Synthese (2021) 198:2973–3013
https://doi.org/10.1007/s11229-019-02258-x

SYMMETRIES AND ASYMMETRIES IN PHYSICS

On symmetry and duality

Sebastian De Haro[1,2,3] · **Jeremy Butterfield**[1]

Received: 25 September 2018 / Accepted: 15 May 2019 / Published online: 29 May 2019

Abstract

We advocate an account of dualities between physical theories: the basic idea is that dual theories are isomorphic representations of a common core. We defend and illustrate this account, which we call a Schema, in relation to symmetries. Overall, the account meshes well with standard treatments of symmetries. But the distinction between the common core and the dual theories prompts a distinction between three kinds of symmetry: which we call 'stipulated', 'accidental' and 'proper'.

Keywords Duality · Symmetry · Theories · Models

Contents

✉ Sebastian De Haro
sd696@cam.ac.uk

Jeremy Butterfield
jb56@cam.ac.uk

1 Trinity College, Cambridge CB2 1TQ, United Kingdom

2 Department of History and Philosophy of Science, University of Cambridge, Cambridge, UK

3 Vossius Center for History of Humanities and Sciences, University of Amsterdam, Amsterdam, Netherlands

1 Introduction

Symmetries have long been a central topic in the philosophy of physics: as witness collections such as Brading and Castellani (2003) and this special issue. But dualities between physical theories have only more recently become a focus of interest.[1] In this paper we bring the topics together, by applying an account of dualities that we have developed elsewhere.

Bringing the topics together is natural. For in the literature, it is agreed by all hands that a duality is like a "giant symmetry": a symmetry between theories. For in physics, the basic idea of a symmetry is a map taking a state of the system into another appropriately related state; and correspondingly mapping physical quantities—details below. And in a duality, an entire theory is mapped into another appropriately related theory.

Our account of dualities will confirm this basic analogy. The leading idea will be: (i) the preservation, or appropriate matching, of a state's values for various quantities, and (ii) this preservation or matching being maintained by the dynamics of the system. (We say 'preservation or matching' so as to respect the distinction between invariance and covariance, and 'dynamics' is to include Euclidean systems: details below.)

In short: our account—we call it a 'Schema'—holds that a duality between two theories requires that:

(a) The two theories share a *common core*; the common core is itself a theory, which we will call the *bare theory*; (for us, a *bare* theory or model is a theory or model stripped of its interpretation; more details in Sect. 2.1); and

(b) The two given theories are *isomorphic models* of this common core, the bare theory.

Here, we understand 'models' as realizations or formulations. They are almost always representations in the sense of representation theory, i.e. homomorphic copies of the bare theory. Representations are of course in general *not* isomorphic. But we say that duality is a matter of two models of a common core, a bare theory, being indeed isomorphic with respect to the structure of that core.

[1] See for example Castellani (2017), Dieks et al. (2015), Fraser (2017), Huggett (2017), Matsubara (2013), Read (2016), Rickles (2017).

We stress that here we mean 'representation' in the mathematical, not the philosophical, sense. The bare dual theories do not interpret the bare common core theory (notwithstanding our use of the word 'model'). Rather they are specific realizations or formulations of the bare theory, like the matrix representations of an abstract group. But agreed: the dual theories, and also often the bare theory, *do* get interpreted. Interpretation, and our choice of jargon, is further explained in Sect. 2.1.

We have developed this Schema in various papers. The most detailed one is our (2018). It also illustrates the Schema with a major example: bosonization, which is a duality between two quantum field theories in two spacetime dimensions—one with bosons, and one with fermions. So here, a bosonic model is isomorphic to a fermionic model; (their common core is a certain infinite-dimensional algebra). Other papers (De Haro 2016a, 2018b; Butterfield 2018) discuss in more detail duality's relation to the interpretation of theories, especially the various senses in which theories get called 'equivalent'. Other papers discuss other advanced, indeed conjectural, examples in string theory (De Haro 2015, 2017; De Haro et al. 2016, 2017). And yet other papers discuss the less formal aspects of dualities: in particular, their heuristic role in finding a theory "behind" the bare theory, of which the dual models are only approximations not representations (De Haro 2018a), and in relations to understanding (De Haro and De Regt 2018a, b).[2]

Thus the aim of this paper is to relate the Schema to symmetries in more detail than we have done previously. (Our previous discussion was in (2018: Sects. 3.1.1, 3.2.4).) The main topic will be the relations between symmetries of the common core theory, and symmetries of one of the dual theories (one of the models of the common core). This topic is an important preliminary to discussing the 'interaction' of symmetries with dualities. It will call for care about what parts of a model "do the representing" of a common core; (cf. Sect. 2.2's distinction between *model root* and *specific structure*). It will also lead to a classification of kinds of symmetry (Sect. 5, summarized in Sect. 5.4). Thus we will give as much emphasis to symmetries as to dualities (usually treating symmetries first).

To be as clear as possible, we will build up the details successively. In effect, Sects. 2 and 3 set the stage for the main claims in the second half of the paper. We begin by explaining the contrast between the two 'levels': the bare theory, and its models (realizations). This contrast is a prerequisite to our discussing symmetries and dualities, since it can (should!) be explained for the case of a single model (realization): i.e. regardless of whatever dualities, or symmetries, may hold good. This we do in

[2] We admit at the outset that because our Schema defines duality formally, without regard to the isomorphism being surprising or scientifically important, it has many illustrations that are unremarkable, and indeed not usually called dualities. We will see examples below: cf. Sect. 2.3. Contrast a case like bosonization. Since bosons and fermions are very different, it is indeed a surprising isomorphism. It is also scientifically important because it pairs situations of strong coupling in one dual (so that problems are in general difficult to solve, since perturbation theory cannot be trusted) with situations of weak coupling in the other dual: so one can sometimes solve a problem in the weak coupling regime of one dual and transfer the result, so as to solve an intractable problem in the other dual. Broadly speaking, dualities' scientific importance depends on these features: the isomorphism being surprising, and its relating strong and weak coupling regimes. But features like importance and surprise are hard to be precise about: hence our decision that it is best to define duality formally, even at the expense of countless trivial examples. For more discussion, cf. De Haro and Butterfield (2018: Sect. 2.1) and De Haro (2018a: Sects. 1, 3.3).

Sect. 2. As we will see, this involves various issues of interpretation, even controversy. And it prompts our distinction between *model root* and *specific structure*.

We are thereby equipped to give a brief exposition of our Schema, and a more detailed prospectus of the second half of the paper—i.e. of our account of the relations between symmetries and dualities. These are in Sect. 3.

This account starts with symmetries in general: without regard to dualities, or even our distinction between theories and the models that realize or represent them; (Sect. 4). Here, we relate our basic idea of symmetry, as a map on states that preserves quantities' values, to some familiar topics: such as dynamical symmetries, how spacetime theories treat symmetries, and what is the subset of quantities that it is "worthwhile" for symmetries to preserve (cf. Caulton 2015). Then we discuss how symmetries fit with the theory-model relation: but again without regard to dualities; (Sect. 5). Here, the differing amounts of structure at the levels of the bare theory and its models—the distinction between model root and specific structure—will prompt definitions of three kinds of symmetry: which we call *stipulated*, *accidental* and *proper*. Finally in Sect. 6, we describe how on our account, dualities preserve symmetries in a straightforward way. Sect. 7 concludes.

A final preliminary: our scope is limited. We will not engage in detail with two debates that have been prominent in the recent philosophical literature about symmetries. For these debates are orthogonal to most of this paper's issues: which centre around the relations between symmetries of the common core, and symmetries of its models. But for completeness, we here briefly note these debates, and our broad view about them. They both concern whether symmetry is always a sign of 'surplus structure', 'redundancy' or 'gauge': for the relation of this to duality, cf. our (2018: Sect. 3.2.4)

(1) Should two symmetry-related solutions of a theory: (a) be interpreted *ab initio* as representing the same physical state of affairs?; or (b) be taken merely to motivate searching for a common ontology that secures such an interpretation? This debate is articulated by Møller-Nielsen (2017) and, in relation to dualities, Read and Møller-Nielsen (2018). These authors defend option (b). Broadly speaking, we agree with them, about symmetries as well as dualities: cf. De Haro (2016a: Sect. 1.3), Butterfield (2018: Sects. 1.2, 3.3) and De Haro (2018b: Sect. 2.3).

(2) Given a theory whose solution-space is partitioned by a group of symmetries—i.e. solutions in the same same cell are symmetry-related—should we: (a) try to write down a 'quotiented' (also known as: 'reduced') theory whose solutions correspond to the cells of the partition; or (b) resist quotienting the given theory, but take its symmetry-related solutions to be isomorphic? This debate is articulated, with (a) and (b) labelled 'reduction' and 'sophistication' respectively, by Dewar (2015: Sects. 4–6, 2017), see also Caulton (2015: pp. 156–157). Dewar defends sophistication. Broadly speaking, we are sympathetic (De Haro 2018b: Sect. 3.2.3; Butterfield 2018: Sect. 2.3): reduction is not to be undertaken lightly. But the Schema does not commend one or the other: we will encounter these two alternatives in Sect. 2.3.2, and we will see that the choice between a 'reduced' or a 'sophisticated' formalism basically comes down to a choice of representations.

2 Theories, models, model roots and specific structure

We said at the beginning of Sect. 1 that on our account of duality, two dual theories share a common core, which is itself a theory, the bare theory—and that they are isomorphic models of it. So to explain that account, we first need to explain our use of the terms 'theory' and 'model': which, after all, have various (conflicting!) uses. We do this in Sect. 2.1. This leads to: distinguishing within a model, what we will call the *model root* (in most cases: a *model triple*) from the *specific structure* (Sect. 2.2); and giving examples of this distinction (Sect. 2.3). Finally, we introduce notation for the values of physical quantities on states (Sect. 2.4). Then we will be ready to define duality as an isomorphism of model roots: details in Sect. 3.

2.1 'Theory' vs. 'model': interpretation

For the words 'theory' and 'model', the first point to make—the main idiosyncrasy in our usage—arises from the fact that since a duality compares two theories, it involves two "levels": the common core "above" (more general and-or more abstract) and the theories "below" (more specific and-or more concrete). On the other hand, one naturally thinks of a theory as general and-or abstract, "standing above", and "being common to", its more specific and-or concrete models. So in order to discuss duality, the question arises: should we allocate the word 'theory' to the common core, or to each of the two dual theories?

We opt for the former. And since the two dual theories (so-called!) are indeed specific and-or concrete realizations (versions, formulations) of this common core that we have opted to call a 'theory', we also co-opt the word 'model' for the duals themselves. So the overall effect of our jargon is to lift the use of the words 'theory' and 'model', "one level up".

As a simple example to illustrate our usage, consider position-momentum duality in elementary Schrödinger-picture quantum mechanics. Although the position and momentum representations might have been discovered, in a counterfactual history, as different "theories" (since the Schrödinger equation takes very different forms in the position and momentum bases, as do the operators and wave-functions), the discovery of the Fourier transformation reveals that these are in fact two representations of a common core theory: namely, quantum mechanics formulated in the basis-free language of Hilbert space. Thus what in the counterfactual history is thought to be two distinct "theories", is seen to be just two models, i.e. representations, of a common structure, which we call elementary quantum mechanics.

2.1.1 'Model'

So beware! This means that our use of 'model' rejects three common connotations of the word.[3] Namely, the connotations that a model is:

[3] In Sect. 5.2, we will recover the first notion, (i), of a 'model', as a *special case* of our usage. That we here reject this connotation means that our models are not *defined* as in (i). But in Sect. 5.2's specific case, "models" in sense (i) turn out to be models in our sense.

(i) a solution of the theory, or a history of the system concerned (often a trajectory through the state-space), as against 'all solutions'—which are the purview of the theory; and-or
(ii) an approximation(s) to what the theory says (maybe about a specific regime or system the theory is concerned with); and-or
(iii) a part of the empirical world—a hunk of reality!—that thus gives an interpretation (or part of an interpretation) of the theory.

We will indeed later be concerned with the ideas (i) to (iii): (in fact, with (iii) in this Subsection). But they are not part of our definition of 'model'.

Instead, a model is for us a specific realization (or version or formulation) of a theory. That is: it 'models' (verb!) another theory "above", which in general also has other such models. Almost always, it is a representation of the theory above, in the sense of mathematical representation theory—'representation' being another word with confusingly diverse uses. Recall the example, in this Section's preamble, of position versus momentum representations.

And as we have announced: the main case for us of "the theory above" will be the *common core*, the *bare theory*, of two dual theories (our 'models'). (It is usually best to think of the bare theory as uninterpreted, or abstract: though it may be interpreted: cf. Sect. 2.1.3). So we say the two given theories are *models* of the bare theory; and their being *isomorphic* models of the bare theory (isomorphic as regards the structure of the bare theory) is what makes them duals.

2.1.2 'Theory'

As to our use of 'theory', the main thing to say is that it is mainstream: i.e. typical of the literature, especially the literature on the semantic conception of theories as applied to physics. Agreed: recent philosophy of science has emphasised many aspects of scientific endeavour that hardly invoke the notion of *scientific theory*, central though this notion was for discussions by both the positivists and their successors. For example, aspects such as experiment (calibration of instruments etc.), causation (mechanistic explanation etc.), and the social dimensions of knowledge (testimony etc.) have recently been discussed with a strong emphasis on models (in a more usual sense than ours!), rather than theories.

We agree that these aspects of scientific endeavour are important for our philosophical understanding of science (De Haro and De Regt 2018a: Sect. 1.1). But even if these aspects do not need the notion of theory, still the notion may well be useful for other aspects. Indeed, we believe it is indispensable for discussion of symmetries and dualities in physics.[4]

More specifically: we can often think of a physical theory as a triple: a state-space $\mathcal{S}$ (e.g. a Hilbert space in a quantum theory), a set of physical quantities $\mathcal{Q}$ (almost always an algebra), and a dynamics $\mathcal{D}$ that describes how the values of quantities (on states) change over time (and-or over space—so as to accommodate Euclidean theories). As

[4] We also believe it useful, even indispensable, in other discussions. One main one is understanding renormalization—a topic for which, again, there has been scepticism about its usefulness: e.g. Kaiser (2005: pp. 377–387). For a defence of the notion, cf. Butterfield (2014: Sect. IV.1).

we said in Sect. 1, symmetries and dualities will concern the preservation or matching of these values.[5] So we will sometimes write a theory (and of course: a model in our sense) as a triple, e.g. $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$. In any specific case, each of $\mathcal{S}$, $\mathcal{Q}$ and $\mathcal{D}$ will of course have further structure. For example:

(i) For a quantum theory, $\mathcal{S}$ will usually be a Hilbert space, or a set of density matrices; for a classical theory, it will usually be a manifold, or a set of probability distributions;
(ii) For almost any theory, $\mathcal{Q}$ will be an algebra over the real or complex numbers, allowing quantities to be added and multiplied; and
(iii) For almost any theory, the dynamics $\mathcal{D}$ can be understood either in Schrödinger picture, with states changing over time and quantities fixed over time, or in Heisenberg picture (vice versa).

Two clarifications about our treatment of a theory as a triple: the first formal, the second interpretative. First: note that we said: 'we can *often* think of a physical theory as a triple', and '*almost any*' in (ii) and (iii). For we agree that not every theory is presented, or best thought of, in this way. Theories in statistical and quantum physics are often formulated in terms of partition functions and-or path integrals with sources, and related concepts like sets of correlation functions, rather than in terms of states and quantities. And in field theories, the dynamics is often presented as field equations holding at each spacetime point—and so not naturally thought of in terms of the Schrödinger or Heisenberg pictures with their "background time". But in this paper, we can think of theories (and models, in our sense) as such triples: all our morals will carry over to these other ways of formulating theories.[6]

Second: We admit that of course, a physical theory is almost never presented to us as a tidy triple of state-space, quantities and dynamics. Almost always, the theory appears to us messier than that: more vaguely defined and-or more complicated. The triple needs to be extracted from that "mess". Indeed, there are two points here.

