(i,t), m(i)\rangle \overset{\text{ALS}}{\rightarrow} \langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathcal{V}_m, \mathbf{g}, x(i,t), (\phi \circ m)(i)\rangle \tag{6.1}$$

From our definition of symmetries in the previous chapter, it follows that such transformations are symmetries of NG iff a model before an ALS is a solution to NG's equations of motion whenever the same model after an ALS is also a solution.

Now, *are* ALS dynamical symmetries of Newtonian Gravitation? The initial answer is 'No'. Martens shows this with his 'comparativist's bucket', an analogue of Newton's famous bucket experiment.[20] As Martens notes, the escape velocity of a particle is given by:

$$v_{esc} = \sqrt{\frac{2Gm_2}{r}} \tag{6.2}$$

where G is the gravitational constant, m_2 is the mass of the body (say, the Earth) the projectile is escaping from, and r is the distance between the two. Now, suppose that in one model, the actual velocity of the projectile v is just over v_{esc}—and so the projectile escapes. Under an ALS, however, the planet multiplies in mass by a factor α, and hence v_{esc} is multiplied by a factor $\sqrt{\alpha}$. As a result, v may no longer

ratio $m_2 : m_1$ is equal to α.

[20]See Baker (2014) for the original version of this thought experiment.

exceed v_{esc}, so the projectile does not escape. Since the actual worldlines $x_i(t)$ of the earth and the projectile are left the same after an ALS, the transformed model is inconsistent with the equations of motion. Therefore, ALS are not symmetries of NG. On the other hand, Active Leibniz Scalings *are* automorphisms of $\mathcal{V}_m$, since $\leqslant$ and $\circ$ are invariant under ϕ. For example, consider arbitrary m_1, m_2 such that $m_1 \leqslant m_2$. By Hölder's representation theorem, $f_r(m_1) \leq f_r(m_2)$. Since α is strictly positive, $\alpha f_r(m_1) \leq \alpha f_r(m_2)$. By another application of the representation theorem, $(f_r^{-1} \circ \alpha f_r)(m_1) \leqslant (f_r^{-1} \circ \alpha f_r)(m_2)$. Since $\phi(m) = (f_r^{-1} \circ \alpha f_r)(m)$, this shows that $m_1 \leqslant m_2$ iff $\phi(m_1) \leqslant \phi(m_2)$, hence $\leqslant$ is invariant under the action of ϕ. The proof for $\circ$ is analogous ('left as an exercise').

This leads to a puzzle. For the fact that ALS are automorphisms of mass value space but do not induce dynamical symmetries suggests a mismatch between the dynamical and kinematical structure of NG. Rather than (the generalisation of) Earman's SP1, which states that all internal symmetries are value space symmetries, we have here a violation of SP2, which conversely states that all value space symmetries are internal symmetries. Earman's argument in favour of SP2 (applied to masses) is that were the principle to fail, "the theory would have to contain names, regarded as rigid designators, of [elements] of [mass value space]" (47, 1989). As we have seen in the previous section, this is exactly what Martens believes: that the laws of 'latch onto' masses via their quiddities. This means that the laws must pick out masses independently from their qualitative features, so in this sense the theory 'names' the mass magnitudes. Of course, the theory does not literally contain a rigid designator for any particular mass. Rather, the gravitational constant G functions to distinguish absolutely indiscernible masses from each other, as we will see in more detail below. Earman, however, sees this as a defect: "But such a difference in lawlike behavior is reason to suppose that $[m_1]$ and $[m_2]$ differ in some structural property that grounds the difference in behavior" (47, 1989). In other words, the fact that gravity acts in a certain way on a certain mass should follow from facts about what that mass is like—not just from the irreducible fact that it is *that* mass! I find these arguments equally compelling here as they are in the context of spacetimes (indeed, I am sympathetic to Myrvold's (2019) suggestion that such principles are analytic truths), and hence this suggests that there is something wrong with Martens' account of intrinsic mass.

6.4 The Gravitational Constant

But we have not yet considered the gravitational constant G. G is a dimensionful constant, proportional to $[M^{-1}]$. A common thought is that if we were to scale G *with* all particle masses, then ALS become symmetries after all. Now, as Martens points out, the fact that G has units of $[M^{-1}]$ by itself does not mean that G ought to scale with particle masses under an ALS. After all, ALS are active scalings of particle masses and not passive changes of unit, so there is no requirement to alter *all* quantities with dimensions of mass.[21] On the other hand, the fact that G is a dimensionful quantity means that it must represent some physical structure. Therefore, we can also consider what would happen *were* we to transform G, in addition to all particle masses. As we will see below, such a scaling *is* a symmetry of NG.

In more detail, I propose that we include G in the models of NG. The reason for including G in our models is simple: these models are *interpretations* (in the logical sense) of the equations of NG, so all terms which occur in these laws require an extension. Because G is usually thought of as a constant it is easy to ignore this, but in fact a model that does not include G is not really a model of NG at all. Moreover, G is a dimensionful constant, so it requires a *physical* interpretation.[22] Let's consider

[21] Both Roberts (2016) and Wolff (2020) aim to defend comparativism from the comparativist's bucket by including G in the Active Leibniz Scalings. For example, Wolff (2020, §8.2.3) argues that we fix the reference of G by ostension, and hence that in mass-scaled worlds G will have a *different* referent. But this argument misses the point, since the question is surely whether mass-scaled worlds have the same laws with respect to *our* language.

[22] Note also that (Luce et al., 2007, 308) prove that mathematically, dimensionful constants such

what sort of object G *is*. The role of G is to determine the contribution of a massive body to the gravitational field at a given distance from said body. This role is encoded in Newton's law of universal gravitation, which we can express as follows:

$$\mathbf{g}(\mathbf{x}, t) = -G \sum_i m(i) \frac{\mathbf{x} - \mathbf{x}_i(t)}{|\mathbf{x} - \mathbf{x}_i(t)|^3} \tag{6.3}$$

So, given a mass $m(i)$ and a displacement vector $\mathbf{r} \coloneqq \mathbf{x} - \mathbf{x}_i(t)$, G tells us the contribution of i's mass to the gravitational field at a distance $\mathbf{r}$ away from i.

Let's consider this in more detail.[23] First we will make explicit some of the theory's implicit structures. Recall that $\mathbb{A}$ is a three-dimensional affine space. As such, we can implicitly define from $\mathbb{A}$ a vector space $\mathbb{V}$ such that for any two points x, y in $\mathbb{A}$, there exists a unique vector $\vec{xy}$ in $\mathbb{V}$ which represents the displacement between these points. Moreover, we can posit a second vector space, $\mathbb{G}$, isomorphic to the first, whose elements represent values of the gravitational field (which by definition is vector-valued in $\mathbb{A}$). Unlike $\mathcal{V}_m$, these structures are not explicitly included in NG's models, because one can define them from the affine space itself. In a sense, they were 'already there' in the theory's models. With these structures in place we can consider in more detail the sort of mathematical entity that G is. For a given mass and displacement, G gives us a value for the gravitational field. This means that we can represent G as a function $G : \mathcal{V}_m \times \mathbb{V} \to \mathbb{G}$. It is more convenient to write $G(m) : \mathcal{V}_m \to (\mathbb{V} \to \mathbb{G})$, that is, $G(m)$ is a function from mass value space into functions from $\mathbb{V}$ to $\mathbb{G}$. We can read this as follows: given a particular mass value m, $G(m)$ tells us the contribution of that mass to the gravitational field for any particular displacement vector. This reveals the sense in which G 'names' particular masses: each mass is assigned a *different* function from $\mathbb{V}$ to $\mathbb{G}$, despite the fact that

as G have the structure of an additional quantity.

[23]The approach here is similar to that of Dewar (2020b), who thinks of the gravitational constant as an isomorphism between value spaces. However, I believe that my exposition more explicitly draws out the physical role played by G.

all elements of mass value space are absolutely indiscernible. In physical terms, this just means that the contribution of a particle to the gravitational field at a fixed distance depends on the particle's mass.

It is important to note that the above only concerns the mathematical *representation* of G. I do not want to suggest that the world itself contains a function which somehow causally contributes to the motions of particles. I will consider the physical nature of the gravitational constant in more detail below. Fortunately the mathematical representation of G depends little on what G *is*, but only on its functional role in the equations of motion, viz. equation (6.3). We can then consider what happens when we transform the mathematical representation of G, postponing the question whether such a transformation is physically allowed to the end of this section.

With this notation in hand, we can add $G(m)$ to the models of NG:

$$\text{NG: } \langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathcal{V}_m, \mathbf{g}, x(i,t), m(i), G(m) \rangle$$

We can now consider what happens when ϕ acts on both $m(i)$ and G. Call this transformation 'ALS+G'. The action of ALS+G on the models of NG is as follows:

$$\begin{aligned} &\langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathcal{V}_m, \mathbf{g}, x(i,t), m(i), G(m) \rangle \\ &\qquad \stackrel{\text{ALS+G}}{\rightarrow} \langle \mathbb{A}, \mathbb{T}, \mathcal{B}, \mathcal{V}_m, \mathbf{g}, x(i,t), (\phi \circ m)(i), \phi^* G(m) \rangle \end{aligned} \tag{6.4}$$

Here, ϕ^* is the pullback of ϕ, defined such that $\phi^* G(m) = G(\phi^{-1}(m))$. It is easy to see that this transformation *is* a dynamical symmetry of NG: when ϕ acts on both G and m, we have $\phi^* G(\phi(m(i))) = G(m(i))$. In other words, $G(m)$ remains invariant under such a transformation. Moreover, since G determines the contribution of $m(i)$ to the gravitational field at some fixed distance from i, an ALS+G transformation leaves the gravitational field invariant too. Therefore, escape velocities are preserved. We can see the same result when we consider a numerical representation of G in some particular units. First, we let $f_r(G(m)) = f_r(G) \cdot f_r(m)$. This product tells us the contribution of a particle with mass $m(i)$ to the gravitational field some unit distance away. From the requirement that $\phi^*G(m) = G(\phi^{-1}(m))$ and the fact that $f_r(\phi(m)) = \alpha \cdot f_r(m)$, it follows that $f_r(\phi^*G(m)) = f_r(G) \cdot f_r(m)/\alpha$. Plugging this into equation (6.2) for the escape velocity of a projectile we immediately see that the division of G by α cancels the multiplication of m_2 by the same factor α, such that the overall expression remains the same. Therefore, an ALS+G is empirically undetectable. Nevertheless, worlds related by ALS+G are *physically distinct*. This follows from the postulation of quiddities: when all masses are scaled, each particle is assigned a different mass. Put differently, it is now a different mass value which contributes exactly one unit of force to the gravitational field at one unit of distance away from the source. This existence of physically distinct but empirically equivalent worlds is just the sort of underdetermination that is familiar from debates around the structure of spacetime. There, a renouncement of intrinsic identities solves the issue. In the next section I will suggest a similar solution for underdetermination of ALS+G.

First, let me consider an objection. Martens (2019a, 12) claims that the above is an illegitimate use of G: "It is not at all surprising that if one were allowed to change (the strength of) the laws at will for each possible world one could get any (or at least many) of the evolutions one may have wanted. That is simply not an option within the rules of the game we are playing." Martens claims that changing G is akin to changing the laws, which is against 'the rules of the game'. But who laid out these rules? It seems to me that there are different conceptions of what a constant of nature *is*, and these have different consequences for whether a change in G is 'allowed'. Roughly, a realist can interpret G as an 'aspect of the laws', or as a 'world-property'.[24] On the former view, the value of G is a fixed parameter in

[24]I borrow this distinction from Zee Perry, personal correspondence.

the laws. This view vindicates Martens' contention that a change in G amounts to a change in the laws. On the latter view, on the other hand, the value of G is a physical property of the world as a whole. But if G is a property of the world, then it seems that we could vary G without changing the laws. Instead, we may think of the value of G as an initial or boundary condition, akin to the total energy or momentum of the universe. The fact that Martens' claim depends on a rather non-trivial choice between these metaphysical views means that the possibility of a symmetry which affects G is too quickly set aside.

I will not evaluate the world-property view versus the aspect-of-the-laws view in detail, but let me briefly add some remarks in favour of the latter. Firstly, note that G is a *dimensionful* quantity. Wolff (2020, 175) points out that this supports the world-property view, since dimensions are always dimensions *of* some physical quantity. This is quite unlike the typical parameters in a theory's equations, such as a factor of π or $\frac{1}{2}$. Secondly, on Martens' aspect-of-the-laws view the constant G is a fixed object in the sense of Earman and Norton (1987): the value of G is kept fixed across models, despite the fact that that value is not determined by the theory's laws. Earman and Norton argue that this constitutes an "unnecessary global assumption". If different values for G are equally well-suited to model the phenomena around us, then it is arbitrary to choose on as uniquely physical. Finally, *even if* a change in G amounts to a change in the laws, it is far from clear whether this is enough to resist the claim that NG has no non-trivial symmetries. Dasgupta (2021) coins the term 'empirical symmetries' for transformations which change the laws in a way that preserves empirical content. On the aspect-of-the-laws view, ALS+G constitutes an empirical symmetry. But this is not much better than the situation in which that same transformation is a dynamical symmetry, since either way, the value of G is underdetermined by the empirical data. The choice for a particular value of G then amounts to a 'gauge fixing' condition. For these reasons I will adopt the world-property view, on which a transformation of G is physically possible.

6.5 Sophistication About Mass

The claim that mass scalings, when G is included, are symmetries of NG at first seems to favour the comparativist, who argues that mass ratios are more fundamental than monadic mass properties. But the comparativist cannot easily account for the physical relevance of G either: G itself is *not* a mass ratio. After all, G is a function from $\mathcal{V}_m$ into $\mathbb{V} \to \mathbb{G}$, but according to comparativism the mass value space $\mathcal{V}_m$ itself is unphysical. While absolutism faces the problem of underdetermination as a result of ALS+G, comparativism cannot account for the way in which G determines the gravitational force. Another way to see this is that we only have empirical access to monadic masses via the product $G \cdot m$ (Martens, 2017). But this product quantity is not a mass ratio, so it is not part of the comparativist's fundamental ontology. Comparativism cannot save the phenomena of Newtonian Gravitation.[25]

For these reasons, I will try a different tack. I believe it is true that Martens' version of absolutism has surplus structure, but I do not believe that the correct response is a form of comparativism. Instead, I will advance an *anti-quidditist* view according to which G is *part of the structure of mass space*; or, rather, G represents cross-value space structure. G is a 'world-property' in the sense that it connects these distinct value spaces. On the structuralist view which I propose, there are no quiddities. For any given distance (i.e. for any element of $\mathbb{V}$) there exists instead a privileged mapping from mass value space into gravitational field space: the constant G. For a fixed unit of distance, each element of $\mathcal{V}_m$ is associated with a unique element of $\mathbb{G}$. Instead of $\langle D_m, \leqslant, \circ, \rangle$, I propose that the structure of mass value space is characterised by $\langle D_m, \leqslant, \circ, G \rangle$. I will show that this allows us to avoid the underdetermination of mass

[25]As Baker (2014) notes, the fact that comparativist misses out on empirically relevant data commits them to indeterminism—this is an additional reason to find a third way out of this dilemma.

scalings via strategy analogous to sophisticated substantivalism.

Recall that Martens argues that we need quiddities for the forces to 'latch on' to the right mass magnitudes: without quiddities, all masses are qualitatively identical. Martens claims that quiddities are "required for the forces to be well-defined, in the sense of uniquely matching up instances of initial conditions, including masses, with, say, accelerations" (2519, 2019). However, the way in which forces latch onto masses on my view is directly via G. Once we include G within the theory's kinematical structure, quiddities are no longer necessary. For example, consider again the escape velocity of a projectile. What velocity the projectile requires to escape from an object with mass m_2 depends on the force which the latter exerts on the former. But for a given distance, this force is equivalent to the value of $G(m_2)$. Once we have added G to the structure of our models, there is no longer any need for quiddities: the gravitational force can simply latch onto masses via G to produce the observed accelerations.

In order to see how this view avoids underdetermination, first note there is now a match between the kinematical symmetries of $\mathcal{V}_m$ and the dynamical symmetries of NG, as dictated by (a suitable generalisation of) Earman's symmetry principles. Recall that a value space symmetry is an automorphism of D_m which leaves $\mathcal{V}_m$ invariant. The non-trivial automorphisms of $\circ$ and $\leqslant$ are just the group of Active Leibniz Scalings. But if an automorphism ϕ must leave *all* of $\mathcal{V}_m$ invariant, then we must also require that $\phi^*G = G$. The only automorphism for which this is true is the identity, since each mass m is associated with a *unique* function from $\mathbb{V}$ into $\mathbb{G}$. Therefore, $\mathcal{V}_m$ has no non-trivial symmetries once we include G. On the other hand, the *dynamical* symmetries of NG—that is, those not affecting the structure of mass space—are also trivial. This is the point of the comparativist's bucket thought experiment: a transformation which acts on particle masses alone does not preserve solution-hood. What we see is that identifying G as part of mass value space creates a match between kinematical symmetries and dynamical symmetries, as SP1 and SP2 ordain.

