Definable Categorical Equivalence

Laurenz Hudetz*†

This article proposes to explicate theoretical equivalence by supplementing formal equivalence criteria with preservation conditions concerning interpretation. I argue that both the internal structure of models and choices of morphisms are aspects of formalisms that are relevant when it comes to their interpretation. Hence, a formal criterion suitable for being supplemented with preservation conditions concerning interpretation should take these two aspects into account. The two currently most important criteria—generalized definitional equivalence (Morita equivalence) and categorical equivalence—are not optimal in this respect. I put forward a criterion that takes both aspects into account: the criterion of definable categorical equivalence.

1. Introduction. This article puts forward a new criterion of formal equivalence of theories. Roughly speaking, theories are formally equivalent when their formalisms are intertranslatable in the sense that they can be logically or mathematically recovered from each other. There are various explications of this idea. The criterion of definitional equivalence is one of the oldest and most well-known ones. Recently, increasing effort has been devoted to finding criteria of formal equivalence that improve on definitional equivalence: the most important proposals are direct generalizations of definitional equivalence (Andréka, Madarász, and Németi 2008; Barrett and Halvorson 2016) and category-theoretic criteria (Halvorson 2016; Weatherall 2016; Halvor-

Received October 2016; revised January 2018.

*To contact the author, please write to: Department of Philosophy, University of Salzburg, Franziskanergasse 1, 5020 Salzburg, Austria; e-mail: laurenz.hudetz@sbg.ac.at.

†I am indebted to Charlotte Werndl for her invaluable support and advice. I also thank Hajnal Andréka, Thomas Barrett, Alexander Bors, Jeremy Butterfield, Adam Caulton, Erik Curiel, Marton Gömöri, Sam Fletcher, Hannes Leitgeb, Sebastian Lutz, István Németi, Karl-Georg Niebergall, Miklos Redei, Bryan Roberts, and Jim Weatherall as well as two anonymous referees for very helpful feedback and discussions.

Philosophy of Science, 86 (January 2019) pp. 47–75. 0031-8248/2019/8601-0003$10.00
Copyright 2019 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

son and Tsementzis 2017). These new criteria have been applied in several case studies.[1]

A well-known criticism of this project is that no criterion that focuses exclusively on formalisms is able to capture full equivalence of theories, while full equivalence is what ultimately matters (see Coffey 2014). Whether two theories are fully equivalent depends on how their formalisms are empirically and ontologically *interpreted*. Different interpretations lead to different equivalence judgments.

Coffey is right that one should not neglect interpretation when addressing the problem of full theoretical equivalence. However, this point is by no means a valid objection against the project of explicating formal equivalence. This project is not committed to the assumption that formal equivalence is necessary and sufficient for full equivalence of theories. Instead, one can view the formal project as laying the groundwork for (sufficient) criteria of full theoretical equivalence.[2] According to this view, full theoretical equivalence is to formal equivalence what interpreted formalisms are to uninterpreted ones. The idea is that full equivalence of theories (qua interpreted formalisms) obtains if there is a way of translating between their formalisms that preserves empirical content as well as theoretical content beyond the empirical (if there is any). This idea has been put forward, for example, by Putnam (1983, 38–39). In order to turn this idea into a rigorous account, we need (1) a criterion of formal equivalence that is suitable for being supplemented with preservation conditions regarding interpretation, (2) an explication of what it is to endow a formalism with an interpretation, and (3) a rigorous formulation of preservation conditions.

This article addresses the first point (but work devoted to the second and third points is already in progress). In section 2, I highlight two aspects of formalisms that are relevant when it comes to interpreting a formalism. I briefly present the two currently most important criteria of formal equivalence—generalized definitional equivalence (GDE) and categorical equivalence (CE)—and point out that each of them takes into account one such aspect while neglecting the other. In section 3, I introduce the criterion of definable categorical equivalence (DCE), which takes into account both relevant aspects. Section 4 investigates the properties of DCE and how it is related to GDE and CE.

Section 4 contains results suggesting that DCE can also be of interest independently of the motivations concerning interpretation described above. For instance, these results shed new light on the criterion of GDE and an-

1. See, e.g., Andréka and Németi (2014), Barrett, Rosenstock, and Weatherall (2015), Hudetz (2015), Rosenstock and Weatherall (2016), Weatherall (2016), Barrett (2017), Barrett and Halvorson (2017), Teh and Tsementzis (2017), and Lefever and Székely (2018).