(i) The complicated appearance is of course in part due to those aspects such as experiment mentioned above. But this complexity, and the need to allow for such aspects (and to assess them philosophically), does not militate against extracting a triple as a concept useful for e.g. understanding symmetries.
(ii) We make no claim that there is always, or even typically, a unique best definition of the triple. So presenting a theory as a triple usually involves: (a) choices about exactly what to take as the state-space etc.; and even (b) judgment about interpretative and perhaps controversial matters. We will see some examples of this variety, already in Sect. 2.2; and we will see that these choices may affect verdicts of interest, e.g. whether there is an isomorphism, or a duality.

[5] In addition, a theory often comes with stipulated symmetries. We will discuss these in Sects. 4.2 and 5.1.

[6] Of course, one can draw connections between formulations with states, quantities and dynamics—our triple conception—and other types of formulation, like partition functions, path integrals and field equations. For example, a partition function with a source that couples to an operator in the Lagrangian of a quantum field theory is standardly used to calculate, by taking functional derivatives, the correlation functions of that operator in the vacuum state. Cf. Sect. 4.4, and De Haro et al. (2017: pp. 75–76).

2.1.3 Interpretation

This brings us to the interpretation of physical theories. A large subject! But here, we only need to report another part of our overall position, which is again mainstream. Namely, we endorse the endeavour of giving theories (and so: models in our sense) a classical referential semantics. That is: we envisage assigning references in the empirical world to appropriate elements of theories—be they states and quantities (i.e. mathematical objects) for a theory presented as a triple, or words and sentences (i.e. linguistic objects) for a theory presented in a language. This even-handedness between mathematical and linguistic objects—and thus between theories as N-tuples and as sets of sentences—is deliberate: for we endorse recent arguments against the traditional firm distinction between the semantic and syntactic conceptions of theory (Lutz 2017; Glymour 2013; van Fraassen 2014).[7]

This even-handedness is one reason we call our endorsement of referential semantics 'mainstream'. There are also two other reasons: the first is familiar in philosophy of science, and the second is familiar in philosophy of language.

First: recall that accepting such referential semantics is independent, or at least largely independent, of the debate over scientific realism. For that debate is largely about what is the right attitude to our theories, not about their semantic content. Thus the arch empiricist, van Fraassen, explicitly accepts a literal construal of the theoretical claims of—i.e. a referential semantics for—scientific theories (1980: p. 14): as do other influential positions that reject realism, such as Fine's 'natural ontological attitude' (1984: pp. 96–99, 1986: p. 130).

Second: we should recall the moral of Lewis' seminal paper, 'Languages and Language' (1975). Lewis begins by rehearsing a thesis and an antithesis: the task of the paper, and his moral, is to reconcile them in a synthesis—which indeed he accomplishes. Thus the thesis begins by saying that *a language* is an assignment 'of meanings to certain sequences of types of sound or of marks…' (p. 3). As this quote hints: the thesis is advocacy of a referential semantics; (indeed an intensional semantics—cf. below). The antithesis begins by saying that '*language* is a social phenomenon wherein people utter strings of vocal sounds ... and wherein people respond by thought or action to the sounds they observe to have been produced' (ibid). As this quote hints: the antithesis is advocacy of an account of language emphasising people's propositional attitudes (intentions, beliefs, desires etc.): both as what is communicated by language, and as what underpins that communication. Thus Lewis' synthesis is his account of what it is for *a* language L, *à la* the thesis, to be *the* language used by a human population, *à la* the antithesis. The main idea is that this is a matter of the population having conventions (in Lewis' sense) of truthfulness in L and trust in L. Thus he knits the thesis and antithesis together in a detailed way (and replies to various objections). He ends by saying: 'According to the proposal I have presented, the philosophy of language is a single subject. The thesis and antithesis have been the property of different schools; but in fact they are complementary essential ingredients in any adequate account either of languages or of language' (p. 35).

[7] And our phrase 'words and sentences' is to signal that the semantics is compositional in the usual sense: viz. the reference assigned to a string of symbols is a function of the references assigned to the symbols.

To which we say: 'Hear, hear!'. That is: we claim a similar reconciliation between our advocating a referential semantics for scientific theories, and various lines of philosophical work that downplay, or even do not mention, reference or theories. Thus some work on some of the aspects of scientific endeavour, mentioned above, seems sceptical, not just about 'theory' as a useful notion, but also about reference. For example: much philosophical writing about experiment (calibration of instruments etc.) emphasizes non-linguistic skills, practices and norms; and much philosophical writing about the social dimensions of knowledge emphasizes the wider and practical world, e.g. the functioning of scientific communities and institutions in e.g. the maintenance of norms, such as accreditation, etc. Such emphases are entirely appropriate, say we. Study of experiment should *of course* emphasize non-linguistic skills; and so on. But such emphases in no way militate against developing a referential semantics for scientific language, and so for scientific theories. In short, we think Lewis' synthesis gives a valuable, because irenic, perspective on this work's relation to referential semantics.[8]

So much by way of defending ourselves as having mainstream views about interpretation. The upshot is that we envisage a referential semantics using interpretation maps, i. These are maps on states or quantities ($s \in \mathcal{S}$ etc.), or on linguistic items, assigning as values (outputs of the map) parts of the empirical world (hunks of reality!). But they are, in general, partial maps; i.e. for some arguments, the map yields no value (output). For some details and examples, see De Haro (2018b).[9]

In the next Section we will contrast *internal* and *external* interpretations, both of which are interpretations in the above sense. Roughly speaking, an internal interpretation only interprets that "part" of the model that is "the homomorphic copy" of the theory, while an external interpretation can interpret all of the model.

2.2 Model roots, specific structure, and their interpretations

As hinted at the end of the previous Section, we need some more jargon and notation about the relation between a bare theory and its models that will allow us to distinguish the parts or aspects of models that are "shadows" of corresponding parts or aspects of the bare theory, from those that are not.

[8] More controversially: we think Lewis' synthesis gives a valuable, because deflating, perspective on the burgeoning literature about scientific representation; viz. along the lines Callender and Cohen's claim that we should analyse representation in science in terms drawn from philosophy of language and mind (2006, especially Sect. 3).

[9] In fact, we endorse a specific programme within the general enterprise of referential semantics. Namely: intensional semantics in the sense developed by Carnap, Montague and Lewis in which:

(i) the notion of 'linguistic meaning' is taken to be ambiguous between what Frege called 'sense' and what he called 'reference', here called 'intension' and 'extension' respectively; and
(ii) intension is modelled as a function taking linguistic items—and for us, states or quantities—to their reference, relative to a possible world.

But in this paper, we will not need the details of this view. A standard exposition is Lewis (1970); besides, pp. 16–17 of his (1975) give a fine sketch, including a defence of semantics adverting to possible worlds. Our own endorsement is in: De Haro and Butterfield (2018: Sect. 2.3), Butterfield (2018: Sect. 3.1), De Haro and De Regt (2018a).

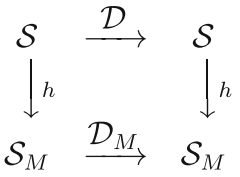

Fig. 1 The model's (Schrödinger) dynamics implements that of the bare theory

The jargon is clearest for the common case, which we will focus on throughout this paper: when the realization of the bare theory proceeds by (mathematical) representation, and the bare theory is a triple comprising a state-space $\mathcal{S}$, a set of quantities $\mathcal{Q}$ and a dynamics $\mathcal{D}$: (cf. Sect. 2.1.2 and footnote 6). Although the bare theory may be interpreted (like its models, the dual theories, usually are), adding the adjective 'bare' signals helpfully that the theory is "above" and can be / is often uninterpreted.[10]

In this common case: by the very definition of 'representation', the model gives a *homomorphic copy* of the bare theory (irrespective of there being a duality). That is: there are appropriate structure-preserving maps from the states, quantities and dynamics of the bare theory to the model's homomorphic copy. To be precise: there is a pair of structure-preserving maps—from states in the bare theory to states in the model, and from quantities in the bare theory to quantities in the model. And there is a meshing condition on the model's dynamics that makes it implement that of the bare theory. The details are as follows: though we can mostly take them in our stride, and just say 'homomorphic copy'.[11]

We can write the bare theory as a triple $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$; and similarly its model $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$, where the subscripts signal that the state-space etc. are different from that of the bare theory. If we think of the dynamics in Schrödinger style as a map on the state-space (more details in Sect. 2.4), and write the homomorphism from $\mathcal{S}$ to $\mathcal{S}_M$ as h, then the meshing condition on the model's dynamics will be the commuting diagram in Fig. 1.

Once we have the distinction of levels, i.e. a bare theory represented by a model, there is an important distinction to be made within the model. (This is important irrespective of there being a duality.) Namely between:

(i) The parts and aspects of the model which together express its realizing the bare theory;
(ii) The parts and aspects which do not express the realization.

We will call (i) the **model root**. And in the common case where the bare theory, and so also the model, is a triple of states, quantities and dynamics, we will call (i) the **model triple**. And we will call (ii) the model's **specific structure**. We can think of it

[10] We take the point, from Read and Møller-Nielsen (2018: Sect. 5.3), that not every common core need be so rich as to lead to a theory (in the present case, not every common core needs to allow being written as a triple). But we will simply restrict our attention to cases in which the common core *is* as rich as a theory, which we submit are the most interesting cases—and indeed they seem to be the most common cases in the literature on dualities. See also De Haro (2018b: Sect. 2.1.1; footnote 30).

[11] So the relation of representation between a bare theory and a model of it will involve not just one map as in, say, group representation theory (the homomorphism from the abstract group to e.g. a set of matrices), but at least two maps. We will see in Sect. 3.1 that these two maps are related to each other, because states and quantities are dual (in mathematicians' sense!) to one another.

as the 'ingredients' or 'building blocks' by which the representation of the bare theory, i.e. (i), is built.

There is a correlative distinction between two kinds of interpretation. Recall from Sect. 2.1.3 that an interpretation is given by interpretation maps, i.e. functions (in general, partial functions) mapping items in the theory to items in the world. Thus we say that an **internal** interpretation is one that only interprets the model root: it maps all of and only the model root to items in the world, regardless of the specific structure. On the other hand, an **external** interpretation also maps (some or all of) the specific structure to items in the world.

Of course, ingredients are present in the cooked dish, and building blocks are present in the built house. Similarly here: often, an item of specific structure is in the model root, though (by definition) it is not part of the representation of the bare theory. And in such a case an internal interpretation does not interpret the item of specific structure. (Sect. 2.3 will give examples.)

Thus the distinction between specific structure ('building blocks'), and model root ('what gets built') is physically significant, in that it constrains interpretation. It is formal in that it can be stated without giving an interpretation: but it has consequences for interpretation. For more discussion, cf. De Haro (2016a: Sect. 1.1.2, 2018a: Sect. 2.2.3), De Haro and Butterfield (2018: Sect. 3.2.2), and De Haro (2018b: Sect. 2.1.2).

But we stress that even within a given model, the distinction is not 'God-given'. It is relative to how exactly we define the bare theory, and thereby the homomorphism from it to the model. And for a physical theory, as usually presented to us informally and even vaguely, there need be no best or most natural way to make this exact definition. For recall comment (ii) at the end of Sect. 2.1.2: how exactly to present a theory as a triple of states, quantities and dynamics is a matter of choice and even judgment.

We will see this flexibility in play later, in Sect. 2.3, where we will have a choice to define the model root either as a single representation of a structure (with a further choice to include or not to include a choice of basis in the model root) or as an equivalence class of representations.

We will also see it in connection with dualities, in Sect. 2.3.2. For as we announced: we say a duality is an isomorphism of models of a bare theory; (details in Sect. 3.1). But this means: an isomorphism with respect to the structure of the bare theory—which is the structure that the models represent. Therefore the isomorphism that is the duality is an isomorphism of model roots. And in the common case of triples of states, quantities and dynamics: it is an isomorphism of model triples.[12] So the flexibility about the definition of a bare theory, and so about what a model must represent, leads to flexibility about exactly what the duality mapping is, i.e. what is the isomorphism between model roots.

It is worth having a notation distinguishing between the model root (which is usually a model triple of states etc.) and the specific structure. So given a model M, we now write m for its model root. This is usually a model triple, which we now write as

[12] A clarification: This is not to say that the specific structure (the 'building blocks') is always 'invisible' to the other side of the duality, i.e. that no part of the specific structure is mapped across by a duality to the other model. Often, some part of the specific structure is mapped across. Indeed, that is unsurprising. For the model root is built from specific structure: so one expects that in order to map the model roots, one into the other, the duality must map at least part of the specific structure, one into the other.

$\langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$. Again, the subscripts signal that the state-space etc. are different from those of the bare theory. We also write $\bar{M}$ for the specific structure. So we write a model M of a bare theory T as

$$M =: \langle m; \bar{M} \rangle = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M; \bar{M} \rangle \,, \tag{1}$$

where the semi-colon in the defined angle-bracket signals the contrast between root and specific structure, and the last equation just expresses the usual case of the root being a triple.

But *beware*: one should not think of M as just the ordered pair of independently given m and $\bar{M}$. For m is built by using the specific structure $\bar{M}$, and so it is not given independently of $\bar{M}$. Rather, m encodes M's representing the bare theory T. So one might well write T_M instead of m, since having a subscript M on the right hand side of the equation $M = \langle T_M; \bar{M} \rangle$ signals that M is not an ordered pair of two independently given items. In other words: the notation T_M emphasises that the model triple: (i) is a representation of T, (ii) is built from M's specific structure viz. $\bar{M}$, and (iii) is itself a theory (hence the letter 'T').

Here is an illustration of this notation in use. As announced: we say that a duality is an isomorphism of model roots, with respect to the structure of the bare theory. So if M_1, M_2 are models of a bare theory T, their being duals means: $m_1 \cong m_2$. And in the usual case of model triples, i.e. $m_i = \langle \mathcal{S}_{M_i}, \mathcal{Q}_{M_i}, \mathcal{D}_{M_i} \rangle$, $i = 1, 2$: this isomorphism of roots will be a matter of two appropriately meshing isomorphisms, one between the state-spaces $\mathcal{S}_{M_i}$ and one between the quantity algebras $\mathcal{Q}_{M_i}$. Details in Sect. 3.1.

2.3 Examples: matrix representations and Galilean transformations

In this Section, we will illustrate our notions of theory and model, and the contrast of model root vs. specific structure. Our first example (Sect. 2.3.1) is from matrix representations; our second example (Sect. 2.3.2) is about Galilean transformations in Newtonian mechanics. The latter example will also illustrate our notion of interpretation from Sect. 2.1.3, especially our internal versus external contrast from Sect. 2.2.

2.3.1 Matrix representations

Perhaps the simplest illustration of these notions, model root and specific structure, comes from defining the bare theory to be just an abstract group G; and as usual, taking realizations to be representations.[13] So let us consider matrix representations of G. More specifically: we consider for a finite group G, a set $\{M_i\}$ of $n \times n$ complex matrices with non-zero determinant (i runs from 1 to the order of G). That is: $M_i \in \mathrm{GL}(n, \mathbb{C})$. In such a representation, the choices of the size n of the matrices and of the ground-field ($\mathbb{R}$ or $\mathbb{C}$) are of course parts of specific structure: for these are building blocks by which we build the homomorphic copy of G.

[13] Strictly speaking, a bare theory must have an appropriate set of maps to the real numbers, to express the values of quantities, and even a dynamics. But for the sake of illustration, we here ignore these maps. Anyway, Sect. 2.3.2 will sketch an illustration from physics.