There is a helpful analogy here with Aristotelian spacetime. Aristotelian spacetime results from adding a privileged worldline to a Newtonian spacetime. This additional structure is surplus in Newtonian Gravitation, but would be necessary to express the laws if forces *were* to depend on absolute position. It is exactly because translations of all material bodies are not symmetries of such an Aristotelian theory that we would need to introduce a worldline which represents the 'centre' of the universe. Similarly, it is because of the comparativist's bucket that there must exist some structure which uniquely distinguishes each point of mass space. Furthermore, it is clear that translations of both all material bodies *and* the centre-of-the-universe worldline are symmetries of the Aristotelian theory; correspondingly, models related by such transformations are isomorphic. Likewise, a scaling of all particle masses *and* the gravitational constant G is a symmetry of Newtonian Gravitation, and models related by such a transformation are isomorphic too. G's function is to *qualitatively discern* elements of mass space.

We now come to the critical point. What is the physical effect of a translation of all material bodies *and* the Aristotelian worldline? If we think of the latter as the property of a certain special collection of spacetime points, it seems that such a translation now assigns this property to a different collection of points. The result then is a qualitatively distinct yet empirically indistinguishable state of affairs. It seems more natural to consider the Aristotelian worldline as, in some sense, *individuating* spacetime points: to be *this* spacetime point is just to be the spacetime point which is located in a certain way relative to the centre of the universe. This means that a translation which includes the central worldline has no physical effect. Compare this with Stachel's (1993) view of the metric in GR as a dynamical individuating field. Both of these views are forms of *anti-haecceitism*: spacetime points have no intrinsic identities, but are individuated by their structural properties.

The analogue to anti-haecceitism for quantities is *anti-quidditism*: physical magnitudes have no intrinsic identities, but are individuated qualitatively. I propose anti-quidditism about mass.[26] The idea is that mass magnitudes have no primitive identities, but are qualitatively identified via their pattern of instantiation. Teller (1991, 393) puts the idea as follows:

> What is it to be the property of having a mass of five grams? Perhaps it is no more and no less than bearing certain mass relations to other masses, or possibly to other exemplified masses. On this account we still take there to be individual mass properties, but we take the principle of individuation of an individual mass to be its mass relations to other masses, so that the mass relations between masses become essential to all of them.

In my view, anti-quidditism is simply the analogue of anti-haecceitism for determinates. Since structural relations between mass values are the same in isomorphic models, anti-quidditism implies that the *same* masses are instantiated across those models. The differences between isomorphic SRMs are representationally irrelevant, since they concern quidditistic facts about *which* particular mass values are instantiated. For this reason we cannot obtain a distinct possible world through uniformly scaling all masses while simultaneously scaling G with them. If there were such a world it would differ from the first only with respect to *which* mass properties are instantiated. Since mass properties have no quiddities, this is a distinction without a difference.

Compare this with a situation in which there are quiddities. There is then a further fact about *which* mass is dynamically privileged *even when the products* $G \cdot m$ *are fixed.* Since the empirical content of NG consists in these product quantities, it

[26]The possibility of anti-quidditism in response to internal symmetries has already been noted in Arntzenius (2012), Dewar (2018) and Martens and Read (2020). Teller (1991) raises the possibility of anti-quidditism about mass specifically.

follows that such quidditistic facts are underdetermined by the data. Martens (2019b) calls this phenomenon *inexpressible knowledge*: we can know the mass of a particle in terms of its dynamical behaviour, but we cannot know *which* mass a particle has in terms of its quidditistic nature. But on an account without quiddities, there is no meaningful answer to the question *which* mass a particle has beyond a position within mass structure. Therefore, the sophisticated account of mass space avoids the underdetermination which troubles absolutism.

6.6 Kinematic comparativism

Martens claims that we have *inexpressible knowledge* of mass magnitudes. I have argued that this is not the case on a sophisticated account of mass, but Martens believes that inexpressibility is a direct consequence of the fact that mass is a *dimensionful* quantity. If so, there must be something wrong with the sophisticated account, which after all asserts that intrinsic mass is both real and dimensionful. But Martens' claim is false. The problem is that Martens incorrectly argues that inexpressibility follows from a condition called *kinematic comparativism*, which he states is characteristic of dimensionful quantities. But it is neither the case that kinematic comparativism is true of all dimensionful quantities, nor does kinematic comparativism entail inexpressibility.

First, here is a statement of kinematic comparativism:

> For any dimensionful determinable, such as mass, the magnitude predicated of any particle can only be reported or expressed, non-dynamically, in terms of how this magnitude relates to the magnitude of another particle having the same determinable property. (Martens, 2019b, 10)

Martens claims that "we should take it not merely as a feature but as a definition of the dimensionfulness of a determinable that it is kinematically comparative". However, there are examples of dimensionful quantities for which kinematic comparativism is false. Consider once more Aristotelian spacetime. The models of Aristotelian spacetime are of the form $\langle M, h_{ab}, t^{ab}, \nabla_a, \sigma^a, \phi, \rho, \xi \rangle$, where ξ is a worldline in M. Unlike Newtonian spacetime, we can express the location of a particle in terms of its distance from the 'centre' ξ. For example: "the Eiffel Tower is located at the centre of the universe at time t". We can thus *express* the location of (material) bodies without recourse to the location of *other* bodies.

Of course, it is impossible to know where the centre of the universe is, since ξ is dynamically inert—so in practice, we can express but not *report* the location of some particle in relation to this structure. But consider a modification of the dynamics: any particle which crosses the centre of the universe EXPLODES (or, for something more realistic, consider adding a central force which radiates outwards from ξ). If we are (un)lucky enough to observe a particle hitting the central worldline, we know exactly where the centre is, hence where bodies are located with respect to it. Therefore, kinematic comparativism is false in this universe, at least for position. But position is dimensionful in the sense that we still need to choose an arbitrary coordinate system to assign numbers to spacetime points. This is a counterexample to Martens' claim that dimensionfulness implies kinematic comparativism.

Furthermore, even *if* kinematic comparativism is true, this does not necessarily imply a failure of expressibility. This is *pace* Martens, who writes that "[t]his failure of expressibility should not surprise us: this is the content of kinematic comparativism, this is what it means, by definition, to be a dimensionful magnitude" (33). Consider Newtonian spacetime. It is true that we can only express the location of a particle with respect to another particle, because Newtonian space is homogeneous. Martens believes this to imply that we have inexpressible knowledge of a particle's absolute location because we cannot detect *which* point it occupies—i.e. the haecceity of a spacetime point.[27] But according to sophisticated substantivalism there are no

[27]In fact, it is not entirely clear to me whether this knowledge *is* inexpressible. We can perfectly express such knowledge with the use of indexicals: the Eiffel Tower is so-times as massive as *that*

haecceities. Instead, points in space are individuated in terms of the bodies which occupy them. There is no further response to the question "Which point does this particle occupy?" other than the answer "The point which is occupied by this particle and stands in such-and-so relations to other points occupied by other particles". In fact, sophistication *never* leads to inexpressible knowledge, since what is inexpressible according to Martens are exactly the intrinsic identities which sophistication rejects.

The discussion carries over to our sophisticated account of mass. Since mass value space is similar to Aristotelian spacetime in that its values are absolutely discerned dynamically, kinematic comparativism fails. In other words, we can express a particle's mass in terms of the force associated to it via G, rather than in terms of the mass of some other particle. Furthermore, the fact that our account of mass is sophisticated means that there is no inexpressible knowledge of *which* masses are instantiated: which mass a particle has depends on the structural features of its location in mass value space, but not on any quiddities. This is a strength of the sophisticated account: that there is no inexpressible knowledge means that there are no in-principle undetectable facts about which masses are instantiated.

6.7 Conclusion

In §2, I presented the debate between absolutism and comparativism about mass in the form of a dilemma: either Active Leibniz Scalings *are* symmetries of Newtonian Gravitation, in which case absolutism is committed to redundant structure; or Active Leibniz Scalings are *not* symmetries, in which case comparativism does not contain enough structure to sustain the dynamics of Newtonian Gravitation. The dilemma is a false one: as I have argued, there is a consistent position which acknowledges that Active Leibniz Scalings are symmetries, once the gravitational constant is accounted

object, for instance the standard mass of the SI. But as I argue below, even if Martens is correct that this is a case of inexpressible knowledge, then this is not a consequence of the fact that mass is dimensionful but rather of the posit of quiddities.

for, but which is nevertheless realist about absolute mass. This position is called *sophistication about mass.* Analogous to sophistication about spacetime, masses are individuated qualitatively via the structure of mass value space, which includes the relations $\leqslant$ and $\circ$ as well as the mapping G. The latter endows mass space with a dynamically privileged mapping onto the value space of the gravitational field. Quiddities are uncalled for, since the forces can latch onto masses directly via G.

The benefits of sophistication about mass are as follows. Compared to absolutism, sophistication is committed to less structure, since it rejects the existence of quiddities. As a result, sophistication does not countenance the possibility of a physically distinct world in which all masses and the gravitational constant G are scaled by a constant factor: such worlds differ merely quidditistically. This means that sophistication avoids the problem of underdetermination associated with non-trivial symmetries. Sophistication also avoids many of the pitfalls of comparativism, for instance, it can explain the so-called *conspiracy of mass relations.* Recall from Chapter 3 that the existence of intrinsic masses ensures that mass relations are transitive, so $m_{ij} = m_{ik} \cdot m_{kj}$. If the comparativist's mass relations are really ratios of absolute masses, this relation is reduced to a mathematical theorem. The same explanation is still valid on a sophisticated account. Of course, we can no longer appeal to a *literalist* understanding of mass magnitudes as uniquely represented by certain numbers-*cum*-units. But the mass space $\mathcal{V}_m$ has a certain structure which allows us to prove representation theorems to the effect that these relations are faithfully represented by various mass scales. For example, the second-order relation 'twice as much as' holds between two masses iff $m_a = 2m_b$. Importantly, *any* representation of the fundamental mass relations satisfies the transitivity of mass relations. If this were not the case—if the equality were to hold, say, in kilograms, but not in pounds—then we would not consider it an objective fact. What explains TMR on a sophisticated account is the fact that mass properties stand in certain relations to each other and that *however we represent this structure numerically*, for any three masses m_i, m_j and m_k it is the case that $\frac{m_i}{m_k} = \frac{m_i}{m_j} \cdot \frac{m_j}{m_k}$.

In conclusion, sophistication about mass is a novel 'best of both worlds' position which escapes from the false dilemma of the absolutism versus comparativism debate. There is no reason to stop there: the process of sophistication can be applied to other symmetry-variant quantities, such as the electrostatic potential. In the next chapter, I will apply sophistication to *local* symmetries, with the Aharonov-Bohm effect as a case study.

Chapter 7

Case Study II: Gauge Quantities

The symmetries discussed so far—shifts and boosts, uniform mass scalings, and so on—are all instances of *global* symmetries. In this chapter I turn my attention to *local* symmetries. The difference is that the former depend on a finite number of parameters, whereas the latter depend on infinitely many parameters. For this reason, Gomes (2019) also calls global symmetries 'rigid' and local symmetries 'malleable'. I will stick to the traditional terminology.[1] In particular, I will focus on local symmetries of internal degrees of freedom, also called *gauge symmetries*.[2] The main example in this chapter is the gauge freedom of scalar electrodynamics.

With local symmetries, many of the same issues arise as with the global symmetries discussed before, such as the presence of undetectable quantities. The reasoning is analogous: since gauge transformations are unobservable, the existence of such symmetries implies a form of underdetermination. Consequently, the values of quantities that vary under those symmetries are unmeasurable. Furthermore, the local character of gauge symmetries entails a particularly harmful form of indeterminism,

[1]For a slightly less formal characterisation of the difference, see Wallace (2002).

[2]The terminology here is particularly confusing: among philosophers, it is not uncommon to use the term 'gauge symmetry' for any symmetry that relates physically equivalent states, whereas physicists reserve the term for local symmetries; see Weatherall (2016b). Furthermore, there is some debate over whether external symmetries, such as the diffeomorphism invariance of General Relativity, count as gauge; see, for example, Wallace (2015) and Dewar (2020a). I will use the term 'gauge symmetries' in the technical sense to denote local internal symmetries.

analogous to the (alleged) indeterminism that features in the Hole Argument in General Relativity. Since gauge symmetries are local, there exist transformations whose action is trivial up to some time t, but non-trivial thereafter. This means that the physical facts up to t do not determine the facts after t: a failure of determinism. This is particularly problematic because these divergent futures are observationally equivalent, so the world does not *appear* indeterministic to us.

As I explained in the introductory chapter, there are three strategies for interpreting symmetry-related models (SRMs). The first is literalism: SRMs represent physically distinct states of affairs. In order to avoid indeterminism, however, one has to supplement literalism with some further claim. For example, one could claim that only one out of an equivalence class of SRMs represents a *possible* state of affairs, namely one that satisfies a certain gauge condition. This claim in effect elevates a particular gauge fixing condition to become an additional law.[3] But such strategies face various problems—most importantly, that the particular choice of 'gauge law' is essentially arbitrary. Therefore, I will not further consider literalism here. Instead, I will focus on the choice between *reduction* and *sophistication*.[4] Recall that reduction aims at a *reduced* theory formulated solely in terms of invariant quantities, such that each model of the reduced theory uniquely corresponds to an equivalence class of SRMs of the old theory. Sophistication, on the other hand, aims at a *restructured* theory such that the new theory's SRMs are *isomorphic*. Since isomorphic models are structurally equivalent, the latter method allows one to interpret SRMs *anti-haecceitistically* or *anti-quidditistically* as also physically equivalent (more on this below). I will defend sophistication over reduction as the correct interpretation of gauge theories. Specifically, I will argue that sophistication makes most sense of physicists' use of the *fibre bundle formalism* in modern formulations of gauge theories, which for a reductionist

[3] This is similar to Maudlin's (1998) 'ONE TRUE GAUGE' approach, although Maudlin does not go so far as to consider the gauge fixing condition as a law of nature.

[4] Dewar (2019) was the first to explicitly draw this distinction; see also Martens and Read (2020) and Jacobs (2021).

seems to possess excess structure.[5]

In §7.1, I use the Aharonov-Bohm effect to illustrate how symmetries pose a problem for the interpretation of gauge theories. In §7.2, I survey and criticise various reductionist responses to this problem, including Healey's (2007) holonomy realism and Wallace's (2014a) deflationary account of the Aharonov-Bohm effect. In §7.3, I introduce the fibre bundle framework, which, I argue, is the appropriate setting for a sophisticated account of gauge symmetries. In §7.4, I consider gauge symmetries from the perspective of the fibre bundle picture. In §7.5, I consider the metaphysics of fibre bundles on a sophisticated account. Specifically, I argue that anti-quidditism about gauge quantities allows us to interpret SRMs as physically equivalent. In §7.6, I discuss whether my account is *local* in various senses. I argue that the fibre bundle account is both local and separable, but that it also implies a form of holism about physical explanations. §7.7 concludes.

7.1 The Aharonov-Bohm Effect

The Aharonov-Bohm effect is essentially a modified double-slit experiment in which a solenoid is placed between the plate and the screen. We assume that the solenoid is impenetrable. The Aharonov-Bohm effect refers to the fact that the interference pattern on the screen changes when we let a current run through the solenoid. Specifically, the peaks are shifted within the interference envelope. This is the case despite the fact that the electromagnetic field vanishes outside the solenoid, whereas the particles cannot penetrate the solenoid. Hence, we cannot simply understand the effect as a result of the force field acting locally on the matter field. This led Aharonov and Bohm to posit the *four-potential*, which does not vanish outside the solenoid, as causally responsible for the effect (Aharonov and Bohm, 1959).

[5] I thus agree with Dewar's (2019) third example of sophistication. I elaborate on his example in offering both a more detailed technical case study, and by filling out the accompanying metaphysical picture.

In more detail,[6] recall that the electromagnetic field tensor $F_{\mu\nu}$ can be expressed in terms of the electromagnetic four-potential A_μ:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{7.1}$$

In terms of $F_{\mu\nu}$ and A_μ, the Lagrangian of scalar electrodynamics for a single particle with wavefunction ϕ in an electromagnetic field A_μ is

$$\mathcal{L} = (D_\mu \phi)(D^\mu \phi)^* - m^2 |\phi|^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{7.2}$$

where $D_\mu \coloneqq \partial_\mu + iqA_\mu$ and q is a scalar quantity which denotes the field's charge. Here, ϕ denotes a classical matter field. The Lagrangian is invariant under the following *gauge transformation*:

$$\begin{aligned} A_\mu &\to A_\mu - \frac{1}{q}\partial_\mu \alpha(x) \\ \phi &\to e^{iq\alpha(x)}\phi \end{aligned} \tag{7.3}$$

where $\alpha(x)$ is a function of the spacetime coordinates x.