2. Putnam (1983, 40) argues that formal equivalence is even necessary for full equivalence.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

swer a technical question about the relationship between CE and GDE posed by Barrett (2018). There is a sense in which DCE builds a bridge between CE and GDE. Thereby, it connects two main strands of research about formal equivalence criteria: the category-theoretic approach and the definability-theoretic approach. The results of this article might be relevant to anyone who is interested in either of these approaches.

2. Two Criteria, Morphisms, and Internal Structure. To understand the new criterion, we first examine the criteria of GDE and CE. These two criteria focus on different aspects of formalisms. GDE focuses on the internal constitution of models, whereas CE focuses on relationships between models. I argue that both of these aspects are relevant when it comes to interpreting a formalism. An adequate criterion should take both into account.

2.1. Generalized Definitional Equivalence. The basic idea behind GDE is to count theories as formally equivalent just in case one can define the vocabulary of each theory within the other theory such that this introduction of vocabulary makes the two sides logically equivalent. This idea can be made precise using only weak presuppositions about the structure of theories. All we need is that a theory T comes with a class of models, $\mathrm{Mod}(T)$, and that its models are $\Sigma(T)$-structures of a signature $\Sigma(T)$.[3] A signature is simply a collection of nonlogical symbols. It is not necessary to identify theories with their classes of models. Hence, this requirement is compatible with both the syntactic and the semantic view of theories. Moreover, signatures do not have to be single sorted and first order. They may just as well be many sorted and higher order.[4]

> **Definition 1.** T_1 is *GD-equivalent* to T_2 iff (if and only if) their classes of models $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$ have stepwise GD-expansions E_1 and E_2 such that $E_1 = E_2$. (For a definition of 'GD-expansion', see app. A.)

In essence, GD-equivalent theories differ only in the choice of primitive vocabulary. Since sort symbols also belong to the vocabulary of a theory, GD-equivalent theories may differ in the sorts of objects they take as primitive. GDE allows that objects of a sort taken as primitive in one theory have to be logically constructed in another theory. Definitional equivalence is the special case of GDE in which theories are required to have the same primitive sorts.

3. So models are taken to be structures in the model-theoretic sense (see Hodges 1993).

4. For details about higher-order signatures, see Awodey (1997, chap. 1, sec. 1), Johnstone (2002, sec. D4.1), or Hudetz (2017).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

For self-contained expositions of the criterion of GDE, see Andréka et al. (2008) or Barrett and Halvorson (2016).[5] Note that Halvorson and Barrett call GDE 'Morita equivalence'.[6] GDE is also closely related to the criterion of general bi-interpretability as described by Hodges (1993, sec. 5.4). There is a conjecture that the two criteria are in fact equivalent. Yet, a proof is still missing.

2.2. Choices of Morphisms as Interpretive Choices. GDE works well in many cases. However, there are cases in which the class of models of a formalism is endowed with additional structure that GDE is not designed to take into account but that is relevant when it comes to possible interpretations. As several authors have pointed out, it can be fruitful to endow a class of models with a collection of morphisms when reconstructing the formalism of a theory (e.g., Halvorson 2016; Weatherall 2016, 2017; Nguyen, Teh, and Wells 2017). Intuitively, morphisms are structure-preserving mappings between models (also called 'transformations' or 'symmetries' in some contexts).

Among other things, this additional structure on a class of models can be used to encode constraints on possible interpretations. Used in this way, a choice of morphisms reflects which features of models are taken to be *meaningful*, that is, intended to play a representational role in possible applications.[7] The idea is that only features invariant under the chosen morphisms count as meaningful.[8] For example, Euclidean geometry has the group of Euclidean isometries as morphisms, whereas affine geometry has the larger group of affine transformations as morphisms. So although distance is meaningful in Euclidean geometry, it is not meaningful in affine geometry since it is not invariant under all affine transformations. However, parallelism is still

5. For applications of GDE to formulations of special relativity, affine geometry, and Newtonian kinematics, see Andréka and Németi (2014), Barrett and Halvorson (2017), and Lefever and Székely (2018).

6. The reason is that GDE is equivalent to the topos-theoretic notion of Morita equivalence for coherent first-order theories (cf. Tsementzis 2017). Coherent theories are *Morita equivalent* iff they have equivalent classifying topoi.

7. Of course, choosing morphisms in not sufficient for interpreting formalisms. It only sets an upper bound on meaningfulness. Formalisms that include morphisms and in which certain formulas are designated as meaningful may be described as 'preinterpreted formalisms'. They are prepared for being interpreted but not endowed with a particular interpretation.