But even in this simple illustration, we can see further options about how exactly to define model root and specific structure. These options relate to the fact that the unitary group $\mathrm{U}(n)$ acts on $\mathrm{GL}(n, \mathbb{C})$, by $M \mapsto UMU^{-1} \equiv UMU^{\dagger}$. Each $U \in \mathrm{U}(n)$ sends a representation $\{M_i\}$ to another representation, called 'equivalent'; and representation theory then of course concentrates on equivalence classes of representations, characterizing them in terms of their invariants, especially characters. The spirit of the enterprise is that a unitary change of basis has no mathematical significance. This means there are two main options, (A) and (B) below, about how to define the model root, and thereby, also the specific structure—and there will be further choices.[14]

(A) *Model root as representation* We can say that a single representation $\{M_i\}$ is the model root. Indeed: multiplication of the matrices M_i realizes a homomorphic copy of G's multiplication. And it is no objection that items of specific structure—the size of the matrices n, and the choice of the ground-field—are 'in' the matrices M_i. For as we said: that an item appears in a model root does not prevent it appearing in the specific structure. After all, the model root is *built from* the specific structure.

On this option, two equivalent representations of G, i.e. a set of matrices $\{M_i\}$ and another set $\{U\ M_i\ U^{-1}\}$, for some fixed $U \in \mathrm{U}(n)$, will be isomorphic, as homomorphic copies of G's multiplication. So on our account of duality, they are duals. (As we will discuss in Sect. 3: this reflects our definition of duality being logically weak i.e. having many instances.)

But there is also a further choice. For the fact that the matrices realize group multiplication is independent of their acting as linear operators on the vector space $\mathbb{C}^n$. That is: although we always think of a matrix as representing (that word again!) a linear operator, it only does so relative to a choice of basis—and the representation we began with, viz. $h : G \to \{M_i\}$, makes *no* such choice. But if we wish, we can adjoin such a choice to our model. That is: we can stipulate that the matrix representation $\{M_i\}$ also has, as part of its specific structure, some specific basis $\mathbf{e}_k$ $(k = 1, \ldots, n)$ of $\mathbb{C}^n$.

Once such a choice is made, the basis vectors $\mathbf{e}_k$ (and thereby all vectors $\mathbf{v} = \sum_{k=1}^{n} v_k\, \mathbf{e}_k$) are of course acted upon—sent to another basis—by the similarity transformations U that act on the matrices M_i. That is: $\mathbf{e}_k \mapsto \sum_{l=1}^{n} U_{kl}\, \mathbf{e}_l$ when $M_i \mapsto U\ M_i\ U^{-1}$. But the idea of the stipulation is that the first-chosen basis $\mathbf{e}_k$ labels the representation: it fixes the interpretation of the $\{M_i\}$ as linear operators. It is just that this labelling basis maps across to the labelling bases of equivalent representations. This illustrates the idea in footnote 12 that—to return to our jargon—specific structure can map across a duality. In this example, the duality maps a labelling basis to another basis: $\{\mathbf{e}_k\} \mapsto \{\sum_{l=1}^{n} U_{kl}\, \mathbf{e}_l\}$ when $M_i \mapsto U\ M_i\ U^{-1}$.

(B) *Model root as equivalence class* We can say that an equivalence class of representations—the entire orbit of a given $\{M_i\}$ under the action of $\mathrm{U}(n)$—is the model root. For indeed: equivalent representations realize the very same homomorphic copy of G's multiplication. Again, the size of the matrices n, and the choice of the ground-field, are 'in' the model root (since they are preserved by equivalence). And

[14] Roughly speaking, (B) takes roots to be equivalence classes of what (A) takes them to be. One might question whether (A)'s "concrete" matrices are really more "basic" than their equivalence classes as in (B). But exposition of the issues is much clearer if we keep the (A) vs. (B) contrast.

again, it is no objection that an item in the model root is also in the specific structure: since the model root is built from the specific structure.

On option (B), there are still duals, on our account of duality—despite the option's having "quotiented" to a more abstract notion of model than that of option (A). For we can vary n; and-or we can vary the ground-field. That is: two different choices of n can provide two model roots—the equivalence class of a representation $\{M_i\} : M_i \in \mathrm{GL}(n_1, \mathbb{C})$, and the equivalence class of a representation $\{M_j\} : M_j \in \mathrm{GL}(n_2, \mathbb{C})$—that instantiate the same homomorphic copy of G. And similarly, we can vary the ground-field, and yet instantiate the same homomorphic copy of G. And similarly, we can consider non-linear/non-matrix realizations/representations of G.

And as in option (A), there is the further choice—again because the fact that matrices realize group multiplication is independent of their acting as linear operators on the vector space $\mathbb{C}^n$. That is: if we wish, we can adjoin to a model root—an entire equivalence class of representations—the choice of some specific basis $\mathbf{e}_k$ $(k = 1, \ldots, n)$ of $\mathbb{C}^n$, which we can then take to be specific structure. Of course, the only natural way to do this is to attach the basis to some arbitrary element of the class, i.e. one set of matrices $\{M_i\}$, and then transport the basis around to the other elements of the equivalence class by the action of $\mathrm{U}(n)$.

2.3.2 Galilean transformations

Newtonian mechanics provides a simple illustration of the notions of model root, specific structure, and indeed duality. And since it is an example from physics rather than pure mathematics, we also get an illustration of Sect. 2.2's distinction between internal and external interpretations.

The idea is as follows. The bare theory T is Newtonian mechanics, of say N gravitating point-particles, set in a Galilean (neo-Newtonian) spacetime: i.e. in a spacetime manifold that is 'globally like $\mathbb{R}^4$', with Euclidean geometry in its instantaneous time-slices, and a flat 4-dimensional connection, but no preferred absolute rest. This bare theory is modelled (in our sense: i.e. realized, represented) by formulations of Newtonian mechanics of N gravitating particles, set in a *Newtonian* spacetime, i.e. in a spacetime that is 'globally like $\mathbb{R}^4$' but that *does* have an absolute rest.

Famously (notoriously!), the difference in such formulations' specifications of absolute rest is not experimentally detectable, since specifications that are each boosted with respect to the other specify the same flat 4-dimensional connection, and a boost maps a solution of the equations of motion to another solution. Or as it is usually put, in more physical terms: no experiment in Newtonian mechanics can distinguish one specification of rest from the others, because the theory is invariant under boosts ('has boosts as a symmetry'). Hence, of course, the debate between Newton and Leibniz, as articulated in the Leibniz-Clarke correspondence, and with its long legacy down to the present day (e.g. Earman 1989).

So in this example, it is natural to say that the specific structure of each model includes its specification of absolute rest. Using this specification, the model defines a flat 4-dimensional connection—viz. the same connection as is defined by the other models—and thereby builds a homomorphic copy of T.

We can make this example simpler and precise, and yet still a worthwhile illustration, by taking the bare theory T to be just the abstract Galilean group Gal(3). This is a 10-dimensional Lie group, whose generators are three spatial rotations, four (space and time) translations, and three boosts. That is: its generators are usually thus described, by way of justifying their commutation relations. But of course the abstract group can be defined by the commutation relations, free of a physical interpretation. Thinking of Gal(3) like this will yield a clear analogy with Sect. 2.3.1's matrix representations of an abstract group G, and with that Section's option (A), i.e. *Model root as representation*.

Gal(3) is usually presented in its fundamental representation. Namely, as a concrete group of transformations on (bijections of) $\mathbb{R}^4$, written in terms of the standard coordinates $(\mathbf{x}, t) \in \mathbb{R}^4$, with $g \in$ Gal(3) represented as the function

$$g(\mathbf{x}, t) := (R \cdot \mathbf{x} + \mathbf{v}_0\, t + \mathbf{r}_0, t + t_0)\ , \tag{2}$$

where R is a 3×3 spatial rotation matrix, $\mathbf{v}_0$ is the velocity boost, $\mathbf{r}_0$ is the spatial translation vector, and t_0 is the time translation. This fundamental representation can also be expressed in a coordinate-free way as an action on the affine space of $\mathbb{R}^4$, i.e. on Euclidean 4-space. But we will not need the details of affine spaces (cf. e.g. Auslander and MacKenzie 1963, Chapter 1).

But just as in Sect. 2.3.1's option (A), where a model root was a matrix representation of a group G, we could adjoin a choice of a basis $\mathbf{e}_k$ $(k = 1, \ldots, n)$ of $\mathbb{C}^n$ as specific structure: so also here, we can adjoin a choice of an inertial coordinate system as specific structure, and we can take this to give the model's specification of absolute rest.[15] The standard coordinate system defined by the components of $\mathbb{R}^4$ itself is then just one choice among many, determining one specification of absolute rest among many. Natural though we find it for writing down the fundamental action of Gal(3), as we did in Eq. (2), the action can of course be written down in any inertial coordinate system. And any such system can be taken to give a model's specification of absolute rest.

And just as in (A) of Sect. 2.3.1, each matrix M_i mapped the adjoined basis $\mathbf{e}_k$ to another basis: so also here, each Galilean boost maps an adjoined choice of absolute rest, represented mathematically by an inertial coordinate system, into another such choice, i.e. another inertial coordinate system.[16]

But there is also, so far, a disanalogy with (A) of Sect. 2.3.1. For so far, we have just one model root, encapsulated in Eq. (2), and various choices of specific structure; whereas (A) of Sect. 2.3.1 had many different model roots—many different matrix representations $\{M_i\}$ of the abstract group G. Correlatively, our fundamental representation of Gal(3) is faithful, i.e. has trivial kernel; while Sect. 2.3.1's $\{M_i\}$ were in general not faithful.

But there are other representations of Gal(3). Indeed, there is a *matrix* representation that is isomorphic, as a model root, to what we have. So it is faithful—and the isomorphism of model roots is, on our account, a duality. Namely, we use

[15] We say 'give'—meaning 'determine'—rather than 'be', simply because a coordinate system includes choices of spatial and temporal origins and units, and of an orientation of spatial axes, as well as the absolute rest, i.e. the timelike congruence of inertial worldlines.

[16] Of course, Galilean transformations that are not boosts keep fixed the choice of absolute rest: cf. footnote 15.

5×5 matrices $G \in GL(5, \mathbb{R})$ (G for 'Galileo' not 'group'!), that act on vectors $(x_A) := (\mathbf{x}, t, 1) \in \mathbb{R}^4 \times \{1\} \subset \mathbb{R}^5$, as follows (cf. Bargmann 1954: pp. 38–41 and Holm 2011: p 10):

$$\tilde{x}_A = \sum_{B=1}^{5} G_{AB}\, x_B$$
$$(G_{AB})_{A,B=1}^{5} := \begin{pmatrix} R & \mathbf{v_0} & \mathbf{r}_0 \\ \mathbf{0}_{1\times 3} & 1 & t_0 \\ \mathbf{0}_{1\times 3} & 0 & 1 \end{pmatrix}. \tag{3}$$

With this representation, a specification of absolute rest is given by a specific choice of vector $x \in \mathbb{R}^4 \times \{1\}$. Namely, $(x_A) := (\mathbf{x}, t, 1) \in \mathbb{R}^4 \times \{1\}$ specifies rest to be the timelike congruence of inertial worldlines parallel to the inertial worldline passing through both the spatiotemporal origin $(\mathbf{0}, 0)$ and the point $(\mathbf{x}, t)$. So here, it is a single vector that gets adjoined as specific structure; (not a whole basis, as in Sect. 2.3.1, and not a whole inertial coordinate system, as above). And each Galilean boost maps an adjoined choice of absolute rest, given by a vector $x \in \mathbb{R}^4 \times \{1\}$, into another such choice, i.e. another vector.

So much by way of how Galilean transformations' illustration of the notions of model root and specific structure is analogous with Sect. 2.3.1's matrices. We end with two comments, [1] and [2], about the physical interpretation of this example (comments which thereby have no analogues about those matrices). For simplicity and brevity, we will restrict both comments to the simple "vacuum" scenario which we have concentrated on: i.e. $\mathbb{R}^4$ as a description of either neo-Newtonian or Newtonian spacetime, without regard to the N gravitating point-particles we mentioned at the start of this Section. Thus recall that we concentrated on taking the bare theory T to be just the abstract Galilean group Gal(3), and considered its action on $\mathbb{R}^4$. But this is only for simplicity: these two comments carry over to the non-vacuum scenario, where there are particles.[17]

[1] *Agreement with usual verdicts* The first comment looks ahead to Sect. 4.4's discussion of symmetries in a spacetime theory, i.e. a theory that postulates a spacetime

[17] It is just that it would take too long to spell out the non-vacuum scenario. To glimpse why, we briefly note some of the issues one confronts. Obviously, one must consider the particles' state-space: which one would build from their configuration space (the "qs"), by adding either velocities ("$\dot{q}$s": defining velocity phase space, in the Lagrangian framework) or canonical momenta ("ps": defining phase space, in the Hamiltonian framework). At first sight, the N particles' configuration space is "just" $\mathbb{R}^{3N}$. But there are subtleties to be dealt with. Indeed: not only the topics mentioned above, of passing to the affine space so as to "rub out" the preferred origin, and whether to have an absolute rest; but also whether to excise collision points, i.e. whether to forbid point-particles to be in the very same place. Assuming these subtleties are dealt with, and the Lagrangian or Hamiltonian state-space is constructed, one would then consider the action of the Euclidean group on this state-space, lifted from its action on $\mathbb{R}^3$. Again, there are subtleties about this lifted action; and to treat boosts and so represent the Galilean group, one needs to "add a time axis", defining what is often called 'extended (velocity) phase space'. For a philosophical introduction to all these subtleties, cf. e.g. Belot (2000: Sects. 3, 4) and Butterfield (2006: Sect. 2.3).
These two comments also bear on the two debates about interpreting symmetries, which we set aside at the end of Sect. 1. Indeed the authors cited take Newtonian gravitation as a main example. But again, these debates are orthogonal to most of this paper's issues; so that it would take too long to spell out, beyond the references we gave there, exactly what these comments imply for them.

with certain chrono-geometric structures like metrics and connection. (Such theories of course also postulate matter and radiation, particles and fields, in the spacetime; but as just announced, we are setting that aside.)

We will see in Sect. 4.4 that in a spacetime theory, a symmetry is usually defined as a map on the spacetime that (once its domain of definition is extended in the natural way to include chrono-geometric structures and matter fields): (i) fixes, i.e. does not alter, the chrono-geometric structures, and (ii) maps a matter-field solution of the equations of motion to another solution. (We will also see how this relates to the more basic and general notion of symmetry we will use from the start of Sect. 4.) Accordingly, boosts are a symmetry of neo-Newtonian spacetime: for a boost preserves the chrono-geometric structures, i.e. the spatial and temporal metrics and the flat affine connection, and maps solutions to solutions. But boosts are not a symmetry of Newtonian spacetime (i.e. a spacetime that is globably like $\mathbb{R}^4$, but that has a specification of absolute rest). For a boost does not fix a specification of rest.[18] These points are, in effect, the modern mathematical expression of the famous (notorious!) point with which this Section began: that no mechanical experiment can discern which is the putatively correct standard of rest.

Our discussion above, and our notions of model root, duality etc., accords with this. In particular, just as in Sect. 2.3.1: an item of specific structure can be "in" the model root, which is, after all, built with specific structure; and so an isomorphism of model roots (on our account, a duality) can map specific structure from one model to another. And this is what Galilean boosts do. In our jargon: they define an isomorphism of model roots that maps one model's specification of absolute rest into another's. Think for example of how the vector $(x_A) := (\mathbf{x}, t, 1) \in \mathbb{R}^4 \times \{1\}$—which specifies rest as parallelism to the inertial worldline through both the origin $(\mathbf{0}, 0)$ and the point $(\mathbf{x}, t)$—is mapped by G_{AB} of Eq. (3) to a vector specifying rest as parallelism to the inertial worldline through both the origin $(\mathbf{0}, 0)$ and the point $(R \cdot \mathbf{x} + \mathbf{v}_0\, t + \mathbf{r}_0, t + t_0)$. [For more detail about how the duality maps in this example are defined by boosts, cf. Butterfield (2018: Sect. 4.1).]