For the Aharonov-Bohm effect, we consider the matter field

$$\phi(x) = \phi_I(x) + \phi_{II}(x), \tag{7.4}$$

where ϕ_I and ϕ_{II} are the components of the field that pass through the left and right slit respectively. Let P denote the source of the field, and Q an arbitrary point on the screen. Relative to the (arbitrary) gauge we started with, we can always re-choose a gauge such that $\alpha(P) = 0$. For a matter field within a simply-connected region in which $F_{\mu\nu}$ vanishes we can similarly always re-choose a gauge such that $A_\mu = \partial_\mu \alpha$

[6] The exposition draws from Healey (2007) and Brown (2016).

and hence write:

$$\alpha(x) = \alpha(x) - \alpha(P) = \int_P^x d\alpha(x) = \int_P^x \partial_\mu(\alpha) dx^\mu = \int_P^x A_\mu dx^\mu \tag{7.5}$$

With this result, we can re-express the transformation of ϕ in (7.3) as follows:

$$\phi(x) \to e^{iq \int_P^x A_\mu dx^\mu} \phi(x) \tag{7.6}$$

When applied to the particular field of interest in (7.4), this results in:

$$\phi'(Q) = e^{iq \int_I A^\mu dx_\mu} \ \phi_I(Q) + e^{iq \int_{II} A^\mu dx_\mu} \phi_{II}(Q) \tag{7.7}$$

$$= e^{iq \int_I A^\mu dx_\mu} \big(\phi_I(Q) + e^{iq \oint A^\mu dx_\mu} \phi_{II}(Q)\big), \tag{7.8}$$

where the second term in (7.8) contains the loop integral of A^μ around a path from P to Q and back which encloses the solenoid. Since $e^{iq \int_I A^\mu dx_\mu}$ is a unit complex number, we can safely ignore it. From an application of Stokes' theorem, it then follows that

$$\oint_{\partial D} A^\mu dx_\mu = \int_D F_{\mu\nu} dS \tag{7.9}$$

where D is the surface bounded by paths I and II respectively. But of course the expression in (7.9) is just the flux Φ through D. Therefore, our final expression for the gauge transformation of ϕ is:

$$\phi(Q) \to \phi_I(Q) + e^{iq\Phi} \phi_{II}(Q) \tag{7.10}$$

The puzzling fact here is that Φ depends on $F_{\mu\nu}$, which vanishes outside the solenoid. The phase shift of the matter field seems to causally depend on a field with which it cannot interact locally.

For this reason, physicists often consider the four-potential A_μ (which does not vanish outside the solenoid) as a physically real field which interacts with the wave-function. Aharonov and Bohm, for instance, wrote that

> [I]n a theory involving only local interactions [...] the potentials must, in certain cases, be considered as physically effective, even when there are no fields acting on the charged particles. (Aharonov and Bohm, 1959, 490)

Similarly, Feynman posits local interaction as a necessary condition for a field to count as physical:

> In our sense then, the A-field is "real." You may say: "But there *was* a magnetic field." There was, but remember our original idea—that a field is "real" if it is what must be specified *at the position* of the particle in order to get the motion. The B-field in the whisker acts at a distance. If we want to describe its influence not as action-at-a-distance, we must use the vector potential. (Feynman et al., 1964, Ch. 15)

The fact that A_μ, unlike $F_{\mu\nu}$, is not a gauge-invariant quantity is problematic for familiar reasons. Solutions to (7.2) related by the transformations in (7.3) are observationally equivalent. Therefore, reifying the four-potential implies the underdetermination of the theory's models by the empirical data. Furthermore, there exist gauge transformations that act as the identity before some time t but non-trivially thereafter (i.e. $\alpha(t, \vec{x}) \neq 0$ iff $t' > t$). The existence of such transformations seems to imply that electrodynamics is indeterministic. This form of indeterminism is particularly problematic because the difference between outcomes is unobservable, so the indeterminism does *not* occur at the level of observables (unlike the indeterminism of quantum mechanics). The literal approach to A_μ according to which gauge-related models represent distinct states of affairs thus has several undesirable features.

7.2 Failures of Reduction

There are various alternatives to A-realism that aim to avoid this indeterminism. The three main proposals are *F-realism*, *holonomy realism* and *field monism*. I will discuss each of these in turn. As I will show, these are all instances of *reduction*. Recall that reduction aims to avoid underdetermination by constructing a new theory—the reduced theory—in terms of invariant quantities of the old theory, such that there exists a unique correspondence between models of the new theory and equivalence classes of SRMs of the old theory. I will argue that the below instances of reduction are unsatisfactory. Specifically, the no miracles argument against comparativism from Chapter 3 applies to the first two of the proposals discussed below. The third proposal, meanwhile, offers a satisfactory reduction of $U(1)$-electrodynamics, but does not extend to more complex gauge theories.

7.2.1 F-Realism

F-realism is the view that the electromagnetic field-tensor $F_{\mu\nu}$ is fundamental. F-realism is a form of reduction, since $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$ is an invariant quantity defined in terms of A_μ; there is a unique correspondence between models in terms of $F_{\mu\nu}$ and equivalence classes of gauge-related models in terms of A_μ (Weatherall, 2016b, 1041-42). Historically, of course, F-realism was not obtained as the reduct of A-realism. Rather, F-realism was the *de facto* account of electrodynamics on which the Aharonov-Bohm effect put pressure. But I consider F-realism as a reduction of A-realism in the spirit of rational reconstruction.

F-realism faces three problems. The first is the well-known fact that an explanation of the Aharonov-Bohm effect in terms of the Faraday tensor implies a violation of the principle of Local Action (Healey, 1997). Since there is no overlap between the electromagnetic field and the wavefunction, the former can only act on the latter at a distance. This is universally seen as sufficient reason to reject F-realism (Feynman et al., 1964; Healey, 1997; Belot, 1998; Lyre, 2004; Wallace, 2014a). I concur with this verdict.

The second issue is that in non-simply connected spaces, the empirical facts are underdetermined by the field tensor (Belot, 1998). In different words, the same values for $F_{\mu\nu}$ are consistent with different experimental results. We can see this if we model the solenoid in the Aharonov-Bohm effect as a 'gap' in space. Since $F_{\mu\nu}$ vanishes outside the solenoid, distributions of A_μ which are *not* related by a gauge transformation correspond to the same electromagnetic field. But the observed Aharonov-Bohm shift is different in these situations, which means that the values of $F_{\mu\nu}$ itself fail to fully determine the observable facts.

The third problem, less well-known but more closely related to the distinction between reduction and sophistication, is that F-realism implies a 'cosmic coincidence' (Dewar, 2019, 498). This follows from the fact that $F_{\mu\nu}$ is in some sense a *comparative* quantity: the derivative which occurs in its definition means that its values depend on the values of A_μ at infinitesimally close points. I argued in Chapter 3 that any comparative form of reduction involves cosmic coincidences. In this case, the coincidence is the Gauss-Faraday law $\partial_{[\mu}F_{\nu\rho]} = 0$. When expressed in terms of A_μ, this law reduces to a mathematical theorem. But if A_μ is merely a mathematical abstraction, then Gauss' law is an additional postulate of the theory. In that case, it is mysterious that $F_{\mu\nu}$ behaves *as if* it is the exterior derivative of a four-potential, even when it is in fact a fundamental quantity. Therefore, F-realism incurs an *explanatory loss* in addition to its violation of Local Action.

7.2.2 Holonomy Realism

While F-realism has virtually no advocates, *holonomy realism* is a popular interpretation of electrodynamics (Belot, 1998; Lyre, 2004; Healey, 2007). According to holonomy realism, the so-called *holonomies* H of A_μ are fundamental:

$$H(l) = \exp\left(-iq \oint_l A_\mu dx^\mu\right) \tag{7.11}$$

On this picture, there is a fundamental *non-localised* property associated to every closed curve in spacetime. Holonomies are not composed of the field-values at each spacetime point, but attach to curves as a whole. When a matter field interacts with these holonomies, it does so 'at once' around a loop. In this sense, interactions for holonomy realism are local, since the holonomies overlap with the matter field.

Like F-realism, holonomy realism is an instance of reduction. Holonomies are gauge-invariant quantities, and an assignment of holonomy values to closed curves uniquely corresponds to an equivalence class of gauge-related A-fields (Barrett, 1991; Rosenstock and Weatherall, 2016).[7] Moreover, holonomies are comparative quantities. The holonomy around a closed path l is a function of the path integrals of A_μ over any pair of open paths l_1, l_2 which compose l:

$$H(l) = e^{-iq \oint_l A_\mu dx^\mu} = e^{-iq \int_{l_1} A_\mu dx^\mu} e^{-iq \int_{l_2} A_\mu dx^\mu} \tag{7.12}$$

Thus, $H(l)$ is relational in the sense that it is defined as a function of *pairs* of objects (in this case, paths), just as distance is defined as a function of pairs of particles. It is for this reason that Arntzenius (2012) calls Healey's view 'gauge relationism'.

Holonomy realism faces three main problems. The first is that the laws are still expressed in terms of A_μ, not in terms of H. As far as I am aware, no one has yet succeeded in writing down a Lagrangian for scalar electrodynamics solely in terms of these holonomies. But this should instill a feeling of suspicion: if holonomies are the fundamental quantities of nature, why is it that we cannot use them to express nature's fundamental laws? Belot (2003b) (who himself does not endorse this type of

[7] In non-Abelian theories, holonomies are not always invariant. Instead, one can choose the *Wilson loops* as fundamental.

argument) relates this view to Quine, who argued that we should regard the posits of our best theory of the world as real. If the theory's laws refer to an electromagnetic potential, then that is what we are ontologically committed to.

The second issue is that an ontology of holonomies is *non-separable*: the intrinsic facts about a region X and the intrinsic facts about another region Y do not uniquely determine all intrinsic facts about the joint region $X \cup Y$. We can easily see this when we consider two partially overlapping regions X and Y close to the solenoid in the Aharonov-Bohm effect. Since X does not enclose the solenoid the flux through its surface is zero, and likewise for Y. But now consider the union $X \cup Y$. This region *does* enclose the solenoid, and so its boundary has a non-zero holonomy value. Therefore, the intrinsic facts about X and Y on their own do not suffice to determine the intrinsic facts about $X \cup Y$: separability fails.

It is debatable whether this is a real issue. Perhaps the world just *is* non-separable. Quantum entanglement already gives us plenty of reason to think this is the case.[8] I therefore find the third issue associated with holonomy realism more serious. The issue is that holonomy realism has to postulate certain structural patterns in the instantiation of holonomies as brute facts. This 'cosmic coincidence' is similar to the one we encountered above for F-realism (and follows the pattern of Chapter 3). Specifically, the holonomies of distinct loops l_1 and l_2 must satisfy the following relation:

$$H(l_1 \circ l_2) = H(l_1)\, H(l_2) \tag{7.13}$$

Here, $l_1 \circ l_2$ is the concatenation of the two loops, that is, the result of first going around l_1 and then going around l_2. Call this feature *composite loop multiplication*

[8] Although as Maudlin (1998) argues, the entanglement here is of a different nature. In addition, Dougherty (2017) argues that an 'untruncated' version of the holonomy view *is* separable, so that holonomies and fibre bundles are on a par in this respect.

(CLM).[9] The fact that CLM holds guarantees that we can represent holonomies as exponentials of loop integrals of A_μ. If CLM fails then it is no longer the case that any assignment of holonomies uniquely defines a potential field up to gauge-equivalence. Put differently, CLM makes it look *as if* holonomies supervene on local fields. Or, as Arntzenius (2012, 195) puts it, "a fairly obvious explanation of why [CLM] hold[s] is that the map H is, roughly speaking, the integration of a connection around a loop". In more detail, if we define $H(l) = \exp\left[-i/\hbar \oint_l \mathbf{A} \cdot d\mathbf{x}\right]$, then

$$\begin{aligned} H(l_1)\, H(l_2) &= \exp\left[-iq \oint_{l_1} A_\mu dx^\mu\right] \cdot \exp\left[-iq \oint_{l_2} A_\mu dx^\mu\right] \\ &= \exp\left[-iq \oint_{l_1 \circ l_2} A_\mu dx^\mu\right] \\ &= H(l_1 \circ l_2) \end{aligned} \tag{7.14}$$

Hereby, we can easily explain the fact that CLM holds if we posit the existence of a local four-potential. On the holonomy interpretation, on the other hand, there is apparently nothing which guarantees CLM. For example, $H(l_1)$ could have been slightly lower than it actually is. In that case, $H(l_1 \circ l_2)$ would either have been different too, or it would remain the same. In the former case CLM could easily have failed, so it seems a conspiracy that the holonomies just happen to be related in this way. In the second case, on the other hand, the value of $H(l_1)$ and $H(l_2)$ are counterfactually connected, despite the fact that l_1 and l_2 only partially overlap. This counterfactual action-at-a-distance is at least as puzzling as non-separability, if not more so. Either way, CLM is a cosmic coincidence which holonomy realism cannot explain.

In response to this objection, Healey appeals to the *loop supervenience* of holon-

[9] In the case of non-Abelian theories, the gauge-invariant quantities satisfy a more complicated set of relations called the Mandelstam identities. These relations seem even more conspiratorial than the one expressed in (7.13). I thank Henrique Gomes for this point, which is also made in Arntzenius (2012, §6.4).

omy properties: "the holonomy properties of any loop $\otimes_i L_i$ are determined by those of any loops L_i that compose it" (Healey, 2007, 123). In other words, composite loops are *not* fundamental, since they are composed of smaller loops. Of course, the same is true for these smaller loops, which are themselves composed of even smaller ones. There is no end to this chain of supervenience, hence no smallest fundamental unit. But let's set aside potential worries this infinite regress may invite. The question then becomes: can loop supervenience explain CLM? I argue that it cannot. For the sense in which smaller loops 'compose' larger loops is unusual. It is not the case that smaller loops constitute composite loops in the same way that the mass of a system's parts determines the mass of the system as a whole, for example. In the latter case, we are simply concerned with mereological composition. But composite loops are not the mereological composite of smaller loops, since the small loops contain parts which do not overlap with any part of the composite loop. Furthermore, loop composition does not have the correct formal properties to count as mereological fusion. For example, mereological parthood is anti-symmetric: distinct objects cannot be proper parts of each other. At the same time, we can prove that loop composition *is* symmetric. Let l_1^{-1} stand for the loop which goes around l_1 in the opposite direction; then $H(l_1 \circ l_2)\, H(l_1^{-1}) = H(l_2)$. If loop composition is identical to mereological fusion, then this would imply that $l_1 \circ l_2$ is part of l_2. But since l_2 is also part of $l_1 \circ l_2$, this implies that the whole is part of a part of the whole: a contradiction. Loop composition fails to determine a unique supervenience *base*, since each loop is part of each other's base. For this reason, loop supervenience cannot explain composite loop multiplication.

7.2.3 Field Monism

Wallace (2014b) introduces an interpretation of electrodynamics which is both local and separable. Since on Wallace's account "the electromagnetic and scalar fields cannot be thought of as separate entities", but jointly "[represent] aspects of a single entity" (15), I will call this view *field monism*. Instead of a complex matter field, Wallace's fundamental fields are the real scalar field $\rho = |\phi|$ and the covariant derivative field $D_\mu\theta = \partial_\mu\theta - A_\mu$. Since both ρ and $D_\mu\theta$ are invariant quantities defined in terms of the old theory's variant quantities, field monism presents us with another example of reduction. The joint distributions of these fields uniquely correspond to equivalence classes of gauge-related models of electrodynamics. Moreover, Wallace's account is comparative in the same sense in which F-realism was, since $D_\mu\theta$ is defined in terms of the partial derivative of θ, which means that in order to know the value of $D_\mu\theta$ at some point one needs to know the values of θ at infinitesimally close points. For this reason, we would expect similar cosmic conspiracies to appear.

But the main issue with field monism is that it does not easily extend to more complex gauge theories. Wallace (2014a, 17) himself admits that this is a problem, writing that "in general, I know of no comparably simple set of local gauge-invariant quantities in the non-Abelian case that can serve as a gauge-invariant representation". This suggests that it is no more than a fortunate accident that we can represent the simple $U(1)$ gauge theory Wallace considers in terms of a unique set of gauge-invariant local quantities. Moreover, even more complex *Abelian* theories need not have a unique gauge-invariant representations (as Wallace admits). Consider a pair of complex-valued fields ϕ and χ with different charges, both of which are coupled to A_μ. The Lagrangian of this system is:

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{7.15}$$

where

$$\mathcal{L}_1 = (\partial^\mu + ie_1A^\mu)\phi^\star(\partial_\mu - ie_1A_\mu)\phi \tag{7.16}$$

$$\mathcal{L}_2 = (\partial^\mu + ie_2A^\mu)\chi^\star(\partial_\mu - ie_2A_\mu)\chi \tag{7.17}$$

$\mathcal{L}_1$ is just the Lagrangian of a single complex field coupled to the gauge field. Therefore, we can follow Wallace and rewrite it in terms of the gauge-invariant quantities $\rho = |\phi|$ and $D_\mu\theta_1 = \partial_\mu\theta_1 - A_\mu$, where θ_1 is the phase of ϕ. But since we have 'replaced' A_μ with $D_\mu\theta_1$ in $\mathcal{L}_1$, we have to make the same substitution in $\mathcal{L}_2$. This results in an ontology which consists of a real-valued field ρ_1 with charge e_1, a complex-valued field χ with charge e_2, and a single connection $D_\mu\theta_1$. It is easy to confirm that this ontology is indeed gauge-invariant.

