Studies in History and Philosophy of Modern Physics 52 (2015) 153–162

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

The role of symmetry in the interpretation of physical theories

Adam Caulton

University of Cambridge, United Kingdom

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form
4 July 2015
Accepted 5 August 2015
Available online 3 September 2015

Keywords:
Symmetry
Interpretation
Reichenbach
Constitutive a priori
Convention

ABSTRACT

The symmetries of a physical theory are often associated with two things: conservation laws (via e.g. Noether's and Schur's theorems) and representational redundancies ("gauge symmetry"). But how can a physical theory's symmetries give rise to interesting (in the sense of non-trivial) conservation laws, if symmetries are transformations that correspond to no genuine physical difference? In this paper, I argue for a disambiguation in the notion of symmetry. The central distinction is between what I call "analytic" and "synthetic" symmetries, so called because of an analogy with analytic and synthetic propositions. "Analytic" symmetries are the turning of idle wheels in a theory's formalism, and correspond to no physical change; "synthetic" symmetries cover all the rest. I argue that analytic symmetries are distinguished because they act as fixed points or constraints in any interpretation of a theory, and as such are akin to Poincaré's conventions or Reichenbach's 'axioms of co-ordination', or 'relativized constitutive a priori principles'.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

Physical theories—if we take them literally—tell us what world is like, or at least what it might have been like. For example, the consensus is that special relativity, if taken as a literal description of the world, tells us that the world is not a sequence of three-dimensional instants, objectively unfolding over time, but rather a four-dimensional continuum, in which there is no objective temporal order between spacelike separated events. Meanwhile, there seems to be little consensus about what quantum mechanics, taken literally, tells us about the world. Clearly, we cannot extract an account of the physical world from a theory unless we first understand it. But how do we come to understand a theory—that is, how do we *interpret* a theory —if (as we suspect) it describes things that we have never before even *thought* of? In other words: How are we supposed to know what a theory is aiming at—what it represents as holding true—if the candidates for its target can only be understood by the descriptions given of them *by that very theory*?

van Fraassen (2008) has recently argued that this predicament prevents us from giving any literal interpretation to a physical theory, except insofar as it describes observable entities. In this paper, I will marshal some considerations from philosophy of language and philosophy of science to argue that a literal interpretation for a physical theory *can* be found—one that describes observables and non-observables alike—under the guidance of the right interpretative constraints. These constraints are Leibnizian in letter, for they recommend the elimination of distinctions without a difference, in much the same way that Leibniz did in his critique of Newtonian absolute space in his correspondence with Clarke (Alexander, 1998). They appeal to *symmetries* of the theory being interpreted, as guides for what to eliminate. (As I will explain, symmetries are transformations on a theory's space of states or solutions that preserve certain quantities defined over that space.) But the constraints are not Leibnizian in spirit, since the motivation for them is semantic rather than metaphysical or theological—in particular, I will not appeal to the Principle of Sufficient Reason. In fact the spirit of this paper owes at least as much to Reichenbach (1958, 1965), and Carnap (1966) as to Leibniz.

The paper is divided into two main parts, each focussed on what I will call 'a triangle of concepts'. In the first part (Section 2), I give a general account of physical theories and their symmetries. There I introduce my first triangle of concepts: these are a theory's states, quantities and symmetries. I propose a dichotomy between symmetries into what I call *analytic* and *synthetic*, so-called because of analogies with the familiar characterisation of propositions in those terms. This dichotomy, applied to symmetries, is crucial for my recommendations in part two for interpreting a physical theory.

E-mail address: adam.caulton@gmail.com

http://dx.doi.org/10.1016/j.shpsb.2015.08.002
1355-2198/

The second part (Section 3) contains a discussion of theory interpretation, and makes proposals for how to interpret a theory that is taken to be about (amongst other things) as-yet-unconceived entities. There I introduce my second triangle of concepts: a theory's formalism, its subject matter, and the representation relation that exists between them. The second triangle is linked to the first: for which symmetries count as analytic and which count as synthetic will hang on the details of this representation relation, and *vice versa*. Therefore a constraint on the symmetries will serve as a constraint on the representation relation, hopefully one strong enough to determine a unique interpretation for the theory.

Here I must immediately qualify the sense of 'interpretation' under which it would be plausible to claim that a unique one might be found. As we shall see in Section 2, my main concern is the representation relation that may hold between possible states, or histories, and physical quantities (i.e. properties and relations) on the one hand, and their mathematical representatives on the other: in particular whether it is one-to-one or one-to-many. In these terms, the sense of interpretation I have in mind must be rather coarse-grained: it is intensional, as opposed to hyper-intensional. As such it will leave many metaphysical questions unanswered—specifically ones which make no difference to the variety of possible states or histories. For example, an interpretation will not settle whether it is better to interpret unit electric charge as a universal or a collection of intrinsically similar tropes, or whether gravitational attraction in Newtonian particle mechanics is a direct relation between particles or mediated by a gravitational field.[1] However, an interpretation in my sense *will* settle questions of ontology insofar as these lead to differences in the corresponding variety of possibilities. These include the existence or otherwise of: absolute location, orientation or velocity; 'haecceities' for spacetime points or elementary particles; and the electromagnetic four-potential. The main purpose of the paper is to argue that, by maximising the number of the theory's analytic symmetries, subject to empirical adequacy, a unique interpretation—in this coarse-grained sense of 'interpretation'—can be found.