But mapping one model's specification of absolute rest into another's is not the same as *fixing* the given specification, i.e. not the same as a symmetry in spacetime theories' usual sense. Thus this example illustrates how a bare theory can have a symmetry, viz. boosts, that (some or even all) its models lack. This will later be a main theme (Sects. 5.1 and 5.2).

This example also has a philosophical moral that is not about symmetry. Namely: duality does *not* imply physical equivalence. Two theories can be duals—in our jargon: models with isomorphic model roots—without their making the very same claims about the world. They can even contradict one another about the world, as do two rival specifications of what is absolute rest. This leads in to the next comment.

[2] *Internal and external interpretations* The example also illustrates Sect. 2.2's distinction between internal and external interpretations. As usual, interpretative issues are underdetermined by formal theory: witness the moral just stated at the end of [1]. However, the example clearly allows us to formulate the disagreement between Newton

[18] On the other hand: spatial rotations and spatiotemporal translations are symmetries of both spacetimes.

(Clarke) and Leibniz—the question whether absolute rest is physically real—in terms of the internal versus external contrast.

For recall that an internal interpretation interprets only the model root, but not the specific structure. More precisely, we define this as meaning that specific structure which is "in" a model root as a building block, does not get interpreted. Therefore models that are isomorphic, i.e. have isomorphic roots, and so differ only in their specific structure, must receive the same interpretation.[19] Thus in our example: an internal interpretation of a model simply does not interpret the specification of rest (whether it is given by an inertial coordinate system considered as rest, as for Eq. (2), or by a vector $x \in \mathbb{R}^4 \times \{1\}$, as for Eq. (3)). In short: the specification of rest is not part of what is physical. Thus an internal interpretation articulates Leibniz' relationist views.

On the other hand, an external interpretation (by definition) does interpret (some or all of) the specific structure. Thus an external interpretation can take any two dual models—any isomorphic model roots with different specifications of rest, x and $\tilde{x}$ say—to have distinct interpretations. This kind of external interpretation thus articulates the Newton-Clarke view: in short, that giving all material bodies the same boost makes a physical difference.

2.4 Values of quantities on states

Before formally defining duality as isomorphism, we need notation for treating states, quantities and dynamics. Suppose we are given a set of states $\mathcal{S}$, a set of quantities $\mathcal{Q}$ and a dynamics $\mathcal{D}$: $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$. (As stressed in Sect. 1: in any specific case, $\mathcal{S}$, $\mathcal{Q}$ and $\mathcal{D}$ will each have a lot of structure beyond being sets—but we will not need these details in what follows. And as admitted in Sect. 2.1.2: not all theories are presented, or best thought of, as such triples—but what we say will carry over to other formulations using e.g. partition functions; cf. footnote 6.)

Then we write the value of quantity Q in state s as

$$\langle Q, s \rangle \tag{4}$$

It is these values that are to be preserved, or suitably matched, by the duality, i.e. by the isomorphism of model triples: cf. Eq. (9) below. And in subsequent Sections' discussion of symmetry, it is these values that are to be preserved by a symmetry map.

For classical physics, one naturally represents (that word again!) a quantity as a real-valued function on states: $Q : s \mapsto Q(s)$. Given such a function representing the quantity, $\langle Q, s \rangle := Q(s) \in \mathbb{R}$ is the system's possessed or intrinsic value, in state s, of the quantity Q. Similarly for quantum physics: one naturally represents quantities as linear operators on a Hilbert space of states, so that $\langle Q, s \rangle := \langle s | \hat{Q} | s \rangle \in \mathbb{R}$ is the system's Born-rule expectation value of the quantity. (But in fact, for quantum

[19] Recall that Newtonian gravitation contains unphysical singularities when two massive point particles coincide, as they can do after a finite time (see footnote 17). Thus internal interpretations of isomorphic model roots of Newtonian gravitation must either give the same "interpretation" to the unphysical singularities (usually: signaling a limitation of the applicability of the theory), or one must deal with the singularities in some other way. See the discussion in De Haro (2018b: Sects. 2.3.1, 2.3.2(C)).

dualities, the duality often preserves off-diagonal matrix elements $\langle s_1|\hat{Q}|s_2\rangle \in \mathbb{C}$: cf. below.)

As to dynamics, here assumed deterministic: This can be usually presented in two ways, for which we adopt the quantum terminology, viz. the 'Schrödinger' and 'Heisenberg' pictures; (though the ideas occur equally in classical physics). We will adopt for simplicity the Schrödinger picture. So we say: $d_{\mathcal{S}}$ is an action of the real line $\mathbb{R}$ representing time on $\mathcal{S}$.

There is an equivalent Heisenberg picture of dynamics with D_H, an action of $\mathbb{R}$ representing time on $\mathcal{Q}$. The pictures are related by, in an obvious notation:

$$\begin{aligned} d_{\mathcal{S}} : \mathbb{R} \times \mathcal{S} \ni (t, s) \mapsto d_{\mathcal{S}}(t, s) =: s(t) \in \mathcal{S} \text{ iff } D_H \\ : \mathbb{R} \times \mathcal{Q} \ni (t, Q) \mapsto D_H(t, Q) =: Q(t) \in \mathcal{Q} \end{aligned} \tag{5}$$

where for all $s \in \mathcal{S}$ considered as the initial state, and all quantities $Q \in \mathcal{Q}$, the values of physical quantities at the later time t agree in the two pictures:

$$\langle Q, s(t)\rangle = \langle Q(t), s\rangle \ . \tag{6}$$

3 The schema: duality as isomorphism of model roots

In this Section, we summarize our account of duality. This account has been developed mainly by De Haro (2016a: Sect. 1, 2016b: Sect. 1), but also more fully by us together (2018a: Sects. 2, 3). As a mnemonic, we label this account a *Schema*. We first define duality (Sect. 3.1); then we give a prospectus for the following Sections (Sect. 3.2). It will be clear that our Schema is logically weak, so that there are countless examples, including elementary ones: a topic taken up in footnote 2's references.

3.1 Duality defined

We can now present our Schema for duality as an isomorphism between model roots (model triples). Let M_1, M_2 be two models, with model roots m_1 and m_2 and specific structure $\bar{M}_1$ and $\bar{M}_2$; so that, with the notation Eq. (1), we have: $M_1 = \langle m_1; \bar{M}_1\rangle$ and $M_2 = \langle m_2; \bar{M}_2\rangle$. We can suppose that M_1, M_2 are both models of a bare theory T. Then we say that M_1 and M_2 are dual if their model roots are isomorphic, i.e. if $m_1 \cong m_2$.

More specifically, if the model roots are triples $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1}\rangle$ and $m_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2}\rangle$, then to say that the model triples m_1, m_2 are isomorphic is to say, in short, that: there are isomorphisms between their respective state-spaces and sets of quantities, that

(i) make values match, and
(ii) are equivariant for the two triples' dynamics (in the Schrödinger and Heisenberg pictures, respectively).[20]

[20] Alternatively, in other formalisms (cf. footnote 6), the dynamics in one model is mapped into the dynamics in the other. For an example, where the dynamics is a set of covariant Euler-Lagrange equations, see De Haro (2018b: Sect. 3).

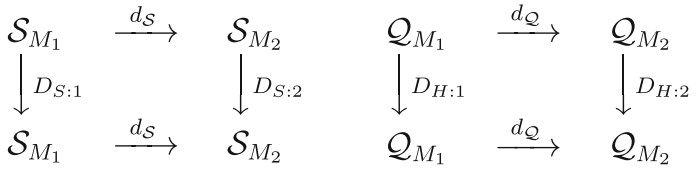

Fig. 2 Equivariance of duality and dynamics, for states and quantities

Thus these maps are our rendering of the correspondence between duals: of, in physics jargon, 'the dictionary'.

Thus we say: A **duality** between $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$ and $m_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2} \rangle$ requires:

(1) An isomorphism between the state-spaces (almost always: Hilbert spaces, or for classical theories, manifolds):

$$d_{\mathcal{S}} : \mathcal{S}_{M_1} \to \mathcal{S}_{M_2} \quad \text{using } d \text{ for 'duality' ; and} \tag{7}$$

(2) An isomorphism between the sets (almost always: algebras) of quantities

$$d_{\mathcal{Q}} : \mathcal{Q}_{M_1} \to \mathcal{Q}_{M_2} \quad \text{using } d \text{ for 'duality' ;} \tag{8}$$

such that: (i) the values of quantities match:

$$\langle Q_1, s_1 \rangle_1 = \langle d_{\mathcal{Q}}(Q_1), d_{\mathcal{S}}(s_1) \rangle_2 \ , \quad \forall Q_1 \in \mathcal{Q}_{M_1} \ , \ s_1 \in \mathcal{S}_{M_1}. \tag{9}$$

and: (ii) $d_{\mathcal{S}}$ is equivariant for the two triples' dynamics, $D_{S:1}, D_{S:2}$, in the Schrödinger picture; and $d_{\mathcal{Q}}$ is equivariant for the two triples' dynamics, $D_{H:1}, D_{H:2}$, in the Heisenberg picture: see Fig. 2.

Eq. (9) appears to favour m_1 over m_2; but in fact does not, thanks to the maps d being bijections.

This definition of duality can be simplified, since the requirement that the values of quantities match, Eq. (9), implies relations between the duality maps $d_{\mathcal{S}}$ and $d_{\mathcal{Q}}$. Thus in the quantum case, the duality maps are related by: $d_{\mathcal{Q}}(Q) = d_{\mathcal{S}} \circ Q \circ d_{\mathcal{S}}^{-1}$ (where $d_{\mathcal{S}}$ is constrained to be unitary).[21] Thus duality comes down to a *single* duality map on states, $d_{\mathcal{S}}$, together with appropriate equivariance conditions on the quantities and the dynamics.

Similarly in the classical case: though representing quantities as real-valued functions on the state-space, rather than as maps on the state-space, means that the relation

[21] The proof is as follows. As mentioned in Sect. 2.4, Eq. (9) generalizes to matrix elements between arbitrary vector states s, s' in the state-space, which in the quantum case is a Hilbert space. Namely, the duality maps must satisfy: $\langle s' | Q | s \rangle_1 = \langle d_{\mathcal{S}}(s') | d_{\mathcal{Q}}(Q) | d_{\mathcal{S}}(s) \rangle_2$. Notice that we can rewrite this purely as an identity in model m_1: $\langle s' |\, Q \,| s \rangle_1 = \langle s' |\, d_{\mathcal{S}}^{\dagger} \circ d_{\mathcal{Q}}(Q) \circ d_{\mathcal{S}} \,| s \rangle_1$. Since the state-space is the whole Hilbert space, $\mathcal{S}_1 = \mathcal{H}_1$, this equation is valid for all $s, s' \in \mathcal{H}_1$ (it is also valid for all $Q \in \mathcal{Q}$); so it implies that the *operators* on the two sides must be equal: $d_{\mathcal{S}}^{\dagger} \circ d_{\mathcal{Q}}(Q) \circ d_{\mathcal{S}} = Q \ \Rightarrow \ d_{\mathcal{Q}}(Q) = d_{\mathcal{S}} \circ Q \circ d_{\mathcal{S}}^{\dagger}$, as claimed in the main text.

between the duality maps $d_{\mathcal{S}}$ and $d_{\mathcal{Q}}$ is a little different. Here, we need the notion of a dual map ('dual' in the mathematical, not physical, sense). Thus recall that given any map between two sets $f : X \to Y$, and any map $g : Y \to Z$, the *dual* (or *transpose*) g^* of g (with respect to f) is defined as the map $g^* : X \to Z$ with $g^*(x) := (g \circ f)(x)$. Putting $X = Y = \mathcal{S}$, $f = d_{\mathcal{S}}$, and $Z = \mathbb{R}$, and taking the function $g : Y \to Z$ to be the quantity (as a real-valued function) $Q : \mathcal{S} \to \mathbb{R}$: this definition of the dual map becomes (with the bra-ket notation now meaning values of quantities as in Eq. (4)): $\langle Q^*, s\rangle := \langle Q, d_{\mathcal{S}}(s)\rangle$. But so far, the notation Q^* does not exhibit its dependence on $d_{\mathcal{S}}$; (just as g^* does not exhibit its dependence on f). So it is clearer to write $d^*_{\mathcal{S}}(Q)$ instead of just Q^*. Thus we write: $\langle d^*_{\mathcal{S}}(Q), s\rangle := \langle Q, d_{\mathcal{S}}(s)\rangle$. Applying this to $d_{\mathcal{S}}^{-1} : \mathcal{S} \to \mathcal{S}$, we deduce that defining $d_{\mathcal{Q}}$ as the dual $(d_{\mathcal{S}}^{-1})^*$ of $d_{\mathcal{S}}^{-1}$ yields Eq. (9), as desired. That is: we have by the definition of 'dual map' :

$$\langle (d_{\mathcal{S}}^{-1})^*(Q)\,,\; d_{\mathcal{S}}(s)\rangle = \langle (Q \circ d_{\mathcal{S}}^{-1}), d_{\mathcal{S}}(s)\rangle = \langle Q, s\rangle\;; \tag{10}$$

which is Eq. (9). So like in the quantum case: duality comes down to a *single* duality map on states, $d_{\mathcal{S}}$, with $d_{\mathcal{Q}}$ being defined as the dual, i.e. transpose, of its inverse $d_{\mathcal{S}}^{-1}$, and appropriate equivariance conditions on the dynamics.

We will see in Sect. 4 that for symmetries instead of dualities, we can similarly concentrate on a map on states—unsurprisingly, since the basic notion of symmetry is the preservation of the values of given quantities.

3.2 The road ahead: duality as a 'giant symmetry'

As discussed in Sect. 1, we have elsewhere related this Schema to various topics, and illustrated it with bosonization and some examples from string theory. The job of the next three Sections is to relate it to symmetries.

As also discussed in Sect. 1, it is agreed in the literature that a duality is like a 'giant symmetry': a symmetry between theories. The main new ingredient that the Schema adds to this agreed idea is its picture of two levels, with the bare theory above and the model roots, the bare theory's homomorphic copies, below. As we will see, these two levels can differ in the amount of structure that a map, such as a symmetry, is required to preserve; and this prompts some distinctions between types of symmetry.

Thus we will proceed in three stages.

(i) We begin with comments about symmetry in general (Sect. 4). They are regardless of both: (a) there being a duality; and (b) the distinction between a theory and its representations (homomorphic copies), or more generally its instantiations. These comments are familiar ground in the philosophy of symmetry: but they are worth making since they will apply, suitably adjusted, to the rest of the paper.

(ii) Then we discuss, regardless of there being a duality, how a symmetry of a theory is related to symmetries of its representations (homomorphic copies), or more generally its instantiations: (Sect. 5). This will yield the distinctions between types of symmetry.

(iii) Finally, we suppose we have a duality in the sense of the Schema, and relate this to symmetries. That is: we show that a duality preserves the symmetries of its model triples (Sect. 6).

4 Symmetries in general

We begin with a usual notion of symmetry: as a map a on states, $a : \mathcal{S} \to \mathcal{S}$, that preserves the values of a salient set of quantities: usually a large set, though not necessarily all the quantities. The map a must also respect the structure of $\mathcal{S}$, e.g. topological or differential structure. (Thus 'a' is for 'automorphism'.) But this requirement will be in the background in the sequel: the emphasis will be on the state s and the image-state $a(s)$ having the same values for quantities in the salient set.