Alternatively, we could have started with $\mathcal{L}_2$ and written *that* Lagrangian in terms of a real-valued field and a connection. This results in an ontology which consists of a complex-valued field ϕ with charge e_1, a real-valued field ρ_2 with charge e_2, and a connection $D_\mu\theta_2$. Crucially, these are different ontologies. Although both theories posit a real-valued field and a complex-valued field, the charges of these fields differ. On the first interpretation the charge of the real scalar field is e_1 and that of the complex field is e_2, while on the second interpretation this is the other way around. Therefore, on one way of understanding Wallace's observation with respect to the unitary gauge, it implies a form of theoretical underdetermination: the choice between these two ontologies is arbitrary.[10] This is hardly better than the underdetermination of theory by the empirical data implied by the existence of gauge symmetries. Therefore, in these more complex scenarios Wallace's account for finding a gauge-invariant representation is inadequate.

7.3 Fibre Bundle Accounts

Instead of reduction, we can try *sophistication* as an approach to gauge symmetries. Recall that the aim of sophistication is to restructure a theory's models such that symmetry transformations become isomorphisms. In the case of local symmetries the

[10]Recall that we saw another instance of such underdetermination in §1.3 of the introduction. It is an interesting question for further work to discuss under which circumstances such underdetermination arises.

appropriate mathematical structures are *fibre bundles.* In this section, I introduce the formalism of fibre bundles. Because my focus is on the physical interpretation of this formalism I do not aim for a comprehensive treatment of the mathematics; for more details, see *inter alia* Baez and Muniain (1994); Isham (1999); Healey (2007); Weatherall (2016a).

I am not the first to suggest that the fibre bundle formalism can aid our interpretation of electrodynamics: Leeds (1999); Nounou (2003); Maudlin (2007); Weatherall (2016a) all appeal to it in one way or another. My account differs from theirs on a few significant issues. Leeds, Nounou and Maudlin all aim to draw metaphysical conclusions from the fibre bundle formalism. Nounou, for instance, argues that topological features of the fibre bundle explain the Aharonov-Bohm effect, whereas Maudlin emphasises the consequences of a fibre bundle picture for the status of universals. However, none of their accounts address the underdetermination associated with gauge symmetries. Leeds explicitly acknowledges that his picture "traffics heavily in non-measurable properties and quantities" (613); Nounou similarly admits that "we also part with determinism in the sense that [...] there are infinitely many gauge fields corresponding to one electromagnetic field" (193). The aim of a sophisticated account of gauge theories, on the other hand, is to rid ourselves of underdetermination by interpreting SRMs as physically equivalent. The proposal thus comes closer to Weatherall, who notes that gauge transformations in electromagnetism are formally similar to those in General Relativity. In the latter case we already have a deflationary interpretation of gauge symmetries (viz. sophisticated substantivalism), so Weatherall suggests that we apply the same interpretation to electrodynamics. But Weatherall remains silent on the *metaphysical* picture this implies: what are the fundamental entities of electrodynamics, and how do they interact? I aim to present a perspicuous metaphysical picture that corresponds to this formalism.

In the remainder of this section I set out the mathematical details of the fibre bundle formalism. For an intuitive idea, start with the concept of a 'value space' discussed in earlier chapters. The value space of a quantity such as mass has a certain structure. We can consider the value space $\mathcal{V}_A$ of the four-potential to have the structure of a four-dimensional vector space. According to A-realism, spacetime points are mapped into this vector space via a function $A(x) : \mathcal{M} \to \mathcal{V}_A$. In Chapter 5, I defined (internal) symmetries as transformations of a theory's models induced by bijections of their value spaces.[11] For example, if $\langle \mathcal{M}, \mathcal{V}_A, A(x) \rangle$ is a model of the theory then a (smooth) bijection ϕ of $\mathcal{V}_A$ is a symmetry iff $\langle \mathcal{M}, \mathcal{V}_A, \phi_* A(x) \rangle$ is also a model. But this set-up does not allow for local symmetries, because $\mathcal{V}_A$ carries no information about *where* particular field-values are instantiated: a transformation ϕ necessarily applies everywhere equally. Concretely, suppose that before a local symmetry transformation is applied the value of A_μ is the same at distinct points x_1 and x_2. In general, there exist local symmetries that non-trivially act on $A(x_1)$ but leave $A(x_2)$ the same. But there is no transformation on $\mathcal{V}_A$ which implements such a symmetry. For if there were such a transformation, then there would exist a bijection ϕ such that $\phi(A(x_1)) \neq \phi(A(x_2))$. But since $A(x_1) = A(x_2)$, the value of $\phi(A(x_1))$ *is* equal to the value of $\phi(A(x_2))$. It follows that we cannot implement local symmetries as the lift of a value space bijection when a field is mapped into the same 'universal' value space at every point. This may seem a flaw in my definition of symmetries, but in fact it is a virtue: that a universal value space for A_μ does not possess enough structure to account for local symmetries (in my sense) illustrates the way in which fibre bundles are a natural setting for gauge theories. The essential idea of a fibre bundle account is to assign a local 'copy' of $\mathcal{V}_A$ to each spacetime point. Instead of a single value value $\mathcal{V}_A$, there is a different value space $\mathcal{V}_A^x$ for each $x \in M$. These local value spaces are called *fibres*. Since each fibre is distinct, we can vary the field values independently at each point. The collection of all fibres forms a manifold, called the

[11] This definition is a little more general than the one found in Dewar (2020a), but our definitions coincide for fibre bundle theories.

fibre bundle. As we will see, local symmetries are (vertical) automorphisms of the bundle, which I will define below.

I will now canvass this picture in some more detail. I provide definitions of all mathematical concepts, but my aim is to explain the essential concepts in relatively non-mathematical terms. Let's start with the concept of a fibre bundle:

Definition (Fibre Bundle)**.** *A* fibre bundle *is a triple* (E, π, M) *where* E *and* M *are smooth manifolds and* $\pi : E \to M$ *is a continuous map, with a space* F *(called the 'typical fibre') such that for each* $x \in M$ *there exists an open neighbourhood* $U \subseteq M$ *and a homeomorphism* $h : U \times F \to \pi^{-1}(U)$ *for which* $\pi(h(x, y)) = x$*.*

So, a bundle consists of a pair of manifolds and a projection π that defines which points on the bundle lie 'above' which points on the base manifold. The bundle is called a *fibre* bundle because locally E looks like the product $U \times F$. In physical terms, we can think of F as the generic representative of a *local* value space. The fibre $\pi^{-1}(x)$ above each point x then 'looks like' the global value space F. But note that there is not necessarily a canonical map from points of E to points of F: there exist distinct structure-preserving maps—*local trivialisations*—from $\pi^{-1}(x)$ into F.

Physical fields are represented by *sections* of fibre bundles:

Definition (Section)**.** *A* section *of a fibre bundle is a map* $s : M \to E$ *such that* $\pi(s(x)) = x$*.*

Put simply, a section assigns to each point x on the manifold a unique point p of the fibre above x. If the fibre over a point represents the possible field values at that point, then a section determines a field value at each point. Thus, sections replace functions $f : M \to \mathcal{V}_A$ from the manifold into some universal value space shared by all points.

We now come to discuss some more specific types of fibre bundles which are relevant in physics. The first of these is a *principal fibre bundle*:

Definition (Principal Fibre Bundle)**.** *A* principal fibre bundle (P, π, M) *is a fibre bundle whose typical fibre is homeomorphic to a Lie group* G*, for which there exists a smooth and* *free right action of G on P such that for any local trivialisation $\xi : U \times G \to \pi^{-1}(U)$, $\xi(p, g)g' = \xi(p, gg')$.*

Equivalently, a principal fibre bundle is a fibre bundle whose typical fibre is a group G, called the *structure group* of P. We require that any local trivialisation preserves the structure of the regular group action. An intuitive way to see this is that the action of G defines the 'difference' between points: if $p, q \in \pi^{-1}(x)$ and $q = pg$, then g is the difference between p and q. The requirement that local trivialisations preserve this structure means that if points on the typical fibre G are some 'distance' g away from each other, then so are the images of ξ at these points.

As with fibre bundles in general, there is no privileged map from the fibres of P onto G. This implies that for points of P on different fibres, it is indeterminate whether these points correspond to the *same* element of G. We can, however, endow a principal bundle with additional structure that defines a notion of 'sameness' across fibres. This is called a *connection*:

Definition (Connection)**.** *Let T_pP denote the tangent space at a point $p \in P$. The* vertical subspace *V_p of T_p is defined as $V_pP = \{\tau \in T_pP : \pi_* \tau = 0\}$, where π_* is the pushforward of π. A* connection *ω on P then assigns a* horizontal subspace *H_pP of T_pP to each point $p \in P$ such that (1) $T_pP = V_pP \oplus H_pP$ and (2) $R_{g*}(H_pP) = H_{pg}P$ (where R_g is the action of G on P and R_{g*} is the pushforward of R_g).*[12]

Effectively, a connection determines which of the vectors tangent to p count as horizontal, such that any vector in T_pP has a decomposition in terms of H_pP and V_pP (where the latter consists of all vectors that point 'along' the fibre). The connection is compatible with the action of G on P, so the horizontal subspace has the same orientation at each point on the fibre.

I will now connect some of these mathematical structures to the theory of scalar electrodynamics discussed above. First, we postulate that the connection on the

[12] Alternatively, one can define a connection algebraically as a Lie algebra-valued one-form that satisfies analogous conditions. These definitions are equivalent (Isham, 1999, 255).

principal bundle represents the electromagnetic potential. In the case of the theory discussed above, the structure group of the principal bundle is U(1). We can represent the connection ω as a vector field A_μ on M relative to a choice of section. This is the familiar vector potential, also called the Yang-Mills field. But A_μ depends on an arbitrary choice of section, whereas ω is a quantity intrinsic to the bundle structure. Therefore, I will focus on ω as the representative of the Yang-Mills field.

We do not yet have a representation of the matter field ϕ on which A_μ acts. These matter fields live on the *associated bundle*:

Definition (Associated Bundle)**.** *Define the* G-product $X \times_G Y$ *of two spaces* X *and* Y *on which* G *has a right action as the space that is obtained from the product space* $X \times Y$ *by identifying points* (x, y) *and* (x', y') *iff* $x' = gx$ *and* $y' = gy$ *for some* $g \in G$. *Let* $[x, y]$ *denote the equivalence classes obtained in this way.*

If P *is a principal* G*-bundle and* F *is a space with a right* G*-action, define* $P_F = P \times_G F$. *The* associated F-bundle *of a principal* G*-bundle then is a triple* (P_F, π_F, M) *where* $\pi_F([p, v]) = \pi(p)$. *If* F *is a vector-representation of* G*, then the associated bundle is a vector bundle.*

The field ϕ is represented by sections of the associated bundle. For example, the structure group of electrodynamics is U(1), and hence the associated bundle is a vector bundle whose typical fibre is isomorphic to $\mathbb{C}$. Locally, a section of this bundle is an assignment of an element of $\mathbb{C}$ to each point in M. But again, there is no canonical map from P_F onto $\mathbb{C}$, and hence a comparison of field-values across points depends on a conventional choice of local trivialisation.

However, the connection on P endows P_F with some further structure which defines a notion of *parallel translation.* First, define the *horizontal lift* of a curve on M to P:

Definition (Horizontal Lift to Principal Bundle)**.** *Let* $\gamma(t)$ *be a smooth curve on* M. *A curve* $\gamma^\uparrow(t)$ *on* P *is a* horizontal lift *of* $\gamma(t)$ *iff* $\pi(\gamma^\uparrow(t)) = \gamma(t)$ *and* $\gamma^\uparrow(t)$ *is horizontal, i.e.*

$\dot{\gamma}^{\uparrow}(t) \in H_pP$. *For each point* $p \in \pi^{-1}(\gamma(0))$, *there is a unique horizontal lift* $\gamma^{\uparrow}(t)$ *of* $\gamma(t)$ *such that* $\gamma^{\uparrow}(0) = p$.

This gives us the lift of a curve on M to the principal bundle P, but we are interested in parallel translation on the associated bundle P_F. We can use the previous definition to define horizontal lifts on P_F as follows:

Definition (Horizontal Lift to Associated Bundle)**.** *Recall that points on the associated bundle are equivalence classes* $[p, v]$. *Let* $k_v(p) = [p, v]$, *and let* $\gamma^{\uparrow}$ *be the unique horizontal lift of a curve* $\gamma(t)$ *on* M *to the principal bundle* P *such that* $\gamma^{\uparrow}(0) = p$. *Then, for any point* $[p, v] \in \pi_F^{-1}(\gamma(0))$, *the horizontal lift of* $\gamma(t)$ *that passes through* $[p, v]$ *is the curve* $\gamma_F^{\uparrow}(t) : k_v(\gamma^{\uparrow}(t)) = [\gamma^{\uparrow}(t), v]$.

We can then define parallel translation as a ternary relation $S[\gamma, p, q]$ between a path $\gamma(t)$ on M and a pair of points $p, q \in P_F$ such that $p \in \pi_F^{-1}(\gamma(a))$ and $q \in \pi_F^{-1}(\gamma(b))$ for some $a, b \in \gamma(t)$. Then q is the parallel translation of p along $\gamma(t)$ iff the horizontal lift $\gamma_F^{\uparrow}(t)$ such that $\gamma_F^{\uparrow}(a) = p$ passes through q, i.e. $\gamma_F^{\uparrow}(b) = q$. Intuitively this means that when one starts at p and travels along γ, one 'ends up' at q. But the relation of parallel translation is *path-dependent*: it is possible that p and q are connected via one path, but not via another. Therefore, parallel translation does not offer a well-defined *universal* notion of sameness across fibres.

In the next two sections, I will show how this picture facilitates a sophisticated account of electrodynamics and the Aharonov-Bohm effect.

7.4 Gauge Symmetries

According to sophistication, one can interpret symmetry-related models as physically equivalent *when those models are isomorphic.* The appropriate structure of a sophisticated theory is one whose dynamical symmetries are isomorphisms. In this section I argue that fibre bundle theories are appropriately structured in this sense.

Recall from Chapter 5 that we can extend Earman's (1989) symmetry principles to cover internal symmetries. The idea is to define *value space symmetries* as automorphisms of a value space $\mathcal{V}$, and *internal symmetries* as solutionhood-preserving transformations of a theory's models induced by bijections of $\mathcal{V}$. Let a model $\langle \mathcal{M}, \mathcal{V}, s\rangle$ consists of a spacetime $\mathcal{M}$, a value space $\mathcal{V}$ and a function $s : M \to \mathcal{V}$ which denotes some physical quantity. Furthermore, let $\mathcal{V} := \langle D, \mathbf{R}\rangle$, so the value space consists of a domain D and a set of relations $\mathbf{R}$ defined over D. In these terms, value space symmetries are automorphisms ϕ of D such that $\phi(R) = R$ for all $R \in \mathbf{R}$. Dynamical symmetries are transformations g of the form

$$\langle \mathcal{M}, \mathcal{V}, s\rangle \xrightarrow{g} \langle \mathcal{M}, \mathcal{V}, \phi_* s\rangle$$

such that $\mathfrak{m}$ is a solution iff $g(\mathfrak{m})$ is (here, ϕ^* is the pullback map induced by ϕ).

The generalisations of Earman's principles then state that:

SP1 If ϕ induces an internal symmetry, then it is a value space symmetry;

SP2 If ϕ is a value space symmetry, then it induces an internal symmetry.

The importance of these principles for our present purpose is that when they are satisfied, symmetry-related models are isomorphic. *Proof.* Consider some relation R on $\mathcal{V}$ and a pair of field values $s(x)$ and $s(y)$. Let ϕ induce a symmetry transformation. From SP1 it follows that $\phi(R) = R$, and hence $\langle s(x), s(y)\rangle \in R$ iff $\langle \phi_* s(x), \phi_* s(y)\rangle \in R$. Since this is the case for all relations $R \in \mathbf{R}$, the symmetry-related models are isomorphic. Conversely, suppose that some symmetry-related models $\langle \mathcal{M}, \mathcal{V}, s\rangle$ and $\langle \mathcal{M}, \mathcal{V}, \phi_* s\rangle$ are isomorphic, so there exists some ϕ such that $\langle s(x), s(y)\rangle \in R$ iff $\langle \phi_* s(x), \phi_* s(y)\rangle \in R$ for all $R \in \mathbf{R}$. But this just means that ϕ is a value space symmetry, so from SP2 it follows that ϕ induces an internal symmetry. Therefore, when a theory satisfies these principles it automatically satisfies the formal requirement for sophistication. $\square$

Here is how this applies to electrodynamics in the fibre bundle formalism. Instead of a universal value space $\mathcal{V}$, we now have a principal bundle P and an associated bundle P_F. The dynamical quantities are the section $s(x)$ which represents the matter field, and the connection ω which represents the Yang-Mills field. Therefore, our theory's models are of the form:[13]

$$\langle \mathcal{M}, P, P_F, s(x), \omega \rangle$$

The equivalent of SP1 says that for any model $\langle \mathcal{M}, P, P_F, s, \omega \rangle$, if the symmetry-related model $\langle \mathcal{M}, P, P_F, \phi_* s, \phi_* \omega \rangle$ is a solution then $\phi(P) = P$ and $\phi(P_F) = P_F$—that is, ϕ is an automorphism of the principal and the associated bundle. Conversely, SP2 says that if ϕ is an automorphism of P and P_F, then $\langle \mathcal{M}, P, P_F, s, \omega \rangle$ is a solution iff $\langle \mathcal{M}, P, P_F, \phi_* s, \phi_* \omega \rangle$ is a solution.