2. States, quantities and symmetries: the first triangle

2.1. What is a physical theory?

I should begin by saying a few words about what I take a physical theory to be. I want to be general enough incorporate a variety of theories, both classical and quantum. My account will not be particularly novel, nor even adequate for all purposes; but it will be adequate for the purposes of this paper. I will use some ideas and techniques associated with the syntactic view of theories (e.g. I partly characterise a theory by its axioms) and some associated with the semantic view (e.g. I partly characterise a theory by its models): I see no reason to opt for one approach over the other.[2]

So I propose to take a physical theory to comprise the following four components:

1. A *language* $\mathcal{L}$: This is intended to be the language in which the equations of motion of the theory are expressed; but it should also incorporate the fragment of simple English that surrounds these equations. (E.g. 'Let $\mathcal{M}$ be a differentiable manifold,' or 'The seventh factor Hilbert space represents the possible states for particle seven.') I propose to give a semantics for this language that follows (at least in spirit) the general scheme laid down by Montague (1970) and Lewis (1970). That is, it is *intensional* and *compositional*.
   Briefly, intensionality means that linguistic items—particularly sentences—are assigned intensions. For many linguistic items (specifically, the ones which may be assigned extensions—i.e., sentences and noun-phrases, whose extensions are truth-values and worldly objects, respectively), the intension is a function from *indices* to extensions. In Montague's theory, an index is a possible world, construed as a Tarskian model, or an ordered sequence comprising a possible world and appropriate context-determining parameters, such as the time and place of utterance.[3] However, for the purposes of interpreting a physical theory, I propose that we here stray from Montague by supplanting Tarskian models with the possible states or histories already belonging to the mathematical formalism of our given theory of interest; see below.
   Compositionality is achieved by breaking down whole sentences (e.g. 'the electron is spin-up') into sub-sentential components whose intensions have a function/argument structure. So, for example, the sentence 'the electron is spin-up' (whose intension maps any possible state in which the electron is spin-up to *true* and all others to *false*) is broken down into the noun-phrase 'the electron' (whose intension maps any possible state to the electron, if there is a unique one, in that state) and the verb-phrase 'is spin-up', whose intension is a function from noun-phrase-intensions to sentence-intensions; specifically, it maps the intension of 'the electron' to the intension of 'the electron is spin-up'.
   An important collection of noun-phrases used in physical theories are those used to refer to determinates, such as '2 kg in mass', 'distance of 3 km', 'spin-up', etc. This suggests that we include amongst our worldly objects not just things like particles or spacetime points, but also masses, distances and vector of spin. These *dimensioned values* lie in some "logical space"—à la van Fraassen (1967) and Stalnaker (1979)—which has a mathematical parametrisation, e.g. by the real line or the vector space of anti-symmetric rank-2 tensors. Of course, any gloss given to these determinates at this stage is at best provisional, since we have not yet interpreted the theory! I return to determinates below, in the discussion of quantities.
2. A space of *mathematical states* $\mathcal{S}$: Almost all theories are equipped with a space of its own "possible worlds".[4] They have many commonalities with Tarskian models and Lewis (1986) possible worlds, most crucial among them being that their role is to represent possibilities regarding the actual (physical) world: i.e., ways the world and its constituents might have been. However, there are two crucial differences between these states and both Tarskian models and Lewisian possible worlds. The first crucial difference is that, while the means by which Tarskian models and Lewisian possible worlds represent possibilities is (at least very often taken to be) well-defined and unproblematic, the representation relation between the mathematical states of a physical theory's formalism and corresponding concrete possibilities is far from obvious. Indeed, a major task in the interpretation of any physical theory is precisely settling on a unique representation relation; I return to this below in Section 3.1. As a result, the interpretation of a given physical theory will not be complete when a Montague semantics is found for its language: for the intensions in any such semantics will make ineliminable reference to mathematical states whose physical correlates have not yet been laid down. It is only when the further task of settling the representation relation

[1] I am grateful to an anonymous referee for this example.

[2] The classic account of the failings of the syntactic view is Suppe (1974). For details of the semantic view, see e.g. French & Ladyman (1999), van Fraassen (2000) and Frigg & Hartmann (2008). For a compelling defence of the syntactic view, see Lutz (2010).

[3] For a comprehensive discussion of the treatment of context gestured at here, see Lewis (1980).

[4] A notable exception may be quantum field theories—at least on the algebraic approach—each of whose possible states cannot be captured by any single separable Hilbert space.