Here, the notion of value is exactly as in the Schema: $\langle Q, s\rangle$, understood as a classical possessed value or a quantum expectation value (or more generally, as a matrix element, $\langle s'|\, \hat{Q}\, |s\rangle \in \mathbb{C}$: see below). The equality of values, for a symmetry a,

$$\langle Q, a(s)\rangle = \langle Q, s\rangle \tag{11}$$

is then analogous to the Schema's matching of values, under transforming both states and quantities by the duality maps $d_{\mathcal{S}}$ and $d_{\mathcal{Q}}$: cf. Eq. (9). (More generally: as in Sect. 3, we take quantum symmetries to also preserve off-diagonal matrix elements:

$$\langle a(s')|\, \hat{Q}\, |a(s)\rangle = \langle s'|\, \hat{Q}\, |\, s\rangle\ , \quad \forall\, s, s' \in \mathcal{S}\ , \tag{12}$$

for a salient subset of operators in $\mathcal{Q}$, usually including the Hamiltonian. And this condition can be weakened, to hold only for a salient subset of states in $\mathcal{S}$.)

This is of course the reason why the Schema confirms the 'giant symmetry' analogy. So far—i.e. before we focus on the two levels, bare theory above and model roots below—there are just two disanalogies between symmetry and duality:

(i) Equation (11) uses the identity map on quantities, while Eq. (9) uses a duality map $d_{\mathcal{Q}}$: corresponding to the jargon 'invariance' vs. 'covariance', and our phrase 'preservation or matching' above;
(ii) Equation (11) typically holds for a salient subset of the quantities, while the duality condition Eq. (9) holds for *all* the quantities: this will be illustrated below.

This notion of symmetry is very simple. But suitably adapted and augmented, it will be sufficient for this paper's purposes. In this Section, we make four comments about it. They are regardless of there being a duality; and of the distinction between levels, i.e. between a bare theory and model roots. So in this Section, we can just think of a theory as a triple of states, quantities and dynamics: $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D}\rangle$. These comments will apply, suitably adjusted, to the rest of the paper. The first, third and fourth comments (Sects. 4.1, 4.3, and 4.4) are about the notion being adaptable, including to dynamics. The second comment (Sect. 4.2) is about the idea of a salient set of quantities.

4.1 Dual maps

We said that a symmetry is a map on states that preserves the values of a salient, usually large, set of quantities. Agreed, it is also usual to think of a symmetry as a map on quantities that preserves values on a salient, usually large, set of states: i.e. for a given state, the value of the argument-quantity equals the value of the image-quantity. Instead of Eq. (11), one would write a symmetry as a map $\alpha : \mathcal{Q} \to \mathcal{Q}$ with:

$$\langle \alpha(Q), s \rangle = \langle Q, s \rangle. \tag{13}$$

But there is no conflict here. The two conceptions are related by duality in mathematicians' sense, not ours (cf. Sect. 3.1). That is: one map is the mathematical dual of the other. Recall that given any map $a : \mathcal{S} \to \mathcal{S}$, its *dual map* on quantities, $a^* : \mathcal{Q} \to \mathcal{Q}$, is defined by requiring that for any $s \in \mathcal{S}$ and $Q \in \mathcal{Q}$: $\langle a^*(Q), s \rangle := \langle Q, a(s) \rangle$. It follows immediately that if $a : \mathcal{S} \to \mathcal{S}$ is a symmetry in our initial sense, i.e. a respects the structure of $\mathcal{S}$, and Eq. (11) holds, then the dual map on quantities, $a^* : \mathcal{Q} \to \mathcal{Q}$ is a symmetry in the corresponding sense as regards quantities, given by Eq. (13):

$$\langle Q, a(s) \rangle = \langle Q, s \rangle \quad \Rightarrow \quad \langle a^*(Q), s \rangle := \langle Q, a(s) \rangle = \langle Q, s \rangle. \tag{14}$$

Besides, while we began with symmetry as a map on states, and conceived symmetries for quantities as dual maps: one could instead equally well start with quantities. For again, one defines dual maps in the same way. Given any map $\alpha : \mathcal{Q} \to \mathcal{Q}$, we say that its dual map on states, $\alpha^* : \mathcal{S} \to \mathcal{S}$, is defined by requiring for all arguments: $\langle Q, \alpha^*(s) \rangle := \langle \alpha(Q), s \rangle$. One could then define symmetries for quantities by Eq. (13), and deduce that if α is a symmetry for quantities, its dual map α^* is a symmetry for states, i.e. obeys Eq. (11).

Likewise in the quantum case, with preservation of matrix elements for a salient set of operators. We can define a symmetry α by:

$$\langle s' | \, \hat{Q} \, | s \rangle = \langle s' | \, \alpha(\hat{Q}) \, | s \rangle \, , \quad \forall \, s, s' \in \mathcal{S} \, . \tag{15}$$

Again, this condition can be weakened, to hold only for a salient subset of states in $\mathcal{S}$. Similar manipulations to the ones in the classical case[22] give that the symmetry, defined as a map α on quantities, induces a symmetry, defined as a map a on states, iff the symmetry map α decomposes in the following way:

$$\alpha(\hat{Q}) = a^\dagger \, \hat{Q} \, a \, . \tag{16}$$

Notice that this correspondence between symmetries α on quantities and symmetries a on states is not one-to-one. For example, if a commutes with a quantity Q, we can

[22] The proof is as follows. Begin with a symmetry $a : \mathcal{S} \to \mathcal{S}$ defined on states, viz. Eq. (12), and rewrite the left-hand side as: $\langle a(s') | \, \hat{Q} \, | a(s) \rangle = \langle s' | a^\dagger \, \hat{Q} \, a | s \rangle$. Comparing this with Eq. (15), we see that this induces a symmetry α, defined on quantities. Namely: $\alpha(\hat{Q}) = a^\dagger \, \hat{Q} \, a$. And the other way around: if the map α on quantities decomposes into the left- and right-action of some operator a, then this induces a symmetry $a : \mathcal{S} \to \mathcal{S}$ on states.

have a non-trivial symmetry on states that gives rise to a trivial symmetry on quantities (i.e. a map a that is non-trivial, while $\alpha = \mathrm{id}$ is trivial).

4.2 Salient quantities and states: stipulated symmetries

We said that a symmetry preserves values for 'a salient, usually large, set of quantities/states'. This general formulation deliberately uses the vague word 'salient', since it varies from case to case which quantities/states it is noteworthy to preserve the values of. But it is worth noticing three sorts of consideration that often mould the choice of quantities/states, i.e. which quantities/states count as salient. The first, which we label *stipulated symmetries*, gives a contrast with how we have written so far; the second and third will get a Subsection of their own (Sects. 4.3 and 4.4).

So far, we have written as if a theory is always given to us with prescribed sets of states and of quantities, so that the set of symmetries is thereby fixed, once some precise meaning of 'salient' is fixed. But as we foreshadowed in footnote 5: often in physics, we "begin our theorizing" with symmetry principles. That is: we define a theory's sets of states and quantities in order that they carry a representation of a given abstract symmetry group: spacetime symmetry groups such as the Poincaré group being a standard example. In such a case, we will say the symmetries are *stipulated*—they are part of the definition of the theory. Usually, we think of these symmetries as maps on states: unitary representations of a spacetime symmetry group on a Hilbert space being a standard quantum example.

In general, then, a theory T that is formulated as a triple, $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$, is said to have a *stipulated symmetry* if it is formulated as having an automorphism of the state-space, $a : \mathcal{S} \to \mathcal{S}$, that preserves some salient subset of the quantities. The stipulated symmetry thus comes with a choice of which quantities count as salient, so that their values are "worth" preserving. This choice is encoded formally in the definition of the triple and its stipulated symmetry, and it of course bears on the theory's interpretation—it moulds the kinds of interpretations that the theory can be given.

But note that stipulating a symmetry does *not* imply that every state has its value preserved for *every* quantity definable on the state-space. (And *mutatis mutandis* if we conceive a symmetry as a map on quantities; cf. Sect. 4.1 above.) For example, one might stipulate rotational symmetry: more precisely, that in a quantum theory the Hamiltonian is rotationally invariant, so that the Hilbert space carries representations of SO(3). But this still allows using quantities Q whose expectation values on some states are not invariant under rotations. Thus even with a stipulated symmetry, there is a question of selecting the salient quantities. This point will recur in Sect. 4.3.[23]

[23] For example, Caulton (2015: p. 156) distinguishes three natural classes, two of which are extreme and trivialising, while the third is the most interesting: (1) symmetries that preserve *all* the quantities $\mathcal{Q}$ of a theory: these are required to preserve everything, which means changing nothing; (2) symmetries that do not preserve any quantities, i.e. the set of all bijections on the state-space $\mathcal{S}$; (3) symmetries that lie between these extremes are the interesting cases: especially symmetries that preserve the 'set of physical quantities, i.e. the quantities that, on the basis of their representing the physical properties and relations, register physical differences' (ibid).

4.3 Dynamical symmetries

We have presented symmetries as maps that preserve the salient quantities' values (and respect the structure of $\mathcal{S}$ or $\mathcal{Q}$: cf. Sect. 4.1). But we have not mentioned time, i.e. the fact that values change over time. It is indeed very usual to define a symmetry as a map that 'preserves the dynamics'. Taking a symmetry, as usual, as a map on states, this means, roughly: if a sequence of states is possible according to the dynamics, so is the sequence of image-states.

We can make this precise by using the framework for dynamics, in both Schrödinger and Heisenberg pictures, given in Sect. 2.4, Eqs. (5) and (6). We will favour the former. (As noted there, a full discussion would also address the treatment of time in theories not best thought of as triples, e.g. theories formulated using partition functions; again, cf. footnote 6.)

On Schrödinger dynamics, a dynamically possible total history of the system is a curve through the state-space $\mathcal{S}$ parameterized by time t: with each point $s(t)$ defining the values $\langle Q, s(t)\rangle$ of the various quantities Q at t. Then we can define a dynamical symmetry as a map a on $\mathcal{S}$ that (a) respects $\mathcal{S}$'s structure (b) maps any dynamically possible history (curve through state-space) to another such history. That is: if $s(t)$ is a dynamically possible history, the sequence $a(s(t))$ of states is also dynamically possible.[24]

On Heisenberg dynamics, the definition of a dynamical symmetry is (we think!) more complicated, because the representation of a dynamically possible history is more complicated. A history is given by a fixed $s \in \mathcal{S}$ and a family of curves through $\mathcal{Q}$, all parameterized by time t: with Q_1, Q_2 on a common curve representing the same physical quantity, e.g. energy, at two times t_1, t_2. So for a single history, there are as many curves through Q in the family as there are physical quantities pertaining to the system. So a dynamical symmetry must be a map whose domain is, not $\mathcal{Q}$, but the set of all such families of curves (or all such families that are indeed dynamically possible, once some $s \in \mathcal{S}$ is chosen). So the map will have to suitably respect, not so much $\mathcal{Q}$'s structure, but the structure $\mathcal{Q}$ induces on this set of families of curves. And for the map to be a dynamical symmetry, it must leave invariant the dynamically possible families (allowing, no doubt, for a change of state $s \in \mathcal{S}$). But in this paper, we can focus on symmetries as maps on states, and so we will not need to further consider the Heisenberg picture.

Given this definition of dynamical symmetry in Schrödinger picture, as a map on $\mathcal{S}$ that commutes with the dynamics (footnote 24), the obvious first question is: how is this related to our initial idea of symmetry as a map on $\mathcal{S}$ that preserves quantities' values, regardless of time?

A priori, they seem very different. Indeed the notion of dynamical symmetry seems weaker in that it requires the transform $a(s(t))$ of each dynamically possible history $s(t)$ 'only' to be itself dynamically possible—it need not have any distinctive relation to $s(t)$, e.g. by being in some sense a 'replica' of $s(t)$. But in fact the notions are drawn together by dynamical symmetry's requirement that the map on histories be induced

[24] Preserving the dynamics in this sense is of course a commutation i.e. equivariance condition. For if we write $s(t) = D_{t,t_0}(s(t_0))$ with D_{t,t_0} representing the dynamics (cf. Eq. 5), then preservation of the dynamics is: $a(s(t)) \equiv a(D_{t,t_0}(s(t_0))) = D_{t,t_0}(a(s(t_0)))$.

by a map on states. That is: writing a history as a set of points in $\mathcal{S}$ for brevity, $\{s(t)\}$: a dynamical symmetry requires that the map on histories $\{s(t)\} \mapsto \{s'(t)\}$ be induced by a map on instantaneous states, i.e. $s(t) \mapsto a(s(t))$. This turns out to be a strong requirement, thanks to the 'sensitivity' of dynamics to the values of many quantities. That is: it turns out to force a to preserve the values of many quantities—leading us back to our initial idea of symmetry. But note that this implication is not *a priori*: it depends on what dynamical evolution, in typical theories, in fact depends on.

The point here is well illustrated by spacetime symmetries, as mentioned in Sect. 4.2. Take for example, spatial translation or spatial rotation in Euclidean space $\mathbb{R}^3$; and consider any of Newtonian mechanics, relativistic mechanics, quantum mechanics, or indeed their 'cousin' field theories. Each of these is of course a framework for theorising: not a specific theory, with specific particle and-or field contents, and their dynamics (equations of motion). But it turns out that most such specific theories that have been empirically successful,[25] once set in Euclidean space $\mathbb{R}^3$, do have spatial translations and spatial rotations as dynamical symmetries. And since a dynamical symmetry is to be induced by a single map a on instantaneous states, e.g. by spatial translation of 1 mile due East applied to every state, the transform $a(s(t))$ of each dynamically possible history $s(t)$ will indeed be a 'replica' of $s(t)$, e.g. spatially translated by 1 mile. Besides, the dynamics is 'sensitive' to the values of many quantities, in the sense that a dynamical symmetry must *not* alter them: so its map a on instantaneous states is indeed a symmetry *à la* our initial idea. Again, spatial translations and spatial rotations give a standard illustration; as follows. As we said, most of the empirically successful specific theories written in any of the above frameworks, and set in Euclidean space $\mathbb{R}^3$, have these as dynamical symmetries. But they "don't allow squeezing"! That is, a dynamical symmetry a must preserve all the relative distances, and relative velocities, between the constituents of the system: if a is, or includes, a spatial translations or rotation, it must be a rigid one, of the system as a whole. It must preserve the values of many quantities—the relative ones.

This discussion of a dynamical symmetry leads in to Sect. 4.4. Also, the cautionary note at the end of Sect. 4.2 above applies again. That is: stipulating that a symmetry be dynamical—indeed, stipulating more specifically: both a symmetry and a dynamics it respects—does not imply that every state has its value preserved for every quantity definable on the state-space. Again, the example of SO(3) in quantum theory suffices. One might stipulate that SO(3) be a symmetry, and that the Hamiltonian be SO(3)-invariant (i.e. in obvious notation: $[H, U_R] = 0$ for all $R \in \mathrm{SO}(3)$). This does not imply that one can only use quantities Q that are rotationally invariant (i.e. $[Q, U_R] = 0$).

4.4 Spacetime theories and their symmetries

In philosophical and foundational discussions of spacetime theories, it is usual to define symmetries in an apparently different way from ours.

[25] 'Successful' within limits, of course: for example, relative velocities small compared with c for Newtonian mechanics to be successful, and typical actions large compared with h for Newtonian or relativistic mechanics to be successful.