In order to check whether electrodynamics satisfies SP1 and SP2, we need to know the theory's symmetries. These are the transformations induced by so-called *vertical principal bundle automorphisms*:

Definition (Vertical Principal Bundle Automorphism)**.** *A principal bundle automorphism is a diffeomorphism $u : P \to P$ such that $u(pg) = u(p)g$ for all $p \in P$, $g \in G$. A principal bundle automorphism is* vertical *iff $\pi(u(p)) = \pi(p)$ for all $p \in P$.*

The vertical principal bundle automorphisms are dynamical symmetries of scalar electrodynamics. But vertical bundle automorphisms are, as the name suggests, *also* symmetries of $(P, \pi, \mathcal{M})$: such automorphisms preserve both the action of G on P and the map π from P to $\mathcal{M}$. Moreover, the same transformations are also symmetries of the associated bundle $(P_F, \pi_F, \mathcal{M})$. This is so because any principal bundle automorphism induces an associated bundle automorphism $u_F : [p, v] \to [u(p), v]$. Therefore, gauge transformations are indeed value space symmetries: SP1 and SP2

[13] Here and below I depart from the notation in §7.3 to let P denote the pair $\langle E, \pi \rangle$.

are satisfied for electrodynamics formulated in terms of fibre bundles.

We can see this when we consider the action of gauge symmetries on the physical quantities s and ω. The transformation of s is straightforward: if $s(x) = [p, v]$, then $u_*s(x) = [u(p), v]$. The connection, meanwhile, transforms via the *pull-back* of u to $u^*\omega$. By definition, u preserves all 'vertical' structure of both s and ω. This leaves us with the 'horizontal' structure, which is captured by their behaviour under parallel translation. Consider a path γ and values of s at points $x, y \in \gamma$. The question then is: is it the case that $S[\gamma, s(x), s(y)]$ iff $S[\gamma, u_*s(x), u_*s(y)]$? The answer is 'Yes'. This is easiest to see algebraically: because ω transforms via the pull-back of u, it is the case that $\omega(s(x)) = u^*\omega(u_*s(x))$, hence the unique horizontal lift of γ through $s(x)$ is mapped onto the horizontal lift of γ through $u_*s(x)$. Therefore, the relation of parallel translation is preserved by gauge transformations. Since gauge symmetries preserve both vertical and horizontal structure, it follows that gauge-related models are indeed isomorphic.

7.5 The Metaphysics of Fibre Bundles

The fact that gauge theories have non-trivial local symmetries is usually seen as a defect. Consider, for example, the matter field as represented by a section $s(x)$ of the associated bundle. Generally, $s(x) \neq u_*s(x)$, where u is a vertical bundle automorphism. However, since these fields are symmetry-related, there are no experiments which can measure the difference between these sections. Thus, the presence of symmetries seems to entail underdetermination. Since the symmetries of gauge theories are local, this also seems to imply that such theories are indeterministic in a way analogous to the alleged indeterminism that features in the Hole Argument.

But these arguments are based on the implicit assumption that $s(x)$ and $u_*s(x)$ represent *distinct* fields. This follows from a literal reading of the fibre bundle formalism: if every point $p \in P_F$ represents a field value, then an assignment of different elements of P_F to points of M must represent a distinct field. However, this claim rests on the assumption that any point p represents the *same* field strength across models. In Chapter 2, I dubbed this assumption the Value-Magnitude Link. But sophistication rejects this link: it claims that *when symmetry-related models are isomorphic*, we can interpret them as physically equivalent. In the case of field theories, this requires an appeal to *anti-quidditism*: the thesis that field values are *qualitatively identified*. In brief, this means that *which* field value is instantiated at some point x depends on the structure of the field over the totality of spacetime. I will elaborate on these claims below.

In this section, I discuss the matter field and the Yang-Mills field in turn. I will argue that both fields are amenable to a sophisticated treatment. This will yield the required result, namely that gauge-related models of classical field theories are physically equivalent. In this way, sophistication avoids underdetermination and indeterminism.

7.5.1 Matter Fields

The analysis of matter fields consists of two steps. The first step is to reify the associated bundle; the second is to identify field values with structural positions within this bundle. I will discuss each step in turn.

Firstly, I propose *associated bundle Platonism*: elements of the associated bundle P_F represents physically real entities, namely the values of the relevant matter field. In the case of classical electrodynamics, these are the values of the matter field ϕ. The sections of the associated bundle then represent physical fields. Thus, spacetime points instantiate field values in the same way that particles instantiate masses, except that in the fibre bundle formalism each spacetime point carries its own set of field values. This implies that field values at different points are numerically distinct, and so different points simply *cannot* possess the same field value. The structure of the associated bundle represents the relations between field values, both vertical and horizontal. For instance, the ternary relation of parallel translation is part of the horizontal structure. There is a physical fact of the matter as to whether the field values instantiated at distinct points are related to each other by parallel translation via some path γ.[14]

The above claims are meant to be taken literally: field values *exist* and stand in second-order relations. We can contrast this with the views of Maudlin (2007) and Arntzenius (2012). For Maudlin, field values are neither universals nor tropes; he rejects the existence of field properties entirely. Maudlin argues that properties ought to induce a notion of *similarity* between objects. For example, tomatoes and strawberries are similar in that both are red. But since the fibre bundle picture implies that the field values at each point are *sui generis*, there simply is no (path-independent) notion of similarity for gauge fields. Maudlin concludes that the category of properties is superseded by that of fibre bundles. The view I call associated bundle Platonism is dismissed in a footnote: "One could suggest that there are still color properties, but that every point in the spacetime has its own set of properties, which cannot be instantiated at any other point. But since such point-confined properties could not underwrite any notion of similarity or dissimilarity [...], it is hard to see what would be gained by adopting the locution." (Maudlin, 2007, fn. 9)

It seems to me that Maudlin's focuses too much on *identity*. True, if distinct objects possess the very same property, this is one way in which they are similar. But consider two objects, one of which is 1 kg and the other is 2 kg. These objects do not possess the same mass value, yet they are similar in that both objects are massive (in other words, both objects instantiate different *determinates* of the same *determinable*). Furthermore, it is clear that both objects are more similar to each

[14]In this way, the present view resembles the one found in Leeds (1999).

other in this respect than to some third object whose mass is 100 kg. The latter claim follows from the fact that mass value space has a certain structure for which mass ratios are well-defined: the closer a mass ratio is to 1, the more similar the objects are with respect to their masses. But just like mass value space, the associated bundle which represents the matter field also possesses a highly non-trivial structure—for instance, the relation of parallel transport defined above. The fibre bundle picture does not allow us to say whether distinct points possess the *same* field value, but this does not mean that we cannot say anything of interest about the field properties at distinct points. Maudlin's objection to fibre bundle Platonism is unsuccessful.

Arntzenius (2012) presents a different interpretation of fibre bundles, which Wolff (2020) calls *locationism*. The core idea of this view is that field values are not Platonic universals, but are of the same metaphysical kind as spacetime points. In addition to their usual spatiotemporal location, particles also have a location in 'field value space'. The advantage of this view, or so Arntzenius argues, is that it satisfies Occam's razor insofar as it posits fewer *kinds* of entities. But the problem with fibre bundle locationism is that it utterly fails in the context of field theories. In modern field theories, spacetime points themselves possess field values. Formally, we can see this because the field distribution is represented by a section, i.e. a function from the manifold into the bundle. On Arntzenius' view, the section determines the location of each spacetime point within the fibre above that point. While we can make sense of the idea that some discrete object has a location both in spacetime and in some distinct value space, it does not make much sense to say that spacetime points themselves are located in a value space: spacetime points *are* locations. I therefore reject Arntzenius' locationism as an alternative to Platonism. This concludes the first step.

The second step is to stipulate certain cross-world identity conditions for field values. In particular, we stipulate that field values are *qualitatively individuated*:

they are nodes in a web of relations.[15] Dewar formulates the claim as follows: "We should be anti-quidditists, and deny that physical properties are modally robust. We should not believe that there are worlds that instantiate the same structure in their laws, and differ only over which properties play which nomological roles" (Dewar, 2019, 505). In more detail, anti-quidditism entails that field values which satisfy the same structural patterns are in fact identical. The relevant structural patterns here are just those codified in the vertical and horizontal structure of the fibre bundle. The fact that gauge-related models are isomorphic means that gauge transformations of field values preserve their structural patterns. From anti-quidditism, it follows that a gauge transformation does not affect which field values are instantiated at any spacetime point.

For a concrete example, consider a pair of gauge-related sections s and s'. On a literalist view, these sections represent distinct field configurations. But from the discussion of §7.4 it follows that s and s' are structurally equivalent. For any point $x \in M$, the values $s(x)$ and $s'(x)$ are qualitatively identical. We can illustrate this with the relation of parallel translation. The relations of parallel translation between field values at distinct points are part of the relevant structural patterns. Since gauge transformations preserve parallel translation, if $s(x)$ and $s(y)$ are related by parallel translation via some path γ then so are $s'(x)$ and $s'(y)$. Therefore, $s(x)$ and $s'(x)$ merely represent the same field value, namely one with a certain qualitative profile.

Consequently, gauge-related models represent the same physical possibility. On a sophisticated interpretation the theory is neither underdetermined nor indeterministic.

[15]The suggestion to apply the metaphysics of anti-quidditism to fibre bundles originates from Dewar (2019); see also Martens and Read (2020). The move here is analogous to the adoption of anti-haecceitism in response to the Hole Argument. For more on the latter, see Pooley (2006) and references therein.

7.5.2 Yang-Mills Fields

Since the Yang-Mills field is represented by a connection on the principal bundle it is tempting to adopt *principal bundle Platonism* in analogy with associated bundle Platonism. But there is an important disanalogy between the matter field and the Yang-Mills field: while the former is represented by a *section* of the associated bundle, the latter is represented by a *connection* on the principal bundle. While we can easily interpret a section as an assignment of field values to spacetime points, the same is not the case for the connection. The connection specifies relations *between* points of the principal bundle—but what do the points of the principal bundle themselves represent?[16]

I will discuss two possible answer to this question. The first is a *deflationary approach*: neither the principal bundle nor the connection on it represent anything physical. Rather, it is the induced connection on the associated bundle that represents the Yang-Mills field. This approach has difficulties in accounting for distinct matter fields coupled to the same Yang-Mills field. The *inflationary approach*, on the other hand, reifies not the principal bundle but the bundle tangent to it. The Yang-Mills field is then represented by a section of this 'bundle of connections'. The inflationary approach is preferable because it can explain the way in which distinct matter fields couple to the same Yang-Mills field.

Deflationary Bundle Realism

Recall from §7.3 that the connection on the principal bundle P induces a connection on the associated bundle P_F. According to the deflationary approach the latter represents the Yang-Mills field, which is therefore not really a *field*. Rather, the connection specifies relations *between* field values. The Yang-Mills field is thus similar to velocity, in the sense that velocities also supervene on relations between nearby

[16] Dewar (2019, fn. 42) also mentions this conundrum.

points of a curve. Wallace (2014a) points out that this yields an essentially dualistic ontology of local field values on the one hand and infinitesimal field relations on the other.

On this picture, the principal bundle is a mathematical abstraction. Weatherall (2016a) defends this view. Weatherall notes that just as it is possible to define an associated bundle from a principal bundle, so one can do the reverse. Furthermore, the principal bundle is mathematically similar to the bundle of frames in General Relativity, whose role is to coordinatise spacetime. On this view, it is only when we consider more than one field that the principal bundle becomes relevant. For if distinct matter fields couple to the same Yang-Mills field, it is useful to represent the latter 'by itself' on a principal bundle. The claim that both matter fields couple to the same Yang-Mills field then translates into the fact that both vector bundles are associated to the same principal bundle.[17]

But it is a problem for this approach that the two fields survey the same connection as a matter of brute fact.[18] There really are two connections: one defined over the first associated bundle, and one defined over the second. These connections are the same only in the sense that we can represent both with the same connection on a single principal bundle. But there is no independently existing Yang-Mills field that the associated bundle connections supervene on. This makes it seem somewhat mysterious that these connections are equivalent. The mysterious coordination between associated bundles begs for a 'common cause' in the form of an independently existing Yang-Mills field.[19] I will therefore consider an alternative view which can explain the

[17]The reason that Wallace's view struggled with such theories (§7.2.3) is now clear: to make sense of two matter fields that couple to the same gauge field, one represents the latter as a connection on distinct associated bundles of each matter field. But on Wallace's view, at least one of the two matter fields lives on a universal real vector space, rather than a fibre bundle.

[18]For a criticism of the technical aspects of Weatherall's deflationary view, see Menon (2018).

[19]There is a similarity here to the debate between the dynamical and geometrical approach to spacetime. On the former approach, it is a brute fact that all matter fields have the same symmetries. But if we assume that matter surveys the structure of spacetime, then the latter can explain the coincidence of symmetries. I thank James Read for this observation.

coincidence of distinct connections in terms of a physically real Yang-Mills field which interacts with both matter fields.

Inflationary Bundle Realism

Recall that a connection defines a horizontal subspace at each point $p \in P$. The function of the connection is to determine which direction counts as horizontal. This strongly suggest that the 'values' of the Yang-Mills field are not elements of P, but of the tangent bundle TP. Specifically, at each point p the 'value' of the Yang-Mills field consists of a subspace H_pP of T_pP which counts as horizontal.

Yet this picture is still redundant, in two ways. Firstly, the connection assigns an entire subspace of T_pP to points of M, rather than a single element of T_pP. In this sense, the connection does not yield a unique value of the Yang-Mills field. Secondly, the connection assigns a subspace of T_pP *for each point* $p \in \pi^{-1}(x)$. This overdescribes the Yang-Mills field, since the connection at one point on a fibre uniquely determines the connection elsewhere on the same fibre via the condition that the connection is compatible with the action of G on P (i.e. $R^{g*}(H_pP) = H_{pg}P$).

We can remove this excess structure via a construction called the *bundle of connections*:[20]

Definition (Bundle of Connections)**.** *Let (P, π, M) be a principal bundle with structure group G; TP is its tangent bundle. Let TP/G denote the quotient of TP by G. Then $(TP/G, d\pi, TM)$ is a fibre bundle over TM (the tangent bundle of the spacetime manifold M), called the* bundle of connections*. There is a one-to-one correspondence between connections on P and linear sections Γ of the bundle of connections.*

The bundle of connections involves two innovations. Firstly, the bundle projects onto the tangent space TM of the manifold M. Instead of a horizontal subspace for each point, the bundle of connections assigns to each vector $v \in TM$ a subspace of T_pP,

[20] For more on this construction, see Kobayaschi (1957, Ch. 4).

such that each vector in this subspace points in the same direction as (the lift of) v. This overcomes the first redundancy. Secondly, in the definition of the bundle of connections we have quotiented by the action of G, such that each vector in TM is assigned a single copy of the tangent bundle T_p at some point p. This overcomes the second redundancy. The result is a construction which assigns a vector space v_pP to each element of TM, i.e. to each pair of infinitesimally close spacetime points. The Yang-Mills field is then represented by a section of the bundle of connections, which assigns to each tangent vector of the spacetime manifold a unique horizontal vector of the principal bundle in the same direction. This vindicates Wallace's claim that the ontology of gauge theories is dualistic: matter fields inhere in points, whereas gauge fields are better understood as relations between points. The advantage of the inflationary view is that the Yang-Mills field is an independently existing entity which we can use to explain the fact that distinct matter fields behave in the same way under parallel translation. The explanation is simply that both matter fields survey the same connection, here understood as a section of the bundle of connections.

Finally, we can apply anti-quidditism to the Yang-Mills field, just as we did for the matter field. In this case, a gauge transformation acts directly on the section of the bundle of connections. But once more, such a transformations preserves all structural patterns. Therefore, the image of ω after a gauge transformation is really the same connection differently represented on the bundle of connections. The conclusion is the same: gauge-related models represent the same physical possibility.

7.6 Locality, Separability, Holism

The sophisticated interpretation of the fibre bundle formalism faces no underdetermination. But an equally puzzling feature of the Aharonov-Bohm effect is that it seems to imply that physics is, in some sense, non-local. Is this also the case for fibre bundle realism? In this section, I will argue that fibre bundle realism is both local and separable. But there is another—less problematic—sense in which the fibre bundle account is *holistic*, which explains the non-local nature of the Aharonov-Bohm effect.

7.6.1 Local Action

In order to see whether fibre bundle realism violates local action, it is useful to return to the Aharonov-Bohm effect. According to literal-minded A-field realism, there is a four-potential field that acts locally on the matter field, shifting its phase as the field propagates across regions within which A_μ is non-trivial. But this account assumes that field values are comparable across points, such that A_μ has an unequivocal influence on ϕ. This is no longer true on a fibre bundle account, since each spacetime point now has its own set of field values. Then how does one model *interaction* on the fibre account?

The idea here is that the connection determines the evolution of the matter field. Locally, the connection defines a notion of 'sameness' across fibres: if q is the parallel translation of p over a path γ, then p and q lie in the same horizontal plane. The dynamics of classical field theories tell us that matter fields propagate along this horizontal direction: the matter field 'surveys' the connection at infinitesimal distances. The account is thus fully local: it is only the connection at the location of the matter field which determines the evolution on the fibre bundle over time.

In the Aharonov-Bohm effect, the connection is such that the parallel translation of the field along a path through the left slit and one along a path through the right shift differ. When both halves of the matter field 'meet' at Q, they find themselves at different points on the fibre above Q: this is just the phase difference that causes the shift in the interference pattern. The degree to which parallel translation around closed curves is closed is called the *curvature*, characterised by $F_{\mu\nu}$. This is indeed a global feature of the bundle, but on our account it supervenes on local values of the connection.

7.6.2 Separability

Is fibre bundle realism also *separable*? Dewar (2019, fn. 56) alleges that the connection on the principal bundle is not, and Martens and Read (2020) concur; a similar claim is also found in Lyre (2004). On the contrary, I believe that the connection *is* separable. Of course, the connection connects infinitesimally close points, so in that sense it is not completely local. But such violations of separability do not seem particularly worrisome, and this is clearly not the sense of non-separability that Dewar intends. There is little reason to believe that even classical theories are truly local in this sense, as Butterfield (2006) has argued. Instead consider *regions* rather than points.[21] Specifically, consider distinct regions U and V which partially overlap. The connection on U determines the horizontal lift of all paths on U, and the connection on V determines the horizontal lifts of all paths on V. But since paths on $U \cup V$ are composed of paths on U and paths on V, their horizontal lifts are now fully determined as well. It seems then that the fibre bundle picture is separable.

In personal correspondence, Dewar has clarified his claim to the contrary as follows. Consider regions U and V which individually do not surround the solenoid, but whose union $U \cup V$ does. The connections on U and V are *locally* isomorphic to a connection which is zero everywhere within these regions (in some local representation), even though the connection on $U \cup V$ is not. Therefore, the connections on U and V *up to gauge transformations* do not determine the connection on $U \cup V$. This is the sense in which Dewar claims connections are non-separable. But this shouldn't come as a surprise, since the gauge-invariant content of the bundle over a region U just consists of the holonomies, which we already saw are non-separable. Fibre bundle realism can resist this conclusion.[22]

[21] So, we are interested in what Myrvold (2011) calls 'patchy' separability.

[22] The same point is made by Dougherty (2017). Dougherty shows that only a 'truncated' version

The problem with Dewar's argument is that it assumes that sophistication considers symmetry-related states of *subsystems* of the universe as equivalent. But sophistication (as I have presented it) is only interested in universe states. After all, the issue of indeterminism only arises when we consider universe symmetries. In the case of gauge theories, this is obscured by the fact that subsystem symmetries which vanish to the identity *are* universe symmetries. But a discussion of sophistication in the context of classical mechanics clarifies this issue. Consider a pair of (dynamically isolated) classical systems; for instance, Rovelli's (2014) fleets of spaceships. Both of these systems are invariant under static shifts, hence the symmetry-invariant content of each system consists only of distance relations. However, the distances between ships of the first fleet and the distances between ships of the second fleet fail to fully determine the joint state of both fleets. After all, the latter also includes the distance between the fleets themselves, which is not contained in the invariant facts about either system individually. From this reasoning, it would seem that even classical mechanics is non-separable—a conclusion which Dewar *et al.* are unlikely to defend. But the mistake in this argument is that sophistication is *not* committed to a symmetry-invariant ontology of distance relations. On the contrary, sophistication embraces realism about non-invariant quantities such as positions on a manifold. In the same vein, Gomes (2019) argues that 'forgetting' symmetry-variant structure of subsystems is fine when we consider such systems in themselves, but causes trouble when we consider the 'gluing' of subsystems. Instead, Gomes advocates 'external sophistication and internal reduction': in order to glue subsystems together, we need to pay attention to their symmetry-variant features.[23]

of the fibre bundle formalism leads to non-separability. I conjecture that Lyre and Dewar have this truncated version of the formalism in mind. But an untruncated version of the same formalism does not entail non-separability. It is this version which I adopt here. Dougherty notes that an untruncated version of the holonomy approach is also separable. But as I said in §7.2.2, the alleged non-separability of holonomies is in any case not the strongest argument against holonomy realism.

[23]This is not to deny that isomorphic models of subsystems may have the same representational *capacities*; but this simply does not imply that such models are, for all purposes, physically equivalent. On this point, see Roberts (2020).

7.6.3 Holism

Yet it remains the case that the Aharonov-Bohm effect *is* distinctly non-local in character. This is clearly brought out when we ask *where* the electron picks up a phase: where does the Aharonov-Bohm effect *come about*? As Healey points out, for any local representation of the connection and any region U which does not enclose the solenoid, there exists a gauge such that A_μ vanishes in U. Therefore, Healey concludes, there is no region in which the matter field non-trivially interacts with the Yang-Mills field—and hence no interaction at all! But Healey's argument is invalid. For while it is true that for any region there exists a gauge in which A_μ vanishes, there is no gauge such that A_μ vanishes everywhere. Therefore, it is a gauge-invariant fact that the (local representative of the) connection is non-trivial *somewhere*. Nevertheless, Healey's argument is instructive in highlighting the *holistic* nature of the Aharonov-Bohm effect.

There is no meaningful answer to the question *where* the matter field picks up a phase because there are no meaningful cross-point comparisons of phase values, since each point of the manifold is associated with its own fibre of field values. In particular, there is no canonical map from local fibres into $U(1)$. There simply *is* no picking up of a phase. The matter field simply moves across the fibre as dictated by the connection. We can only compare field values at the same point, that is, when the two halves of the matter field meet at the point Q on the screen.The holism of the Aharonov-Bohm effect consists of the fact that there is not enough physical structure to express phase differences along open paths, but only around closed paths.

The following analogy is helpful.[24] Consider the Twin Paradox: one twin remains in her inertial rest frame on earth, while the other travels to Alpha Centauri and back. Since the latter measures less proper time, she is younger on return than her sister. Just as we are interested in where the phase difference occurs in the Aharonov-Bohm

[24]For another helpful analogy (with currency exchange rates), see Maldacena (2014).

effect, we may wonder when, and at which rate, the age difference between the twins comes into existence. This is easy enough to answer with respect to certain planes of simultaneity. For example, from the perspective of the earth-twin, the rocket-twin's clock runs slow at a constant rate. But it is often[25] thought that simultaneity is *conventional*. If that is indeed the case, different simultaneity choices yield different results for the differential ageing of the twins (Debs and Redhead, 1996). For example, one can choose a convention such that the rocket-twin does not age slowly compared to the earth-twin at all on the outbound leg of her trip, but then ages even more slowly on her return leg. This result is analogous to the fact that one can choose a gauge such that A_μ is zero over any open path. Indeed, Rynasiewicz (2012) has shown that choice of simultaneity convention is formally equivalent to a choice of 'gauge' in General Relativity. Just as there are no cross-point comparisons of field values, so there are no objective cross-point simultaneity relations.

This implies that effects such as the Twin Paradox and the Aharonov-Bohm effect are holistic in this sense: although the total effect size (the age difference or interference shift) is measurable, there is no fact of the matter as to how this effect comes about as the result of small local differences. The final age difference between the twins is not the result of many small age differences that accrue locally. Likewise, the final phase difference in the Aharonov-Bohm effect is not the sum of the phase differences over infinitesimal paths. Although puzzling, such holism is simply a consequence of the fact that value spaces are localised. This is not a defect in our theories, but a consequence of the novel metaphysics of fibre bundles.

7.7 Conclusion

I have defended a sophisticated interpretation of gauge theories in the fibre bundle formalism. The interpretation is local, separable and deterministic. I have explained

[25] But not universally—see Malament (1977). I will not enter this debate here.

the puzzling nature of the Aharonov-Bohm effect in terms of a certain form of holism which is a consequence of the fact that each point comes attached with its own set of field values. The account easily generalises to non-Abelian gauge theories, since nothing in the above depends on the fact that the $U(1)$ structure group of electrodynamics is Abelian. This makes sophistication preferable over Wallace's deflationary account.

There is also a broader lesson here, namely that symmetries are an important guide to the structure of physical quantities. This is especially clear in the fibre bundle framework, since it is essentially the local $U(1)$ symmetry-group of electrodynamics which determines the structure of the principal bundle, and hence of the associated bundle which represents matter fields. Far from a redundancy in our description of the world, symmetries are carriers of physical information. Instead of eradicating this structure from our theories, sophistication does justice to its significance.

Chapter 8

Conclusion: Why Symmetries?

Symmetries are of great significance in the philosophy of physics. But *what* their significance is is a matter of profound disagreement. The orthodox view is expressed by Ismael and van Fraassen's (2003) slogan that symmetries are a "guide to superfluous structure". On this view, there exists a symmetry-invariant reality that is obscured by 'descriptive fluff' in our models. The function of symmetries is to reveal this superfluous structure. This claim dovetails with *reduction*, the view that we ought to construct novel theories which are expressed solely in terms of the theory's invariant quantities. The hope is that such a reduced theory is free from any superfluous structure. *Sophistication*, the alternative to reduction that I have advocated, challenges this orthodoxy: on a 'sophisticated' view, symmetries are part of a *positive* account of the structure of reality. We can draw out the difference between these two views by asking *why* our theories have symmetries. Why is it the case that virtually every physical theory of the past five hundred years, from Newtonian mechanics to the standard model of particle physics, exhibits non-trivial symmetries? This is the question Rovelli (2014) asks in his paper 'Why gauge?'.[1] In these concluding remarks, I aim to answer Rovelli's question from the perspective of sophistication.

[1]Rovelli uses the term 'gauge' in the philosophical sense to mean any symmetry that relates physically equivalent states. This definition includes, but is not restricted to, the local (internal) symmetries for which I have reserved the term.

Of course, there is another explanation of this fact that is both obvious and wholly unsatisfactory. This is the view I have called *literalism*: symmetry-related models represent physically *distinct* states of the world. The explanation of the fact that there are symmetries then is just that the dynamically relevant facts underdetermine the physical facts. (Of course, that claim may require its own explanation in turn.) However, it is exactly because of the resulting underdetermination that literalism is unsatisfactory: as I discussed in Chapter 2, such underdetermination leads to unmeasurable quantities; a failure of theoretical terms to uniquely refer; and in the case of local symmetries, indeterminism. Therefore, literalism is almost universally rejected. Instead, it is asserted that symmetry-relate models are physically equivalent: a thesis known as Leibniz Equivalence. It is only once we reject literalism that Rovelli's question really arises: if symmetry-related models *are* physically equivalent, then why does their existence seem indispensible to physics? Why are our theories such that they allow us to represent the same physical possibility in many different ways?

Before sketching my answer, let me set out the orthodox view (reduction) in more detail. According to this view, the quantities that appear in our theories—spatial location, intrinsic mass, gauge fields—are not fundamental. Instead, there are some underlying *invariant* quantities that are fundamental. These fundamental quantities are redundantly represented by the variant quantities that actually appear in our theories. For example, some comparativists claim that fundamental mass relations are represented as ratios of monadic mass quantities: the fundamental mass relation $r(i, j)$ between particles is represented as the ratio $m(i)/m(j)$, where $m(i)$ and $m(j)$ express monadic properties of massive bodies. But this representation is redundant, because distinct pairs of masses can yield the same mass ratio. Similarly, on the orthodox view the loop holonomies $H(l)$ we encountered in Chapter 7 are the fundamental quantities of scalar electrodynamics, expressed as integrals of the four-potential field A_a. Again, different distributions of the A_a-field can lead to the exact same values for the holonomies. The orthodox view embraces the Invariance Principle, which states that only symmetry-invariant quantities such as $r(i,j)$ or $H(l)$ are physically real.

The orthodoxy explains the existence of symmetries as follows. On the one hand, the fundamental quantities are invariant, so there are no dynamical symmetries at the physical level; that is, there are no systematic transformations of the invariant quantities that yield physically distinct yet empirically equivalent states of affairs. But the invariant quantities are conveniently expressed in terms of non-invariant ones, and it is this expression that is one-to-many. For example, there are different pairs of intrinsic masses $m(i)$ and $m(j)$ that correspond to the same mass ratio $r(i,j)$. The reason that our theories possess symmetries, then, is that those theories are not expressed in terms of the fundamental quantities. This infelicity introduces descriptive redundancies. In a sense, a world with intrinsic masses or localised gauge fields is richer than a purely invariant world, since one can embed many different states of the former type into a unique state of the latter. The function of symmetries, on this view, is essentially negative. They are, in Ismael and van Fraassen's words, a "guide to superfluous theoretical structure". What this means is that symmetries act as a *decision criterion* for physical reality. If some piece of structure is symmetry-invariant, then it is physically real; if it is not, then it is superfluous. In the same vein, Dasgupta (2016) coins the term 'symmetry-to-reality inference'. The idea is that we can infer the (un)reality of some quantity from its (non-)invariance under dynamical symmetries. Finally, the fact that symmetries reveal superfluous structure provides the motivation for finding novel, reduced theories without those symmetries.

The orthodox view faces a number of objections. These come in a Quinean and a Humean flavour. The Quinean objection is that our best theories do not quantify over the invariant quantities that the orthodoxy sees as fundamental, but over the variant quantities that it deems redundant. According to Quine (1948), we are ontologically committed to whatever entities are quantified over in our best systematisation of the empirical facts, but this claim is in tension with the ideas of symmetries as superfluous. After all, our best physical theories *are* expressed in terms of spacetime points, intrinsic masses and local fields, rather than distances, mass ratios and holonomies. Many reductionists realise the importance of finding laws formulated in terms of the fundamental quantities, but progress in this area has been limited. For instance, Earman's (1989) programme of Einstein algebras turned out to be subject to the same form of underdetermination as manifold substantivalism (Rynasiewicz, 1992). This is an embarrassment for the orthodoxy if we accept the view that the laws that govern the behaviour of fundamental quantities ought themselves be expressed in terms of those quantities. In a similar vein, Arntzenius (2012) distinguishes the 'easy' from the 'hard' nominalist project: the former consists in finding a nominalist *ontology* for physics, the latter requires an expression of the laws in terms of this novel ontology. If we replace 'nominalist' with 'invariant', we can say that reductionism has at best succeeded at the easy project, but not yet at the hard one.

I think the Quinean objection is both strong and unanswered,[2] but in this thesis I have emphasised the Humean objection. According to Hume's dictum, there are no necessary connections between distinct existents. This is often interpreted as the claim that the fundamental stuff of the world is *freely recombinable*. But, as I argued in Chapter 3, invariant quantities are often not freely recombinable. This means that they are a poor candidate for fundamentality. This is the case when the invariant quantities are *relational*. For example, mass relations are typically transitive: if I am equally massive as Pascal, and Pascal is equally massive as Eve, then Eve is—must be—equally massive as I am. The mass relations between Pascal and me and between Eve and me thus constrain the mass relation between Pascal and Eve. Of course, the

[2]The most plausible answer I know of is Melia's (2000) idea of 'weaselling away' structure, but this essentially amounts to a rejection of the Quinean view.

relationist could argue that constraints such as the transitivity of mass ratios are further laws of physics. But this both makes their theory more complicated and introduces a form of non-local influence, since it requires that mass relations across space are calibrated in just the right way. On the other hand, the transitivity of mass ratios (and its analogue for other relational quantities) is easily explained if we assume that mass ratios in fact supervene on intrinsic masses, for then the mathematical theorem $\frac{m(i)}{m(k)} = \frac{m(i)}{m(j)} \cdot \frac{m(j)}{m(k)}$ is implied by the additive structure of intrinsic mass. I have defended this line of reasoning from objections in Chapter 3.

To sum up, the orthodox view claims that all fundamental quantities are invariant under the theory's symmetries. This entails that symmetries relate physically equivalent states. The only reason that our theories possess symmetries is that the invariant quantities are redundantly represented by variant ones. We can then use symmetries as the basis for 'symmetry-to-(un)reality' inference. But the claim that only invariant quantities are fundamental faces a pair of critical issues: our theories are not expressed in terms of these quantities, and even if they were, such quantities fail the condition of free recombination. Both are reasons to suspect that invariant quantities are not fundamental. If invariant quantities are not fundamental, then the orthodox explanation of symmetries fails.