Besides, it is usual to define such theories, not as a triple $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ as we have done, but as a set of, so to speak, possible universes. That is: as a set of n-tuples, consisting of a spacetime manifold M, equipped with both chrono-geometric structure (encoded in a metric field, a connection etc.) and matter fields (encoded in tensor and spinor fields obeying equations of motion). Each such n-tuple represents a (total, 4-dimensional) solution of the theory: a 'possible universe'.[26] A symmetry is then usually defined along the following lines. (Recall the example of Galilean transformations in Sect. 2.3.2.) It is a bijection of the manifold that:

(a) respects its topological and differential structure (technically: is a diffeomorphism); and whose induced maps on tensor fields, connections etc.:
(b) fix the chrono-geometric structure (i.e. maps the metric field, connection etc. into themselves) and also
(c) send the matter fields into another solution of the equations of motion—another sequence of values over time that is dynamically allowed/possible.

So we need to link our construal of a theory as a triple, and of a symmetry as a map on a state-space, to these ideas. The main link is of course that while a physical theory usually has as its subject-matter some limited kind of system, for which we think of the instantaneous state (values of quantities) changing over time, a spacetime theory takes the universe-throughout-all-time as its subject-matter. So in our construal, a dynamically possible total history of the system is, on Schrödinger picture dynamics, as discussed in Sect. 4.3: a curve through $\mathcal{S}$ parameterized by time t, with each point $s(t)$ defining the values $\langle Q, s(t) \rangle$ of the various quantities Q at t.[27] But in a spacetime theory, a dynamically allowed/possible total history of the system is just an n-tuple.

One natural way to link to our construal is to make a space-vs.-time split within the spacetime theory's manifold.[28] That is: we take the spacetime theory to have:

[a] a state-space $\mathcal{S}$ of the instantaneous states of a notional 3-manifold Σ, which we take as a fiducial spacelike slice of the spacetime manifold;
[b] a set of quantities $\mathcal{Q}$ defined on Σ (local densities of matter fields etc.); and
[c] a dynamics $\mathcal{D}$ determining the evolution of the instantaneous state of Σ. Then a dynamically possible total history is a foliation of spacetime, whose leaves are time-evolutes of Σ, equipped with fields. In other words: it is a curve through $\mathcal{S}$ parameterized by a time t labelling the leaves of the foliation.

[26] So these n-tuples are usually called 'dynamically possible models' of the theory: but we will resist yet another use of the over-worked word 'model'!

[27] As in Sect. 4.3 , one can have a corresponding discussion using the Heisenberg picture. But here, we set this aside.

[28] An alternative way to link to the usual covariant formulations of spacetime theories takes a model of T to have as its state-space the set of all admissible spacetimes that are solutions of the theory's dynamical equations (e.g. in general relativity, the Einstein equations). But adopting either of these ways, here and in the main text, one faces issues, if a state is a whole spacetime, about how to "get inside a spacetime", so as to distinguish a quantity's differing values at different spacetime points. Thus one widely adopted approach is to associate quantities not with points, but with extended regions of different types, i.e. one gets 'quasi-local quantities': see Penrose (1982) and Brown and York (1993); for reviews, see Wang (2015) and Szabados (2009). There are also other issues about which much could be said, such as: rigorously defining quasi-local quantities; fixed fields; the definition of diffeomorphism invariance and its interplay with boundary conditions. See for example Pooley (2017: p. 117) and De Haro (2017).

Combining [a]–[c] with our discussion of dynamical symmetries in Sect. 4.3, we can now see that our initial simple idea of a symmetry, as a map a on the system's state-space $\mathcal{S}$ that respects its structure and preserves the values of a salient set of quantities, is, after all, similar to the usual definition of symmetry for spacetime theories, (a)–(c) above. For making the space-vs.-time split, [a]–[c], renders this usual definition of symmetry, (a)–(c), like Sect. 4.3's definition of dynamical symmetry. Indeed, most of the familiar (and empirically successful) theories set in the framework of Newtonian mechanics, or relativistic mechanics or quantum mechanics (or their 'cousin' field theories) mentioned in Sect. 4.3 can be written down as spacetime theories, i.e. as postulating a spacetime manifold equipped with chrono-geometric structure and matter fields: with the different notions of symmetry being linked by using a space-vs.-time split.

There are of course several issues here, about which much more could be said. Among them are:

(i) the interplay of the structures (topological, differential, metrical etc.) of physical space, spacetime, and state-space;
(ii) the justification for singling out, in the spacetime definition of symmetry, the chrono-geometric structure of the manifold as having to be fixed (requirement (b)), while the matter fields need not be (especially in the context of general relativity); and
(iii) the justification for making the space-vs.-time split, especially in the context of relativity theory (especially general relativity: cf. footnote 28).

And of course, such issues are mutually related. For example, (i) and (ii): one might argue that the requirement (b), to fix the chrono-geometric structure, reflects the requirement in our initial idea of symmetry, that a must respect the structure of the state-space $\mathcal{S}$, and the fact that $\mathcal{S}$'s structure is largely determined by the chrono-geometric structure of spacetime. However, for this paper we will not need to explore these issues: sufficient unto the day ...

5 Symmetries of theories and of models

We now combine the ideas of Sect. 4 with Sect. 3's distinction between a bare theory T and its *models* in our sense, viz. representations of T, each with a specific structure of its own. Our notation distinguishes the model's representation of the bare theory's triple, from its specific structure $\bar{M}$. It gives a subscript M to the former to signal that it is built out of the latter: $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M; \bar{M} \rangle$. Cf. Eq. (1). We also wrote this as $M =: \langle m; \bar{M} \rangle$, where $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$ is the model triple.

As we announced in the prospectus (Sect. 3.2), the differing amounts of structure at the two levels, bare theory and model, prompt some distinctions, and some definitions of types of symmetry.

We will begin with the idea of symmetries of a theory that are implemented in all the models, which leads to the idea of *stipulated symmetries* (Sect. 5.1). Then we consider symmetries that *fail* to be implemented in some of the models, which leads to the idea of *accidental symmetries* (Sect. 5.2). We will see that this failure can happen for two

diverse reasons. Roughly speaking: either the theory's homomorphic copy has "lost the structure with which to exhibit the symmetry", or the model has "extra structure that blocks the symmetry".

Then in Sect. 5.3, we will discuss cases where the specific structure in a model has a symmetry that is "invisible" in the weaker structure of the model triple (the homomorphic copy of the bare theory). Here, 'invisible' will mean, roughly speaking: 'cannot be defined, or reduces to the trivial symmetry, i.e. the identity map'. We call such a symmetry a *proper symmetry* of the model. Note that these cases yield a contrast with the other, more common, usage of 'model' as an individual solution of a theory. On that usage, a model is in general *less* symmetric than its theory—as is often remarked, with the buzz-word 'symmetry-breaking'. A solution of a dynamics with a spherically symmetric Hamiltonian need not be spherically symmetric; a cubical crystal lattice with one particular placing of its lattice points, and one particular orientation of its edges, can be a solution of a dynamics that is translation-invariant and isotropic; and so on. So *beware*: our cases of proper symmetries of a model are cases 'lying in the opposite direction' from what one expects from the other, more common, usage of 'model'.

Finally, Sect. 5.4 will sum up the Section. It emphasises that these three types of symmetry are mutually exclusive. (It also mentions a fourth type: which will not be important, but which completes our classification, i.e. makes the four types exhaustive.)

This Section builds on Sect. 4. For in this Section:

(i) We will again take a symmetry as a map a on states, though we could equally well take it as a map on quantities;
(ii) We will return to the idea of stipulated symmetries; and
(iii) One can, at each stage of the discussion, require the symmetry to be a dynamical symmetry; but we will not explicitly mention this;

and these three points echo Sects. 4.1, 4.2 and 4.3, respectively.

5.1 Guaranteeing that a bare theory's symmetry is implemented in a model: stipulated symmetries

For any theory T and any of its models M, there is a natural condition for a symmetry of T to be itself realized in M: for it to have, so to speak, a "shadow" in the model M. It is that an obvious diagram should commute (cf. Fig. 3). Here, we write representation as a map h ('h' for homomorphism). So treating symmetries as maps on states: a symmetry of $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ is a map $a : \mathcal{S} \to \mathcal{S}$ that preserves the value of all quantities in a salient subset of the quantities $\mathcal{Q}$. So we take h as an appropriate structure-preserving map: from $\mathcal{S}$ in the theory T itself, to $\mathcal{S}_M$ in the representing model triple $m = T_M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$. Then the condition—that the symmetry a is itself realized in M—is that there should be a map $a_M : \mathcal{S}_M \to \mathcal{S}_M$, such that the diagram in Fig. 3 commutes.

But this condition is not automatic: even in simple cases, like elementary group theory. In Sect. 5.2 we will discuss how the condition can fail: here we consider ways to secure it.

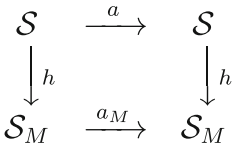

Fig. 3 Commutativity diagram of the symmetry a with the representation map h

There are two obvious ways to do so: i.e. ways to secure that a bare theory's symmetry is implemented in a model of it. The first is a formal triviality; the second is a stipulation.

First: consider the case where the model triple is *isomorphic* to the bare theory. (In the language of representation theory: the representation is *faithful*, i.e. the homomorphism has trivial kernel.) With the map h an isomorphism, the map a_M is bound to exist; and the diagram in Fig. 4 will trivially commute. In this case, a symmetry of the theory has a "duplicate" or "replica" (not just a "shadow") in the symmetries of the model—so that in effect, the symmetries of the theory are a *subset* of the symmetries of the model. We say 'in effect' just because of the different domains of definition: $\mathcal{S}$ vs. $\mathcal{S}_M$.

Second: consider, so to speak, the power of stipulation. That is: recall the idea of *stipulated symmetries*, as in Sect. 4.2. The idea there was that often in physics, we "begin with", i.e. are guided by, symmetry principles, and therefore *require* the state-space of our theory to carry a representation of a group, with values of the salient quantities being preserved under the action of the (representation's) group elements. One standard example is requiring the states, of whatever type of particle or field, to carry a representation of a spacetime symmetry group, such as the Poincaré or Galilean group.

Carrying this idea over to our two levels, of bare theory and models: it means requiring each model triple to have a symmetry that is a shadow of the bare theory's symmetry. In terms of symmetries as maps on states: each model triple is simply required to have a symmetry a_M on $\mathcal{S}_M$ that makes Fig. 3 commute. In short: one has commutation by stipulation. We shall call such symmetries—both a on the bare theory that 'does the requiring', and the 'required' a_M on the model M—*stipulated symmetries*. And similarly for a whole group of symmetries. And again, a spacetime symmetry group gives a standard example.

This second idea can be implemented in practice as follows: (i) Stipulate the symmetries to be part of the *definition* of the bare theory, T (cf. footnote 5). (ii) Reduce the set of models to only those models that instantiate the symmetry, i.e. those for which the diagram in Fig. 3 commutes. Here, (i) implies (ii).

Thus, for a field theory based on the homogeneous Lorentz group, one can systematically construct the fields by looking at the irreducible representations of this group.[29] Thus we begin (cf. Weinberg 1995: Chapters 5.2–5.6) with the trivial representation of the Lorentz group, which gives a massive scalar field. The next (in terms of its dimension) irreducible representation is the Dirac spinor representation describing a fermion. The next irreducible representation is the (massive) vector representation,

[29] For textbooks on quantum field theory that take this approach from their very first pages, see Weinberg (1995: Chapters 2.4–2.7, 5) and Maggiore (2005: Chapter 2).

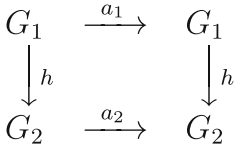

Fig. 4 Commutativity of group automorphism a_1 with group homomorphism h

i.e. a four-vector (which contains spin zero and spin one states); and we can have various tensor representations, etc.

Often, in practice, and in particular for spacetime symmetry groups: these two ways to secure the commutation—to secure that a bare theory's symmetry is implemented—come together. For the stipulation made is not *just* that the state-space of the model triple carry a symmetry, or carry some or other representation of a group of symmetries. It is also stipulated that it carry a *faithful* representation of the group. That is: the representation map h is required to be an isomorphism. So the upshot is: a symmetry of the bare theory has a duplicate or replica (not just a shadow) in the symmetries of *each* model (but of course represented differently in the theory and in each of the models)—so that in effect, the stipulated symmetries of the theory are a *subset* of the symmetries of the models.[30]

5.2 A bare theory's symmetry might fail to be implemented in a model: accidental symmetries

As we said at the start of Sect. 5.1, Fig. 3 need not commute—even in simple cases, like elementary group theory. In fact, the condition can fail in either of two ways: because the theory's homomorphic copy has "lost the structure with which to exhibit the symmetry"; or because the model has "extra structure that blocks the symmetry". We will give an example of each of these: the first from group theory, the second from elementary spacetime theory.

So first, let the bare theory be just a group G_1,[31] with an automorphism $a_1 : G_1 \to G_1$; and suppose a group G_2 represents G_1 thanks to the existence of a homomorphism $h : G_1 \to G_2$. So $G_2 \cong G_1/\ker h$. Then for there to be a homomorphism of G_2, $a_2 : G_2 \to G_2$ (even homomorphism: let alone automorphism), that realizes a_1 (counts as a_1's "shadow" in G_2) requires commutation: i.e. for all $g_1 \in G_1, h(a_1(g_1)) = a_2(h(g_1))$. Or as a diagram, see: Fig. 4. But a_1 and h need not mesh in this way: commutation of the diagram is not guaranteed. The smallest counterexample we have found is presented in the Appendix.

A second example of a theory's symmetry failing to exist at the level of its models is given by Sect. 2.3.2's neo-Newtonian spacetime as a bare theory, modelled by various Newtonian spacetimes each with their own specification of absolute rest. Here it is "extra structure in the models" that blocks the implementation of the theory's

[30] Again, we say 'in effect' just because of the different domains of definition: $\mathcal{S}$ vs. $\mathcal{S}_M$. And as mentioned in this Section's preamble: note the contrast with the other, more common, meaning of 'model' as an individual solution of a theory. On that meaning, a model is in general *less* symmetric than its theory.

[31] To be a bare theory in our sense, we would need to add a set of maps to the real numbers, to express evaluation of the quantities, and even a dynamics. But to give the simplest case, we ignore these maps.

symmetry. Thus we take neo-Newtonian spacetime to be so defined that the Galilean group is its symmetry group: i.e. the set of diffeomorphisms of the spacetime that fix the spatial metric (giving Euclidean geometry in the instantaneous slices), the temporal 1-form (ordering the slices like $\mathbb{R}$), and the flat affine connection (giving a notion of absolute acceleration, without a notion of absolute rest or absolute velocity). (Cf. Sect. 4.4.) Thanks to avoiding an absolute rest, the boosts are symmetries. But in a Newtonian spacetime, with the extra structure of a specified absolute rest, the boosts are not symmetries: the symmetry group is reduced to the Euclidean group (i.e. rotations and spatial translations) combined with time translations. In our jargon, whereby a model comprises specific structure as well as a model triple: the specific structure prevents the symmetry.

Accidental symmetries: We have seen that, in general, the symmetries of a theory are not instantiated in its models. But this of course does not forbid special situations in which a theory happens to have symmetries that are not stipulated (they are not part of the definition of the bare theory). And these symmetries may (though of course they need not) have "shadows" in some (but not all) of their models. This is the situation where the diagram in Fig. 3, for a specific model M, commutes "for free", i.e. without the symmetry a being part of the definition of the bare theory T. We call these *accidental symmetries*.