Rovelli himself is often seen as propounding the orthodox view (for instance by Teh (2015)). Rovelli (2014) asks us to imagine a fleet of N spaceships whose dynamics depend solely on the ships' velocities. We can write down a set of variables x_n for the ships' positions, but because of the invariance of the dynamics under static shifts the invariant quantities are the relative positions $a_n = x_{n+1} - x_n$. We now imagine a *second* fleet of spaceships with position variables y_n and (invariant) distances $b_n = y_{n+1} - y_n$. It seems that these quantities fully characterise each fleet. If we now apply a static shift to the second fleet only, the difference is clearly observable: the distance between the two fleets has increased. But this difference is not registered by the invariant observables a_n and b_n of each fleet, hence we need to introduce an additional variable: $c = y_1 - x_N$. Tellingly, c is defined in terms of a pair of gauge-*dependent* variables. The lesson relationists draw from this is that gauge-dependent variables are useful to represent the interaction between subsystems. Gomes (2019) calls this position 'external sophistication and internal reduction': variant quantities are convenient to describe a system from the outside (i.e. to describe the first fleet from the perspective of the second). Quantities that are not physical when considered from within a system *become* physical once that system is considered as a subsystem of the universe. Nevertheless, this view holds that from an intrinsic point of view the invariant quantities remain fundamental: c is physically real, but y_1 and x_N are not.

I believe that Rovelli himself rejects this view. Although Rovelli agrees that it is *possible* to describe the composite system in terms of gauge-invariant distances, he argues that "restricting to gauge-invariant observables makes us blind to a fine structure of the world" (103). I thus disagree with Teh's (2015) reading of Rovelli as a relationist. In support of my interpretation, note that Rovelli claims that "[g]auge quantities cannot be predicted, but can often be *measured*" (101, emphasis mine). Furthermore, while Rovelli writes that "nature is described by relative quantities" (101), he defines relative quantities as relations between variant observables, such as c defined above. Far from anti-realism about such observables, Rovelli believes that physics is all about the relations *between* them. Instead, I submit that Rovelli's views lead us to an alternative answer to the question 'Why symmetries?', which is closer to the spirit of sophistication. Consider how Rovelli describes the essence of the spaceship scenario:

> The position of the individual ships is redundant in the theory insofar as we measure only relative distances among these. *But in the physical world each ship has a position nevertheless*: this becomes meaningful with respect to one additional ship, as this appears. In other words, the existence of a gauge expresses the fact that a reference is needed to measure

> the position of a ship. It expresses the fact that we measure position relationally. (Rovelli 2014, 95, emphasis mine)

Rovelli states that each ship *has a position*; he thus rejects the view that distances are fundamental. According to Rovelli, symmetries arise from the fact that we cannot measure the position of a ship without relating its position to that of the other ships. Rovelli puts this in terms of measurement, but we can rephrase the point in terms of descriptions: *we cannot absolutely individuate positions without reference to the objects which occupy them.*[3] In other words, there are no features of spacetime which allow us to answer the question: where is this object located? Of course, we can say that the object is located at the unique point at which it is located, but such an answer is clearly trivial. For instance, there is no point that answers to the description: *the (unique) point at the centre of the universe.* The same is the case for quantities such as mass. There are no definite descriptions of mass properties, such as *the smallest mass* or *the unique unit mass.* It is in this sense, and only in this sense, that the world is fundamentally relational.

This provides us with the seed of a better explanation of symmetries. The key Rovellian insight is that there are no characteristics that absolutely individuate the values of physical quantities, without reference to where and when these values are instantiated. We now combine this insight with the metaphysics of sophistication. Recall from Chapter 1 that sophistication is committed to all and only the invariant content of a theory. In particular, sophistication embraces *anti-haecceitism*: the thesis that there are no non-qualitative fundamental facts. When possibilities are qualitatively identical, they are one and the same. The aim of sophistication is to restructure the theory such that symmetries preserve all qualitative features. In formal terms, this is the requirement that the theory's models are isomorphic. If symmetries

[3]In more precise terms: there exists no one-place formula which quantifies at most over the set of spacetime points, which applies uniquely to some particular spacetime point. For more on absolute (in)discernibility in relation to symmetries, see Caulton and Butterfield (2012).

preserve all qualitative structure, then it follows from anti-haecceitism that symmetry-related models represent the same state of affairs. In this way, sophistication avoids symmetry-induced underdetermination. On the other hand, sophistication affirms the reality of variant quantities. Therefore, it does not fall prey to the Quinean nor to the Humean objections discussed above.

This returns us to the initial question: if symmetry-related models merely represent the same state of affairs, why symmetries in the first place? The final answer has to do with how we use mathematical models to represent physical possibilities. Recall that a model consists of various structured domains: a domain of particles, a domain of spacetime points, a domain of mass values. Each structured domain $\mathcal{D}_i$ consists of a base set D_i and a set of relations $\mathbf{R}_i$, which form the pair $\langle \mathrm{D}_i, \mathbf{R}_i \rangle$. The elements of the base set represent physical entities, such as determinate masses. The relations represent the structure which holds between these entities. For instance, mass value space has a relation $\leqslant$ which defines an order on the set of mass values. This allows us to say that some object is more or less massive than another. Finally, we have a set of functions f between these domains. For instance, the mass function $m(i) : \mathcal{B} \to \mathcal{V}_m$ is a function from the set of bodies $\mathcal{B}$ into 'mass value space' $\mathcal{V}_m$. From these elements, we build a theory's models: mathematical structures which represent ways the world could be if the theory were true. Consider the toy model $\langle \mathcal{B}, \mathcal{V}_m, m(i) \rangle$, which represents a collection of massive bodies. If i denotes some body, then $m(i)$ denotes an element m of $\mathcal{V}_m$ which represents a's mass. Of course, a more realistic model will also include a spacetime in which these bodies are located, as well as a set of dynamical equations which dictate their motions. But for our purposes, this toy model suffices.

Recall the Rovellian insight: we cannot individuate the values of physical quantities without reference to the objects which instantiate them. This fact is reflected at the mathematical level: we cannot uniquely refer to an element of $\mathcal{V}_m$ via a description of its position within the mathematical structure determined by the relations in $\mathbf{R}_m$. For instance, $\mathcal{V}_m$ contains no smallest element, nor does it contain a preferred unit element. Compare this with the more familiar mathematical structure of a vector space. The vector space has no preferred basis, since any set of linearly independent vectors serves equally well as a basis within which we can express all other vectors. The same is true of the mathematical structures that represent physical quantities. But the fact that a value space has no preferred unit entails that it has non-trivial automorphisms. Here, an automorphism is a systematic permutation of the value space's domain which preserves the relations between them. For instance, a uniform scaling of all mass values preserves both their order and their composition, so is an automorphism of mass value space. The existence of non-trivial automorphisms is thus a consequence of the fact that value spaces represent physical quantities which are qualitatively individuated. But it is exactly the existence of such automorphisms that entails the presence of symmetries. For an automorphism of one of the theory's structured domains induces an isomorphism of the theory's models. In particular, we can 'pull-back' the function $m(i)$ by such an automorphism. If we denote said automorphism with ψ, and the pull-back of $m(i)$ by ψ with $\psi^*m(i)$, this induces a transformation on the models of this theory of the form $\langle \mathcal{B}, \mathcal{V}_m, m(i)\rangle \rightarrow \langle \mathcal{B}, \mathcal{V}_m, \psi^*m(i)\rangle$. The symmetries of a sophisticated theory just are these induced isomorphisms.

By definition, isomorphic models have the same structure, that is, isomorphisms preserve the relations between the elements of the domain. Therefore, isomorphic models represent qualitatively identical possibilities. From sophistication's commitment to anti-haecceitism, it follows that isomorphic models represent numerically identical possibilities. But isomorphic models themselves *are* numerically distinct. After all, an isomorphism results from a systematic permutation of the theory's underlying domain, so after such a permutation different objects occupy different 'nodes' in the web of structural relations. But since the elements of the theory's domain are mere place-holders for the roles within a mathematical structure, it doesn't matter which element occupies which node. In the words of Caulton and Butterfield (2012), "[m]odel theory allows without demur that an object may be individuated independently of its qualitative role in a structure" (79). The existence of symmetries is therefore a necessary consequence of the way we use mathematics to model the physical world.[4]

Finally, what is the *function* of symmetries on this picture? On the orthodox view, symmetries were seen as tell-tale signs of excess structure. Theories with symmetries were said to employ derivative quantities; the remedy was to adopt a novel set of invariant quantities instead. On the picture I have sketched, on the other hand, a theory's symmetries reflect the structure of nature's fundamental quantities themselves. The fact that the values of these quantities are only characterised in terms of their structure—"up to isomorphism"—means that we can always represent the same physical pattern in many different ways. If philosophers of physics were sculptors and a theory's models their sculptures, then the orthodoxy views symmetries as a guide to which pieces of stone one ought to remove. If a sculpture is mirror-symmetric, then the orthodoxy wants to cut the sculpture in half. But for the sophisticated sculptor, symmetries act as constraints on the final form of the sculpture: a statue with the right symmetries counts not as a failure but as a satisfying success.

The symmetries of a theory's laws are a consequence of symmetries in nature. But in practice, we only know the former. The positive role of symmetries then is to reveal the latter: the theory's dynamical symmetries reveal the kinematical symmetries of physical quantities. In Chapter 5, I outlined a set of principles which connect these symmetries. In brief, I argued that the symmetries of the theory's laws (or *dynamical symmetries*) ought to coincide with the symmetries of the theory's value spaces (or

[4]I believe this conclusion is in line with Weatherall (2018), who argues that we ought to treat mathematically equivalent structures *as if* they are the same. See also Bradley and Weatherall's (2020) second definition of surplus structure, which is distinct from Ismael and van Fraassen's.

kinematical symmetries). If the theory has fewer dynamical than kinematical symmetries, then it is committed to dynamically irrelevant excess structure. This is a return to literalism, on which symmetry-related models represent worlds which differ only in the values of dynamically inefficacious quantities. If the theory has more dynamical than kinematical symmetries, it endows the world with less structure than is required to sustain the laws. In that case, the primitive identities of physical entities matter to the dynamics, contrary to our rejection of haecceitism. But when a theory's dynamical and kinematical symmetries coincide, it has neither too much nor too little structure. This means that we can use the dynamical symmetries to determine the kinematical symmetries. The kinematical symmetries represent the structure of physical quantities. Hence the slogan: *symmetries are a guide to the structure of physical quantities.*

Bibliography

Adams, R. M. 1979. Primitive Thisness and Primitive Identity. *The Journal of Philosophy 76*(1): 5–26.

Aharonov, Y. and D. Bohm. 1959. Significance of Electromagnetic Potentials in the Quantum Theory. *Physical Review 115*(3): 485–491.

Armstrong, D. M. 1988. Are Quantities Relations? A Reply to Bigelow and Pargetter. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 54*(3): 305–316.

Arntzenius, F. 2012. *Space, Time, and Stuff.* Oxford, New York: Oxford University Press.

Arntzenius, F. and C. Dorr. 2012. Calculus as Geometry. In *Space, Time, and Stuff.* Oxford University Press.

Baez, J. and J. P. Muniain. 1994. *Gauge Fields, Knots and Gravity.* Singapore ; River Edge, N.J: World Scientific Publishing Company.

Baker, D. J. 2010. Symmetry and the Metaphysics of Physics. *Philosophy Compass 5*(12): 1157–1166.

Baker, D. J. 2014. Some Consequences of Physics for the Comparative Metaphysics of Quantity. http://philsci-archive.pitt.edu/12674/.

Bar-Hillel, Y. 1951. A Note on State-Descriptions. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 2*(5): 72–75.

Barrett, J. W. 1991. Holonomy and path structures in general relativity and Yang-Mills theory. *International Journal of Theoretical Physics 30*(9): 1171–1215.

Barrett, T. W. 2017. What Do Symmetries Tell Us About Structure? *Philosophy of Science* (4): 617–639.

Barrett, T. W. 2020. How to count structure. *Noûs n/a*(n/a).

Belot, G. 1998. Understanding Electromagnetism. *The British Journal for the Philosophy of Science 49*(4): 531–555.

Belot, G. 2000. Geometry and Motion. *British Journal for the Philosophy of Science 51*(4): 561–95.

Belot, G. 2003a. Notes on Symmetries. In *Symmetries in Physics: Philosophical Reflections*, eds. K. A. Brading and E. Castellani, 393–412. Cambridge University Press.

Belot, G. 2003b. Symmetry and Gauge Freedom. http://philsci-archive.pitt.edu/1513/.

Belot, G. 2011. *Geometric Possibility.* Oxford University Press.

Belot, G. 2013. Symmetry and Equivalence. In *The Oxford Handbook of Philosophy of Physics*, ed. R. Batterman, 318–339. Oxford University Press.

Belot, G. 2018. Fifty Million Elvis Fans Can't be Wrong. *Noûs 52*(4): 946–981.

Bigelow, J. and R. Pargetter. 1988. Quantities. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 54*(3): 287–304.

Black, R. 2000. Against quidditism. *Australasian Journal of Philosophy 78*(1): 87–104.

Bourbaki, N. 1957. *Theory of Sets.* Elements of Mathematics. Berlin Heidelberg: Springer-Verlag.

Boyd, R. 1991. On the Current Status of Scientific Realism. In *The Philosophy of Science*, eds. R. Boyd, P. Gasper, and J. D. Trout, 195–222. MIT Press.

Bradley, C. and J. O. Weatherall. 2020. On Representational Redundancy, Surplus Structure, and the Hole Argument. *arXiv:1904.04439 [physics].*

Bricker, P. 1993. The Fabric of Space: Intrinsic vs. Extrinsic Distance Relations. *Midwest Studies In Philosophy 18*(1): 271–294.

Brown, H. 2016. The (magnetic) Aharonov-Bohm effect [lecture notes].

Butterfield, J. 1984. Relationism and Possible Worlds. *The British Journal for the Philosophy of Science 35*(2): 101–113.

Butterfield, J. 2006. Against Pointillisme about Mechanics. *The British Journal for the Philosophy of Science 57*(4): 709–753.

Butterfield, J. 2011. Emergence, Reduction and Supervenience: A Varied Landscape. *Foundations of Physics 41*(6): 920–959.

Carnap, R. 1928. *Der Logische Aufbau Der Welt.* Meiner Verlag.

Carnap, R. 1950. *Logical Foundations of Probability.* University of Chicago Press.

Carnap, R. 1951. The Problem of Relations in Inductive Logic. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 2*(5): 75–80.

Caulton, A. 2013. A Humean alternative on what there is – PERSP, Universitat Autonoma de Barcelona (Presentation).

Caulton, A. 2015. The Role of Symmetry in the Interpretation of Physical Theories. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 52: 153–162.

Caulton, A. and J. Butterfield. 2012. On Kinds of Indiscernibility in Logic and Metaphysics. *The British Journal for the Philosophy of Science 63*(1): 27–84.

Chen, E. K. 2018. The Intrinsic Structure of Quantum Mechanics. http://philsci-archive.pitt.edu/15140/.

Dasgupta, S. 2013. Absolutism vs Comparativism about Quantity. In *Oxford Studies in Metaphysics: Volume 8*. Oxford University Press.

Dasgupta, S. 2016. Symmetry as an Epistemic Notion (Twice Over). *The British Journal for the Philosophy of Science 67*(3): 837–878.

Dasgupta, S. 2018. How to Be a Relationalist. *Unpublished*.

Dasgupta, S. 2021. Symmetry and Superfluous Structure: A Metaphysical Overview. In *The Routledge Companion to Philosophy of Physics*, eds. E. Knox and A. Wilson. Routledge.

Debs, T. A. and M. L. G. Redhead. 1996. The twin "paradox" and the conventionality of simultaneity. *American Journal of Physics 64*(4): 384–392.

Dees, M. 2015. Maudlin on the Triangle Inequality. *Thought: A Journal of Philosophy 4*(2): 124–130.

Denby, D. A. 2001. Determinable Nominalism. *Philosophical Studies 102*(3): 297–327.

Dewar, N. 2015. Symmetries and the Philosophy of Language. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 52*(Part B): 317–327.

Dewar, N. 2018. Maxwell Gravitation. *Philosophy of Science 85*(2): 249–270.

Dewar, N. 2019. Sophistication about Symmetries. *The British Journal for the Philosophy of Science 70*(2): 485–521.

Dewar, N. 2020a. General-Relativistic Covariance. *Foundations of Physics 50*(4): 294–318.

Dewar, N. 2020b. On Absolute Units. http://philsci-archive.pitt.edu/18711/.

Dewar, N. 2022. *Structure and Equivalence*. Cambridge University Press.

Dirac, P. A. M. 1930. *The Principles of Quantum Mechanics*. Clarendon Press.

Dorr, C. and J. Hawthorne. 2013. Naturalness. In *Oxford Studies in Metaphysics: Volume 8*, eds. K. Bennett and D. Zimmerman, 1. Oxford University Press.

Dougherty, J. 2017. Sameness and Separability in Gauge Theories. *Philosophy of Science 84*(5): 1189–1201.

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

Earman, J. 2004. Laws, Symmetry, and Symmetry Breaking: Invariance, Conservation Principles, and Objectivity. *Philosophy of Science 71*(5): 1227–1241.

Earman, J. and J. Norton. 1987. What Price Spacetime Substantivalism? The Hole Story. *The British Journal for the Philosophy of Science 38*(4): 515–525.

Eddon, M. 2013a. Fundamental Properties of Fundamental Properties. In *Oxford Studies in Metaphysics, Volume 8*, ed. K. B. D. Zimmerman, 78–104.

Eddon, M. 2013b. Quantitative Properties. *Philosophy Compass 8*(7): 633–645.

Ellis, B. 1966. *Basic Concepts of Measurement*. Cambridge University Press.

Esfeld, M. and V. Lam. 2008. Moderate Structural Realism About Space-Time. *Synthese 160*(1): 27–46.

Feynman, R. P., R. B. Leighton, and M. Sands. 1964. *The Feynman Lectures on Physics*, Volume III. Reading: Addison-Wesley.

Field, H. 1984. Can We Dispense with Space-Time? *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1984: 33–90.

Field, H. H. 1980. *Science without Numbers: A Defence of Nominalism*. Princeton, N.J: Princeton University Press.

Fine, K. 2012. Guide to Ground. In *Metaphysical Grounding*, eds. F. Correia and B. Schnieder, 37–80. Cambridge University Press.

Fitting, M. 1996. *First-Order Logic and Automated Theorem Proving*. New York, NY: Springer New York.

Fletcher, S. C. 2020. On Representational Capacities, with an Application to General Relativity. *Foundations of Physics 50*(4): 228–249.

Forrest, P. 1995. Is Space-Time Discrete or Continuous? — An Empirical Question. *Synthese 103*(3): 327–354.

French, S. and J. Ladyman. 1999. Reinflating the semantic approach. *International Studies in the Philosophy of Science 13*(2): 103–121.

Friedman, M. 1974. Explanation and Scientific Understanding. *Journal of Philosophy 71*(1): 5–19.

Friedman, M. 1983. *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton University Press.

Funkhouser, E. 2006. The Determinable-Determinate Relation. *Noûs 40*(3): 548–569.

Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought* (New Ed edition ed.). Cambridge, Mass.: MIT Press.

Gomes, H. 2019. Holism as the significance of gauge symmetries. http://philsci-archive.pitt.edu/16499/.

Greaves, H. 2011. In Search of (Spacetime) Structuralism. *Philosophical Perspectives* 25: 189–204.

Greaves, H. and D. Wallace. 2014. Empirical Consequences of Symmetries. *The British*

Journal for the Philosophy of Science 65(1): 59–89.

Halvorson, H. and D. Tsementzis. 2018. Categories of Scientific Theories. In *Categories for the Working Philosopher*, ed. E. Landry. Oxford University Press.

Healey, R. 1997. Nonlocality and the Aharonov-Bohm Effect. *Philosophy of Science 64*(1): 18–41.

Healey, R. 2006. Symmetry and the Scope of Scientific Realism. In *Physical Theory and Its Interpretation: Essays in Honor of Jeffrey Bub*, eds. W. Demopoulos and I. Pitowsky, The Western Ontario Series in Philosophy of Science, 143–160. Dordrecht: Springer Netherlands.

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

Healey, R. 2009. Perfect Symmetries. *British Journal for the Philosophy of Science 60*(4): 697–720.

Hetzroni, G. 2019. Gauge and Ghosts. *The British Journal for the Philosophy of Science* (axz021).

Hölder, O. 1901. Die Axiome der Quantität und die Lehre vom Maß. *Berichte über die Verhandlungen der Königlich Sächsichen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physikalische Classe* 53: 1–64.

Huggett, N. 2006. The Regularity Account of Relational Spacetime. *Mind 115*(457): 41–73.

Isham, C. J. 1999. *Modern Differential Geometry for Physicists: Second Edition* (2nd Revised edition edition ed.). Singapore ; River Edge, N.J: WSPC.

Ismael, J. MS. Simplicity as a Guide to Scientific Metaphysics.

Ismael, J. and J. Schaffer. 2020. Quantum holism: Nonseparability as common ground. *Synthese 197*(10): 4131–4160.

Ismael, J. and B. C. van Fraassen. 2003. Symmetry as a Guide to Superfluous Theoretical Structure. In *Symmetries in Physics: Philosophical Reflections*, eds. K. Brading and E. Castellani, 371–92. Cambridge University Press.

Jacobs, C. 2020. Absolute Velocities Are Unmeasurable: Response to Middleton and Murgueitio Ramirez. *Australasian Journal of Philosophy*.

Jacobs, C. 2021. Invariance or Equivalence: A Tale of Two Principles. *Synthese*.

Janssen, M. 2002. COI Stories: Explanation and Evidence in the History of Science. *Perspectives on Science 10*(4): 457–522.

Janssen, M. 2009. Drawing the Line Between Kinematics and Dynamics in Special Relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 40*(1): 26–52.

Kitcher, P. 1981. Explanatory Unification. *Philosophy of Science 48*(4): 507–531.

Knox, E. 2014. Newtonian Spacetime Structure in Light of the Equivalence Principle. *The British Journal for the Philosophy of Science 65*(4): 863–880.

Kobayaschi, S. 1957. Theory of connections. *Annali di Matematica Pura ed Applicata 43*(1): 119–194.

Kosso, P. 2000. The Empirical Status of Symmetries in Physics. *The British Journal for the Philosophy of Science 51*(1): 81–98.

Krantz, D. H., P. Suppes, and R. D. Luce. 1971. *Foundations of Measurement I: Additive and Polynomial Representations.* New York: Academic Press.

Ladyman, J., D. Ross, D. Spurrett, and J. Collier. 2007. *Every Thing Must Go: Metaphysics Naturalized.*

Leeds, S. 1999. Gauges: Aharonov, Bohm, Yang, Healey. *Philosophy of Science 66*(4):

606–627.

Lewis, D. 1970. How to Define Theoretical Terms. *Journal of Philosophy 67*(13): 427–446.

Lewis, D. 1973. *Counterfactuals.* Malden, Mass: Wiley-Blackwell.

Lewis, D. 1983. Extrinsic Properties. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 44*(2): 197–200.

Lewis, D. 1986. *Philosophical Papers Volume II.* Oxford University Press.

Luce, R. D., D. H. Krantz, P. Suppes, and A. Tversky. 2007. *Foundations of Measurement Volume III: Representation, Axiomatization, and Invariance.* Mineola, N.Y: Dover Publications.

Lyre, H. 2004. Holism and Structuralism in U(1) Gauge Theory. http://philsci-archive.pitt.edu/1831/.

Malament, D. 1977. Causal Theories of Time and the Conventionality of Simultaneity. *Noûs 11*(3): 293–300.

Malament, D. B. 2012. *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory.* Chicago: Chicago University Press.

Maldacena, J. 2014. The symmetry and simplicity of the laws of physics and the Higgs boson. *arXiv:1410.6753 [hep-ex, physics:hep-lat, physics:hep-ph, physics:hep-th, physics:physics]*.

Martens, N. C. M. 2017. Regularity Comparativism about Mass in Newtonian Gravity. *Philosophy of Science 84*(5): 1226–1238.

Martens, N. C. M. 2019a. Machian Comparativism about Mass. *The British Journal for the Philosophy of Science*.

Martens, N. C. M. 2019b. The (un)detectability of absolute Newtonian masses. *Synthese*.

Martens, N. C. M. and J. Read. 2020. Sophistry about symmetries? *Synthese*.

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. 1998. Healey on the Aharonov-Bohm Effect. *Philosophy of Science 65*(2): 361–368.

Maudlin, T. 2007. Suggestions from Physics for Deep Metaphysics. In *The Metaphysics Within Physics.* Oxford University Press.

Melia, J. 2000. Weaseling Away the Indispensability Argument. *Mind 109*(435): 455–479.

Menon, T. V. 2018. *The Dynamical Foundations of Physical Geometry.* D.Phil thesis, University of Oxford.

Middleton, B. and S. M. Murgueitio Ramírez. 2020. Measuring Absolute Velocity. *Australasian Journal of Philosophy 0*(0): 1–11.

Møller-Nielsen, T. 2017. Invariance, Interpretation, and Motivation. *Philosophy of Science 84*(5): 1253–1264.

Mundy, B. 1986. On the General Theory of Meaningful Representation. *Synthese 67*(3): 391–437.

Mundy, B. 1987. The Metaphysics of Quantity. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 51*(1): 29–54.

Myrvold, W. C. 2011. Nonseparability, Classical, and Quantum. *The British Journal for the Philosophy of Science 62*(2): 417–432.

Myrvold, W. C. 2019. How could relativity be anything other than physical? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 67: 137–143.

North, J. 2018. A New Approach to the Relational–Substantival Debate. In *Oxford Studies* *in Metaphysics: Volume 11*, eds. K. Bennett and D. Zimmerman, Chapter Oxford Studies in Metaphysics Volume 11. Oxford University Press.

Nounou, A. M. 2003. A Fourth Way to the Aharonov-Bohm Effect. In *Symmetries In Physics: Philosophical Reflections*, eds. K. Brading and E. Castellani. Cambridge University Press.

Nozick, R. 2001. *Invariances: The Structure of the Objective World.* Cambridge, Mass.: Harvard University Press.

Pooley, O. 2006. Points, Particles, and Structural Realism. In *The Structural Foundations of Quantum Gravity*, eds. D. Rickles, S. French, and J. T. Saatsi, 83–120. Oxford University Press.

Pooley, O. 2013. Substantivalist and Relationalist Approaches to Spacetime. In *The Oxford Handbook of Philosophy of Physics*, ed. R. Batterman. Oxford University Press.

Psillos, S. 1999. *Scientific Realism: How Science Tracks Truth.* Routledge.

Putnam, H. 1975. *Mathematics, Matter and Method: Volume 1, Philosophical Papers.* CUP Archive.

Quine, W. V. 1948. On What There Is. *The Review of Metaphysics 2*(5): 21–38.

Quine, W. V. 1951. Ontology and Ideology. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 2*(1): 11–15.

Quine, W. V. O. 1960. *Word and Object.* MIT Press.

Read, J. and T. Møller-Nielsen. 2018. Motivating Dualities. *Synthese*.

Read, J. and T. Møller-Nielsen. 2020. Redundant epistemic symmetries. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.

Reichenbach, H. 1956. *The Direction of Time.* Dover Publications.

Rindler, W. 1977. *Essential Relativity: Special, General, and Cosmological.* Springer Science & Business Media.

Roberts, B. W. 2020. Regarding 'Leibniz Equivalence'. *Foundations of Physics 50*(4): 250–269.

Roberts, J. T. 2008. A Puzzle about Laws, Symmetries and Measurability. *The British Journal for the Philosophy of Science 59*(2): 143–168.

Roberts, J. T. 2016. A Case for Comparativism about Physical Quantities – SMS 2016, Geneva (Presentation).

Rosen, G. 2010. Metaphysical Dependence: Grounding and Reduction. In *Modality: Metaphysics, Logic, and Epistemology*, eds. B. Hale and A. Hoffmann, 109–36. Oxford University Press.

Rosenstock, S. and J. O. Weatherall. 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.

Rovelli, C. 2014. Why Gauge? *Foundations of Physics 44*(1): 91–104.

Russell, B. 1903. *The Principles of Mathematics.* University Press.

Russell, J. S. 2018. Quality and Quantifiers. *Australasian Journal of Philosophy 96*(3): 562–577.

Rynasiewicz, R. 1992. Rings, Holes and Substantivalism: On the Program of Leibniz Algebras. *Philosophy of Science 59*(4): 572–589.

Rynasiewicz, R. 1994. The Lessons of the Hole Argument. *British Journal for the Philosophy of Science 45*(2): 407–436.

Rynasiewicz, R. 1996. Is There a Syntactic Solution to the Hole Problem? *Philosophy of Science* 63: S55–S62.

Rynasiewicz, R. 2012. Simultaneity, convention, and gauge freedom. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 43*(2): 90–94.

Saunders, S. 2003a. Indiscernibles, General Covariance, and Other Symmetries: The Case for Non-Reductive Relationalsm. In *Revisiting the Foundations of Relativistic Physics*, ed. A. Ashtekar, 151–173.

Saunders, S. 2003b. Physics and Leibniz's Principles. In *Symmetries in Physics: Philosophical Reflections*, eds. K. Brading and E. Castellani, 289–307. Cambridge University Press.

Saunders, S. 2007. Mirroring as an A Priori Symmetry. *Philosophy of Science 74*(4): 452–480.

Saunders, S. 2013. Rethinking Newton's Principia. *Philosophy of Science 80*(1): 22–48.

Schaffer, J. 2009. On What Grounds What. In *Metametaphysics: New Essays on the Foundations of Ontology*, eds. D. Manley, D. J. Chalmers, and R. Wasserman, 347–383. Oxford University Press.

Sider, T. 2011. *Writing the Book of the World*. Oxford University Press.

Sider, T. 2020. *The Tools of Metaphysics and the Metaphysics of Science*. Oxford, New York: Oxford University Press.

Spivak, M. 1994. *Calculus* (3rd ed.). Cambridge University Press.

Stachel, J. 1993. The Meaning of General Covariance. In *Philosophical Problems of the Internal and External Worlds: Essays on the Philosophy of Adolph Grünbaum*, eds. J. Earman, A. Janis, and G. Massey, 129–60. University of Pittsburgh Press.

Stalnaker, R. 1979. Anti-Essentialism. *Midwest Studies In Philosophy 4*(1): 343–355.

Suppe, F. 1989. *The Semantic Conception of Theories and Scientific Realism*. University of Illinois Press.

Suppes, P. 1960. A Comparison of the Meaning and Uses of Models in Mathematics and the Empirical Sciences. *Synthese 12*(2-3): 287–301.

Suppes, P. 1967. What Is a Scientific Theory? In *Philosophy of Science Today*, ed. S. Morgenbesser. New York: Basic Books.

Teh, N. J. 2015. A Note on Rovelli's 'Why Gauge?'. *European Journal for Philosophy of Science 5*(3): 339–348.

Teitel, T. 2019. Holes in Spacetime: Some Neglected Essentials. *The Journal of Philosophy 116*(7): 353–389.

Teller, P. 1987. Space-Time as a Physical Quantity. In *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, eds. P. Achinstein and R. Kagon, 425–448. MIT Press.

Teller, P. 1991. Substance, Relations, and Arguments About the Nature of Space-Time. *Philosophical Review 100*(3): 363–397.

van Fraassen, B. C. 1967. Meaning Relations among Predicates. *Noûs 1*(2): 161–179.

van Fraassen, B. C. 1980. *The Scientific Image*. Clarendon Library of Logic and Philosophy. Oxford, New York: Oxford University Press.

van Fraassen, B. C. 1989. *Laws and Symmetry*. Oxford University Press.

Wallace, D. 2002. Time-Dependent Symmetries: The Link Between Gauge Symmetries and Indeterminism. In *Symmetries in Physics: Philosophical Reflections*, eds. K. Brading and E. Castellani, 163–173. Cambridge University Press.

Wallace, D. 2014a. Deflating the Aharonov-Bohm Effect. *arXiv:1407.5073 [quant-ph]*.

Wallace, D. 2014b. Review of Interpreting Quantum Theories, by Laura Ruetsche. *The British Journal for the Philosophy of Science 65*(2): 425–428.

Wallace, D. 2015. Fields as Bodies: A unified presentation of spacetime and internal gauge symmetry. *arXiv:1502.06539 [gr-qc, physics:hep-th]*.

Wallace, D. 2019a. Isolated Systems and their Symmetries, Part I: General Framework and Particle-Mechanics Examples. http://philsci-archive.pitt.edu/16623/.

Wallace, D. 2019b. Observability, redundancy and modality for dynamical symmetry transformations. http://philsci-archive.pitt.edu/16622/.

Wallace, D. 2019c. Who's Afraid of Coordinate Systems? An Essay on Representation of Spacetime Structure. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 67: 125–136.

Wallace, D. and C. G. Timpson. 2010. Quantum Mechanics on Spacetime I: Spacetime State Realism. *The British Journal for the Philosophy of Science 61*(4): 697–727.

Wang, J. 2016. Fundamentality And Modal Freedom. *Philosophical Perspectives 30*(1): 397–418.

Weatherall, J. O. 2015. Are Newtonian Gravitation and Geometrized Newtonian Gravitation Theoretically Equivalent? *arXiv:1411.5757 [gr-qc, physics:physics]*.

Weatherall, J. O. 2016a. Fiber bundles, Yang–Mills theory, and general relativity. *Synthese 193*(8): 2389–2425.

Weatherall, J. O. 2016b. Understanding Gauge. *Philosophy of Science 83*(5): 1039–1049.

Weatherall, J. O. 2017. A Brief Comment on Maxwell(/Newton)[-Huygens] Spacetime. *arXiv:1707.02393 [physics]*.

Weatherall, J. O. 2018. Regarding the 'Hole Argument'. *The British Journal for the Philosophy of Science 69*(2): 329–350.

Wilson, J. 2010. What is Hume's Dictum, and Why Believe It? *Philosophy and Phenomenological Research 80*(3): 595–637.

Wilson, J. 2017. Determinables and Determinates. In *The Stanford Encyclopedia of Philosophy* (Spring 2017 ed.)., ed. E. N. Zalta. Metaphysics Research Lab, Stanford University.

Wolff, J. E. 2020. *The Metaphysics of Quantities*. Oxford, New York: Oxford University Press.