To illustrate the practical usefulness of the notion of accidental symmetries, we now give a very simple example that we will generalise in the next Section. Consider the bare quantum theory of the hydrogen atom, with a Hamiltonian that only contains a kinetic term for the electron, and a potential that only contains the Coulomb attraction, i.e. the basic textbook example of the hydrogen atom, neglecting all other interactions. Thus the states of our theory are square-integrable wave-functions, the quantities are the set of operators of the hydrogen atom (in particular the Hamiltonian), and the dynamics is the Schrödinger equation (thus we adopt the Schrödinger picture). This theory has many models, in our logically weak sense of 'model' from Sect. 2.1.1. Indeed, each square-integrable wave-function that solves the Schrödinger equation with given initial conditions (along with the set of operators that act on it, and its time evolution) is a representation of the theory—even if a very non-faithful representation.[32] Let us call these models $M_{n\ell m}$, where n, ℓ, m are the quantum numbers of the hydrogen atom.[33] But the set of all such models, i.e. the set of all square-integrable wave-functions that solve the Schrödinger equation, is itself also a model. Let us call this model $M_{\text{total}} := \{M_{n\ell m}\}$, where n, ℓ, m have the usual ranges.

Let us now discuss the spherical symmetry of this bare theory of the hydrogen atom. Particular solutions (except for special, spherically symmetric, solutions) are of course not spherically symmetric. Thus, most of the individual models $M_{n\ell m}$ do *not* enjoy the SO(3) symmetry of the theory (although they are mapped into each other by the action of SO(3)). On the other hand, M_{total} *does* carry non-trivial representations of SO(3), which acts as an automorphism of the set of states. Namely, SO(3) maps

[32] This recovers the usual meaning of 'model', (i) in Sect. 2.1.1, as a *special case* of 'model', in our sense. Cf. footnote 3.

[33] Since each state is determined by a set of initial conditions, it is in general not the case that only states with fixed values of n, ℓ, m are allowed. The general allowed state is a linear combination of these, depending on the boundary conditions. But for notational simplicity, we label the models by quantum numbers.

states with the same n and ℓ but different values of m, so that ℓ labels the distinct representations of SO(3) (and, by focussing on a suitable set of eigenfunctions of the Laplacian that form a complete set, one can get the irreducible representations). In other words, SO(3) is a symmetry of M_{total}. Of course, all of the above generalises to the case where the potential is specified only to be a spherically symmetric function, i.e. of the form $V(r)$, and we consider other spatial dimensions: as we will do in the next Section.

Thus rotational symmetry is an accidental symmetry of this particular theory. It is a symmetry of the theory that is represented non-trivially in one of the models of the theory, but not in all the models.

5.3 Proper symmetries of models

In the last two Sections, we considered symmetries of a theory that are stipulated to be symmetries of all the models (Sect. 5.1) or that fail to be realized in at least some of the models (Sect. 5.2). In this Section, we consider symmetries that are *not* defined in the theory, but that are nevertheless symmetries of a model. For a model's specific structure $\bar{M}$—its 'content' that goes beyond its being a model/realization of T—can make the model have symmetries *additional* to those that are "shadows", or even "duplicates", of the symmetries of T. And we expect that if these additional symmetries are well-defined on the model triple, or if they naturally induce a symmetry there, that symmetry is trivial, i.e. just the identity map on the model triple.

Our prototypical cases of representations of a group or algebra give examples. Perhaps the simplest is as follows. Let T be the real numbers $\mathbb{R}$; and let M be the complex numbers $\mathbb{C}$ which of course represents $\mathbb{R}$ as the real axis, i.e. the complex numbers with zero imaginary part, $\{z \in \mathbb{C} \,|\, z = x + i0,\ x \in \mathbb{R}\}$. So this latter set, the real axis, is like the model triple. Then M has the symmetry of complex conjugation $z \mapsto \bar{z}$ which is indeed well-defined on the real axis: but there, it is just the identity map.[34]

This prompts the idea of a *proper symmetry* of a model: where the word 'proper' is to connote 'specific to the model'. The idea is that such a symmetry depends, at least in part, on the model's specific structure—which, recall, "lies beyond" the model triple's representing the bare theory: just like complex conjugation in the example just given.

To give a general definition, it is clearest to introduce another notation for a model, as a triple of its own set of states, quantities and dynamics. So we write: $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$. Contrast the way our previous notation in Eq. (1) separated out the specific structure $\bar{M}$. Thus we wrote: $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M; \bar{M} \rangle = \langle m; \bar{M} \rangle$. With the new notation,

[34] There are also examples in interesting cases of dualities. In gauge-gravity dualities, De Haro (2017) showed that a certain subgroup of the diffeomorphism group of the gravity model of the theory (roughly, the diffeomorphisms which preserve the asymptotic boundary conditions) was 'invisible' to the gauge model of the theory, in the sense of not representing any difference on that model: and so these diffeomorphisms are not in the common core between the two models, and they are trivially represented on the bare theory. The same verdict was made in De Haro (2016b: Sect. 2.2.3) for the 'gauge symmetries' of the gauge side of the duality. These are not visible on the gravity side: they are symmetries of the formulation of the gauge model of the theory, and are trivially represented on the bare theory.

we can discuss 'symmetry of a model' in just the same way that, already in Sect. 4, we discussed 'symmetry of a theory': namely, as a map a on the state-space $\bar{\mathcal{S}}$ that preserves the values of a salient, usually large, set of quantities.

This suggests the following definition. Given a (bare) theory T that is represented by various models and presented as a triple, we will say that a symmetry a of one such model $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$, $a : \bar{\mathcal{S}} \to \bar{\mathcal{S}}$, is a *proper symmetry* of M, if there is some other model of T, say $M' = \langle \bar{\mathcal{S}}', \bar{\mathcal{Q}}', \bar{\mathcal{D}}' \rangle$ for which, in effect, a cannot be defined; and a can, in effect, also not be defined for the theory T. (Here, we say 'in effect' for the same reason as in Sect. 5.1: i.e. just because of the different domains of definition: $\bar{\mathcal{S}}$ vs. $\bar{\mathcal{S}}'$.) More precisely: in the model $M' = \langle \bar{\mathcal{S}}', \bar{\mathcal{Q}}', \bar{\mathcal{D}}' \rangle$, there is no natural definition of a symmetry map $a' : \bar{\mathcal{S}}' \to \bar{\mathcal{S}}'$ that is a "cousin" of a, except perhaps as the identity map.

Again, the complex numbers vs. the real numbers give what is perhaps the simplest example, with the real line now taken as another model $M' := \mathbb{R}$, along with the complex plane $M := \mathbb{C}$: rather than $\mathbb{R}$ being the theory T. In this example, complex conjugation is a proper symmetry of $M := \mathbb{C}$: it reduces to the identity map on the other model, $M' := \mathbb{R}$.

Consider again the example from the previous Section, but in generalised form. We stipulate that our bare theory is the quantum mechanics of the abstract two-body problem in an arbitrary number of spatial dimensions. Consider now models that specify:

(i) the number of spatial dimensions, n, and the idealization that one body is a test-body, which secures rotational invariance, and so an SO(n) symmetry; and
(ii) the form of the potential.

Specifically, we will take the model to specify the dimension of space to be 3, and the potential to be the Coulomb potential. With these definitions, the SO(3) rotational symmetry is not a symmetry of the bare theory but it is a symmetry of the model. Furthermore, the model's specifying the potential to decay like $1/r$ means that the SO(3) symmetry is enlarged to SO(4), where the additional symmetry corresponds to a new conserved quantity—the Runge-Lentz vector.[35] On the other hand, notice that there are countless other choices of potential that do not share the SO(4) symmetry. Thus, in fact, this model has an SO(4) symmetry that the bare theory and countless other models do not share: it is a proper symmetry of this model.

5.4 Kinds of symmetry

Let us sum up this Section's discussion of how the symmetries of a bare theory relate to those of its models. We have made two main points, as follows.

(i) A bare theory T is represented by one of its model triples, m. The model M then consists of m and some specific structure $\bar{M}$; (cf. Sect. 2). But representation requires only a homomorphism, not an isomorphism. Hence our articulating in

[35] Recall that the rotational invariance of the Hamiltonian implies the conservation of angular momentum, which means that the classical orbits lie on a plane. Likewise, the SO(4) symmetry implies the conservation of the Runge-Lenz vector, namely $\mathbf{M} := (\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p})/2m - e^2\mathbf{r}/r$: i.e. it commutes with the Hamiltonian.

Sects. 5.1 and 5.2 the condition—in terms of a commuting diagram—for a symmetry of T to be itself realized in m. As Sect. 5.2 described, this condition can fail. So when this condition *is* met for some of the models—not by stipulation but "accidentally"—we say that there is an *accidental symmetry*.

(ii) When the diagram is stipulated to commute, so that each model has a 'duplicate' or 'replica' of the symmetry a, we have a *stipulated symmetry*. There are two obvious ways to implement this stipulation: defining the theory to have a symmetry, and accordingly restricting the set of models to those that respect the symmetry, or requiring the model triples to be isomorphic to the theory T (the latter condition being strictly stronger). But even in such cases, the *model* (as against the model triple) has its own specific structure $\bar{M}$—which may have symmetries that m, and the other models "know nothing of": these are *proper symmetries of the model* (Sect. 5.3).

We can also now see that our three types of symmetry are mutually exclusive. Stipulated symmetries and accidental symmetries are mutually exclusive because the former are stipulated while the latter are not; and stipulated symmetries and proper symmetries of models are mutually exclusive because, while the former has a shadow (i.e. can be defined as a symmetry) in all of the models, the latter cannot be defined in at least some models (and it is also not defined in the theory).

Accidental symmetries and proper symmetries are mutually exclusive because an accidental symmetry is a symmetry of the theory (although it is not stipulated); while a symmetry of the theory cannot be a proper symmetry of a model.

The only possibility that is not covered by the above three kinds of symmetries is the possibility that a symmetry is a symmetry of all the models but is not a symmetry of the theory. Such a symmetry can for example be obtained when a proper symmetry of the models is shared by some models (but of course not all), and then the models for which this symmetry is not realized are excised from the set of models. We could call such a symmetry an *improper symmetry of the models* (improper, in the sense that it is a symmetry of all the models but not of the theory).

Indeed, it is now true that any symmetry (in our sense) belongs to one of these four categories. For, given a theory and a set of models, a symmetry can either be a symmetry of the theory and of all the models (hence it is stipulated); or it is a symmetry of all the models but not the theory (an improper symmetry); or it is a symmetry of the theory and of only some models (an accidental symmetry); or it is a symmetry of some models but not of the theory and not of some other models (a proper symmetry of the models); or it is only a symmetry of the theory (also an accidental symmetry).

6 Duality preserves symmetries

It is straightforward to confirm that on Sect. 3.1's definition of duality, a duality preserves any symmetry (including dynamical symmetries) of its model triples. There are two points here. First: there is a commuting square diagram of isomorphisms. Second: there is the issue of the values of a quantity being equal on a given state, and on its transform under a symmetry. The first point will lead in to the second.

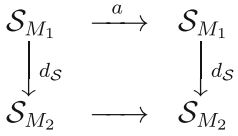

Fig. 5 Commutativity of duality and symmetry for states

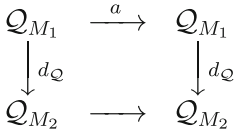

Fig. 6 Commutativity of duality and symmetry for quantities

First: The duality maps $d_{\mathcal{S}}, d_{\mathcal{Q}}$ are not only bijections, but isomorphisms: $d_{\mathcal{S}} : \mathcal{S}_{M_1} \to \mathcal{S}_{M_2}$, and $d_{\mathcal{Q}} : \mathcal{Q}_{M_1} \to \mathcal{Q}_{M_2}$. And although we did not have to spell out the exact structures of $\mathcal{S}_{M_i}, \mathcal{Q}_{M_i}$ that these isomorphisms are to preserve, it is obvious from the fact that 'is isomorphic to' is both a symmetric and a transitive relation, that the diagram above, with a understood to be any automorphism of $\mathcal{S}_{M_1}$, commutes (cf. Fig. 5).

And of course, this diagram of isomorphisms is just what we mean by saying a duality d preserves an automorphism of the state-space $\mathcal{S}_{M_1}$ in its domain model triple, and preserves $\mathcal{S}_{M_1}$'s structure. Namely, d carries the automorphism—a map a on $\mathcal{S}_{M_1}$—to a corresponding automorphism of states in the codomain (indeed: range) model triple. The diagram defines this corresponding automorphism, i.e. the map forming the fourth side of the square: $d_{\mathcal{S}} \circ a \circ (d_{\mathcal{S}})^{-1} : \mathcal{S}_{M_2} \to \mathcal{S}_{M_2}$.

There is obviously a corresponding point about quantities, as against states. Since $d_{\mathcal{Q}}$ is required to be an isomorphism of quantities, the diagram above, with a now understood to be any automorphism of $\mathcal{Q}_{M_1}$, must commute, cf. Fig. 6.

And again, this diagram is just what we mean by saying a duality d preserves an automorphism of the quantities in its domain model triple, and preserves $\mathcal{Q}_{M_1}$'s structure. Namely, d carries the automorphism—a map a on $\mathcal{Q}_{M_1}$—to a corresponding automorphism of quantities in the codomain (indeed: range) model triple. The diagram defines this corresponding automorphism: $d_{\mathcal{Q}} \circ a \circ (d_{\mathcal{Q}})^{-1} : \mathcal{Q}_{M_2} \to \mathcal{Q}_{M_2}$.

In short: Figs. 5 and 6 show that duality commutes with automorphisms of the states and of the quantities.

Second: But in physics, the notion of symmetry of course involves more than the notions of automorphism of the state-space, and of the set (usually algebra) of quantities. It involves the pairing whereby states s and quantities Q assign each other a value: $\langle Q, s \rangle$. For these values (for a large and salient set of quantities) must be preserved under the symmetry.

But satisfying this is automatic, for a duality as defined in Sect. 3.1. That is: For a duality to respect this aspect of symmetry was already built in to our definition of duality: namely in condition (i), that the values are equal between states and quantities that correspond by the duality. Recall Eq. (9), which we here repeat:

$$\langle Q_1, s_1 \rangle_1 = \langle d_{\mathcal{Q}}(Q_1), d_{\mathcal{S}}(s_1) \rangle_2 \ , \quad \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1}. \tag{17}$$

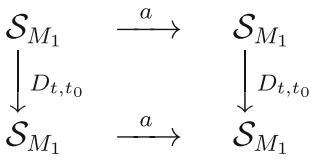

Fig. 7 Commutativity of symmetry and dynamics

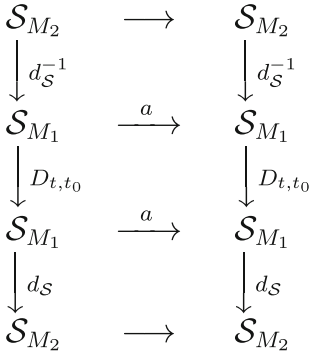

Fig. 8 Commutativity of duality, symmetry, and dynamics

Thus let us show that the map $a_2 := d_{\mathcal{S}} \circ a \circ (d_{\mathcal{S}})^{-1} : \mathcal{S}_{M_2} \to \mathcal{S}_{M_2}$ is a symmetry on states in our sense, i.e. preserves values; cf. Fig. 5. (A similar argument works for the symmetry on quantities, as in Fig. 6.) In effect, we just use the duality maps and Eq. (17) to carry back the discussion of preservation of values to the automorphism $a : \mathcal{S}_{M_1} \to \mathcal{S}_{M_1}$ in model M_1. Thus we consider $\langle Q_2, s_2 \rangle_2$, and we write Q_1, s_1 for the inverse images, under $d_{\mathcal{Q}}$ and $d_{\mathcal{S}}$ respectively, of Q_2, s_2. Then we have (using Eq. (17) for the second and fourth equations, and a being a symmetry for the third equation):

$$\begin{aligned} \langle Q_2, s_2 \rangle_2 &= \langle d_{\mathcal{Q}}(Q_1), d_{\mathcal{S}}(s_1) \rangle_2 = \langle Q_1, s_1 \rangle_1 = \langle Q_1, a(s_1) \rangle_1 \\ &= \langle d_{\mathcal{Q}}(Q_1), d_{\mathcal{S}}(a(s_1)) \rangle_2 = \langle Q_2, a_2(d_{\mathcal{S}}(s_1)) \rangle_2 \equiv \langle Q_2, a_2(s_2) \rangle_2 \,. \end{aligned} \tag{18}$$

Finally: the same verdict—that a duality preserves any symmetry of its model triples—applies to dynamics, i.e. to dynamical symmetries. Recall from footnote 24 (in Sect. 4.3) that a dynamical symmetry is a commutation i.e. equivariance condition. So for the Schrödinger picture of dynamics, the diagram for the 'first' side of a duality, i.e. $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$, is, with a the dynamical symmetry, as in Fig. 7.

So we now compose this diagram with Fig. 5, which represents that a duality preserves a symmetry. But since in Fig. 7, the 'first' side, '1', of the duality occurs twice, on both top and bottom rows, we now need to compose Fig. 7 with Fig. 5 twice: both on its bottom row; and also on its top row (with the duality arrow in Fig. 5 reversed). The resulting diagram (Fig. 8) shows that the duality isomorphism on state-spaces $d_{\mathcal{S}}$ carries the dynamical symmetry a on the '1' side of the duality, to a dynamical symmetry on the '2' side: namely, the symmetry $d_{\mathcal{S}} \circ a \circ d_{\mathcal{S}}^{-1}$ (cf. either the top or bottom square). The Schrödinger picture dynamics on $\mathcal{S}_{M_2}$ is (reading down the columns in the Figure): $d_{\mathcal{S}} \circ D_{t,t_0} \circ d_{\mathcal{S}}^{-1}$.

To sum up: we have shown that a duality always preserves a symmetry of its model triples.

7 Conclusion

In this paper we have investigated the relations between dualities and symmetries, using our Schema for dualities. The Schema begins with a conception of theories and of models that is widespread in philosophy of science (namely, a theory as a formal structure or a set of axioms, and models as instantiations or representations of the theory) but that departs from the common usage, in philosophy of physics, of models as individual solutions of the theory's dynamical equations. The flexibility in our usage allows both a natural definition of a duality as an isomorphism between models *and* the recovery, as a special case, of the more familiar notion of 'model' in philosophy of physics: by appropriately tuning the level of the description—more precisely, by carefully selecting the relevant representations.

In Sect. 2, we emphasised (i) the distinction within a model (in our sense, viz. representation of a bare theory), between model root and specific structure; and (ii) the distinction between internal and external interpretations. This gave us a clear formulation of disagreements like those between Newton (Clarke) and Leibniz, about the status of absolute rest: which we illustrated with Galilean transformations. We also used matrix representations of a group to illustrate the different ways one can define 'model root' and 'specific structure'.

The distinction (i) also underpins the Schema's exact definition of duality as an isomorphism of model roots (Sect. 3). This equipped us, after reviewing symmetries in general (Sect. 4), to understand in detail how dualities and symmetries relate to one another.

This we undertook in the last two Sections. In Sect. 5, the Schema's conceptions of theory and model prompted us to introduce four types of symmetries: of which two are symmetries of the theory (stipulated symmetries and accidental symmetries) and two are symmetries of models only (proper and improper symmetries). These four types are mutually exclusive and jointly exhaustive of all the possibilities for symmetries, given the Schema's distinction between theory and models, and our construal of 'symmetry'. Then in Sect. 6 we showed that, just as one would hope, the Schema's construal of duality meshes with whatever symmetries the duals, i.e. the isomorphic model roots, may have.

Acknowledgements We thank the organizers and audience at the 2017 conference in Hannover; audiences at conferences in Bad Honnef and London in 2018; and Niels Martens, James Read and two anonymous reviewers for comments. SDH's work was supported by the Tarner scholarship in Philosophy of Science and History of Ideas, held at Trinity College, Cambridge.

Appendix: A group automorphism lost under homomorphism

We present a simple case of Sect. 5.2's topic, of a symmetry of a theory being "lost" in a model. As discussed in Sect. 5.2: for the simple case of group theory, this is a matter of the diagram in Fig. 4 not commuting. That is: an automorphism $a_1 : G_1 \to G_1$, and a homomorphism $h : G_1 \to G_2 \cong G_1/\ker h$ need not mesh as the diagram shows, so as to give a well-defined homomorphism $a_2 : G_2 \to G_2$.

To see what is involved, we can identify G_2 with its isomorph $G_1/\ker h$; and let us write K as short for $\ker h$. Then the commutation requirement $h(a_1(g_1)) = a_2(h(g_1))$ is the requirement that the definition of a_2 by saying $a_2(Kg_1) := h(a_1(g_1))$ be well-defined. That is: the given automorphism a_1 must respect the cosets of h. That is: we require that if $h(g_1) = h(g_1')$, i.e. $Kg_1 = Kg_1'$, i.e. there is a $k \in K$ such that $g_1' = kg_1$, then also there is a $k \in K$ such that $a_1(g_1') = ka_1(g_1)$. But a_1 and h need not mesh in this way: commutation of the diagram is not guaranteed.

The smallest counterexample we have found takes G_1 as the dihedral group D_4. It has D_2 as a normal subgroup, $D_2 \lhd D_4$, with $D_4/D_2 \cong C_2 = \{0, 1\}$. So we take G_2 as (up to isomorphism) the cyclic group C_2 and $h : G_1 \to G_2$ as the canonical projection taking each element of D_4 to its D_2-coset. Then we define $a_1 : D_4 \to D_4$ to take an element b of the normal subgroup D_2 *out* of D_2. This implies that the diagram cannot commute. For on the left of the diagram $h : D_4 \ni b \mapsto D_2 = e_{D_4/D_2}$, so that any bottom-row homomorphism a_2 must map $D_2 = e_{D_4/D_2} \mapsto D_2 = e_{D_4/D_2}$; while on the right of the diagram $h : D_4 \ni a_1(b) \mapsto (D_2)a_1(b) \neq e_{D_4/D_2}$.

The details are as follows. D_4 is the symmetry group of the square, and is generated by two elements: rotation out of the plane of the square by π about the axis in the plane that horizontally bisects the square; and rotation in the plane of the square by $\pi/2$ about the axis normal to the plane through the centre of the square. We label these b and c respectively. So with e the identity transformation, $b^2 = c^4 = e$. One checks that bc and bc^3 are the rotations out of the plane of the square by π about the two diagonal axes, and bc^2 is the rotation about the axis in the plane that vertically bisects the square. So these elements are of order 2. Also one checks that $cb = bc^3$, $c^2b = bc^2$, $c^3b = bc$. And so D_4 has eight elements, which we write as $\{e, c, c^2, c^3, b, bc, bc^2, bc^3\}$, with generating equations $b^2 = c^4 = (bc)^2 = e$. D_4 has a copy of the dihedral group D_2—viz. $\{e, c^2, b, bc^2\}$ with generating equations $b^2 = (c^2)^2 = (bc^2)^2 = e$—as a normal subgroup: $D_2 \lhd D_4$, with $D_4/D_2 \cong C_2$. So define $h : D_4 \to D_4/D_2 \equiv \{D_2, D_2(bc)\}$ as the canonical projection. That is: each element of D_2 is sent to $D_2 \in D_4/D_2$; and the other four elements, c, c^3, bc, bc^3, are each sent to their common coset $D_2(bc)$. Finally, we define the automorphism $a_1 : D_4 \to D_4$ as:

(i) the identity map on the rotations in the plane (the powers of c); but also
(ii) mapping b to bc, i.e. $b \in D_2$ is mapped out of D_2, and
(iii) mapping bc to bc^2, bc^2 to bc^3, and bc^3 to $bc^4 \equiv b$ (so that an element out of D_2 is mapped into D_2).

A tedious check shows that a_1 thus defined is an automorphism. So by the argument in the preceding paragraph, the diagram in Fig. 4 does not commute.

References

Auslander, L., & MacKenzie, R. (1963). Introduction to differentiable manifolds. McGraw-Hill: Dover reprint 2007.

Bargmann, V. (1954). On unitary ray representations of continuous groups. *Annals of Mathematics*, *59*(1), 1–46.

Belot, G. (2000). Geometry and motion. *British Journal for the Philosophy of Science*, *51*, 561–595.

Brading, K., & Castellani, E. (Eds.). (2003). *Symmetries in physics philosophical reflections*. Cambridge: Cambridge University Press.

Brown, J. D., & York, J. W, Jr. (1993). Quasilocal energy and conserved charges derived from the gravitational action. *Physical Review D*, *47*(4), 1407–1419.

Butterfield, J. (2006). On symplectic reduction in classical mechanics. In J. Earman and J. Butterfield (eds.) *The handbook of philosophy of physics* (pp. 1–131). North Holland 2006. arXiv:physics/0507194 and http://philsci-archive.pitt.edu/2373/

Butterfield, J. (2014). Reduction, emergence and renormalization. *The Journal of Philosophy*, *111*, 5–49. arXiv:1406.4354. http://philsci-archive.pitt.edu/10762/.

Butterfield, J. (2018). On dualities and equivalences between physical theories. In N. Huggett, & C. Wüthrich (Eds.), *Space and time after quantum gravity* (**Forthcoming**).

Callender, C., & Cohen, J. (2006). There is no special problem about scientific representation. *Theoria*, *21*, 67–85.

Castellani, E. (2017). Duality and particle democracy'. *Studies in History and Philosophy of Modern Physics*, pp. 100–108. https://doi.org/10.1016/j.shpsb.2016.03.002

Caulton, A. (2015). The role of symmetry in the interpretation of physical theories. *Studies in History and Philosophy of Modern Physics*, *52*, 153–162.

De Haro, S. (2015). Dualities and emergent gravity: Gauge/gravity duality. *Studies in History and Philosophy of Modern Physics*, *59*(2017), 109–125. https://doi.org/10.1016/j.shpsb.2015.08.004.PhilSci11666.

De Haro, S. (2016a). Spacetime and physical equivalence. In N. Huggett, & C. Wüthrich (Eds.), *Space and time after quantum gravity*. (**Forthcoming in**). http://philsci-archive.pitt.edu/13243.

De Haro, S. (2016b). Duality and physical equivalence. http://philsci-archive.pitt.edu/id/eprint/12279 (This is an expansion of (2016); and the title has changed).

De Haro, S. (2017). The invisibility of diffeomorphisms. *Foundations of Physics*, *47*(11), 1464–1497.

De Haro, S. (2018a). The heuristic function of duality. *Synthese*. https://doi.org/10.1007/s11229-018-1708-9. arXiv:1801.09095 [physics.hist-ph].

De Haro, S. (2018b). Theoretical equivalence and duality. This volume.

De Haro, S., & Butterfield, J. N. (2018). A schema for duality, illustrated by Bosonization. In J. Kouneiher (Ed.), *Foundations of mathematics and physics one century after Hilbert*. Berlin: Springer.

De Haro, S., Mayerson, D., & Butterfield, J. N. (2016). Conceptual aspects of gauge/gravity duality. *Foundations of Physics*, *46*(11), 1381–1425. https://doi.org/10.1007/s10701-016-0037-4. [arXiv:1509.09231 [physics.hist-ph]].

De Haro, S., Teh, N., & Butterfield, J.N. (2017). Comparing dualities and gauge symmetries. *Studies in History and Philosophy of Modern Physics*, *59*, 68–80. [arXiv:1603.08334 [physics.hist-ph]]. http://philsci-archive.pitt.edu/12009

De Haro, S., & De Regt, H. W. (2018). Interpreting theories without a spacetime. *European Journal for Philosophy of Science*, *8*(3), 631–670. https://doi.org/10.1007/s13194-018-0204-x.

De Haro, S., & De Regt, H. W. (2018a). A precipice below which lies absurdity? Theories without a spacetime and scientific understanding. *Synthese*.https://doi.org/10.1007/s11229-018-1874-9.

Dewar, N. (2015). Symmetries and the philosophy of language. *Studies in History and Philosophy of Modern Physics*, *52*, 317–327.

Dewar, N. (2017). Sophistication about symmetries. *British Journal for Philosophy of Science*, **70**(2), 485–521.

Dieks, D., van Dongen, J., & de Haro, S. (2015). Emergence in holographic scenarios for gravity. *Studies in History and Philosophy of Modern Physics*, *52*, 203–216. https://doi.org/10.1016/j.shpsb.2015.07.007.

Earman, J. (1989). *World enough and spacetime*. Cambridge: MIT Press.

Fine, A. (1984). The natural ontological attitude. In J. Leplin (Ed.), *Scientific realism* (pp. 83–107). Berkeley: University of California Press.

Fine, A. (1986). *The shaky game*. Chicago: University of Chicago Press.

Fraser, D. (2017). Formal and physical equivalence in two cases in contemporary quantum physics. *Studies in History and Philosophy of Modern Physics*, *59*, 30–43.
Glymour, C. (2013). Theoretical equivalence and the semantic view of theories. *Philosophy of Science*, *80*, 286–297.
Holm, D. D. (2011). *Geometric mechanics. part II: rotating, translating and rolling* (2nd ed.). London: Imperial College Press.
Huggett, N. (2017). Target space ≠ space. *Studies in History and Philosophy of Modern Physics*, *59*, 81–88. https://doi.org/10.1016/j.shpsb.2015.08.007.
Kaiser, D. (2005). *Drawing theories apart: The dispersion of Feynman diagrams in postwar physics*. Chicago: University Press.
Lewis, D. (1970). General semantics. *Synthese22*, 18–67. reprinted in his *Philosophical Papers: volume 1* (1983), Oxford University Press.
Lewis, D. (1975). Languages and language. In Keith Gunderson (Ed.), *Minnesota studies in the philosophy of science* (Vol. VII, pp. 3–35). Minneapolis: University of Minnesota Press.
Lutz, S. (2017). What was the syntax-semantics debate in the philosophy of science about? *Philosophy and Phenomenological Research, XCV*, *2*, 319–352.
Maggiore, M. (2005). *A modern introduction to quantum field theory*. Oxford: Oxford University Press.
Matsubara, K. (2013). Realism, underdetermination and string theory dualities. *Synthese*, *190*(3), 471–489.
Møller-Nielsen, T. (2017). Invariance, interpretation and motivation. *Philosophy of Science*, *84*, 1253–1264.
Penrose, R. (1982). Quasi-local mass and angular momentum in general relativity. *Proceedings of the Royal Society of London, A*, *381*, 53–63.
Pooley, O. (2017). Background independence, diffeomorphism invariance, and the meaning of coordinates. In D. Lehmkuhl, G. Schiemann, & E. Scholz (Eds.), *Towards a theory of spacetime theories*. Basel: Birkhäuser.
Read, J. (2016). The interpretation of string-theoretic dualities. *Foundations of Physics*, *46*, 209–235.
Read, J., & Møller-Nielsen, T. (2018). Motivating dualities. *Synthese*. http://philsci-archive.pitt.edu/14663.
Rickles, D. (2017). Dual theories: same but different' or different but same? *Studies in History and Philosophy of Modern Physics*, *59*, 62–67. https://doi.org/10.1016/j.shpsb.2015.09.005.
Szabados, L. B. (2009). Quasi-local energy-momentum and angular momentum in general relativity. *Living Reviews in Relativity*, *12*, 4.
van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
van Fraassen, B. C. (2014). One or two gentle remarks about hans halvorson's critique of the semantic view. *Philosophy of Science*, *81*, 276–283.
Wang, M. -T. (2015). Four lectures on quasi-local mass. arXiv:1510.02931 [math-ph].
Weinberg, S. (1995). *The quantum theory of fields, volume I: Foundations*. Cambridge: Cambridge University Press.