8. I only assume that the choice of morphisms is coupled with interpretive decisions. This does not imply a commitment to a morphisms-first approach, according to which morphisms are prior to the identification of surplus structure (cf. Ismael and van Fraassen 2003), or to an interpretation-first approach, according to which the interpretive identification of surplus structure is prior to the choice of morphisms.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

meaningful in affine geometry. Note that even if two formalisms have exactly the same models, they may differ in which morphisms they include. Also note that nonmeaningful (surplus or auxiliary) features of models can be of great practical usefulness, and it may be hard and even undesirable to eliminate them.

Since GDE does not take into account morphisms, it is unable to capture certain cases of equivalence or inequivalence that are due to choices of morphisms between models. Weatherall (2016) discusses examples of such equivalences. An example regarding inequivalences will be discussed in section 4.3.

So in order to deal with formalisms that are enriched with morphisms, we need a criterion that takes morphisms into account. CE is such a criterion.

2.3. Categorical Equivalence. In contrast to GDE, the criterion of CE compares classes of models endowed with morphisms rather than bare classes of models of theories. A class of models endowed with morphisms forms what is called a 'category'. A category C comprises a class of objects, Ob(C); a class of morphisms (arrows) between these objects, Mor(C); an associative composition operation for morphisms; as well as an identity morphism from each object to itself. Category theory provides a rigorous notion of equivalence of categories, on which the criterion of CE is based. In order to define this notion, we need a few important category-theoretic concepts.

A *functor* from one category to another maps objects to objects and morphisms to morphisms in a way that preserves composition of morphisms and identity morphisms. So functors are morphisms of categories. *Natural transformations*, however, are morphisms of functors, and *natural isomorphisms* are invertible natural transformations.[9] Using these concepts, we can define when categories are equivalent. Roughly speaking, categories are equivalent just in case they are related by functors that are quasi-inverse to each other (and hence forget nothing that matters from a categorical perspective).

Definition 2. C_1 and C_2 are *equivalent categories* iff there are functors

$$F: C_1 \to C_2$$

$$G: C_2 \to C_1$$

such that their composition GF is naturally isomorphic to 1_{C_1}, and FG is naturally isomorphic to 1_{C_2} (where 1_{C_i} is the identity functor on C_i for $1 \leq i \leq 2$).

9. For details about these notions, see any textbook on category theory, e.g., MacLane (1998) or Awodey (2006).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

A closely related notion is that of duality. Two categories are said to be *dual* to each other (also *dually equivalent* or *contravariantly equivalent*) iff one category is equivalent to the opposite of the other category, where the opposite of a category is the result of reversing all of its arrows.

With these concepts at hand, we can state the criterion of CE for theories that come with an associated category of models. If T is such a theory, let 'Cat(T)' denote its category of models.

> **Definition 3.** T_1 is *C-equivalent* to T_2 iff $\mathrm{Cat}(T_1)$ and $\mathrm{Cat}(T_2)$ are equivalent or dually equivalent categories.

Philosophers of physics have applied this criterion in several case studies (see Barrett et al. 2015; Hudetz 2015; Rosenstock and Weatherall 2016; Weatherall 2016; Barrett 2017; Teh and Tsementzis 2017).

2.4. Internal Structure of Models Matters for Representation. The criterion of CE only takes into account how models are related by morphisms and neglects other properties of models. In particular, it does not take into account how models are internally structured. Yet, the internal constitution of models matters when it comes to establishing representation relations between models and their target systems.

Let me illustrate this point using a simple example. It is well known that the category of finite-dimensional vector spaces with linear mappings as morphisms is equivalent to the category that has natural numbers as objects and real $m \times n$-matrices as morphisms. Up to isomorphism, there is exactly one vector space of dimension n. So one replaces all vector spaces of dimension n with the natural number n and linear mappings with corresponding matrices. Differences in how the objects of two equivalent categories are internally structured are irrelevant from a purely category-theoretic perspective. Nonetheless, such differences are by no means irrelevant for all scientific purposes. Whenever vector spaces are used to model features of real-world systems, we cannot simply replace them with natural numbers. Vector spaces possess internal structure: a set of vectors, the operations of vector addition and scalar multiplication, a zero vector (and an underlying field of scalars). In applications, some of these features serve representational purposes. For example, vectors might be used to represent forces acting on objects, and the net force on an object may be represented by a sum of vectors. It is utterly impossible that a natural number n could serve the same representational purposes as a vector space of dimension n. The number n simply lacks the internal structure necessary for fulfilling such representational roles.

These observations suggest two things. First, a formalism intended to serve representational purposes can only be equivalently replaced by another for-

https://doi.org/10.1086/701047 Published online by Cambridge University Press

malism if their models have the same representational capacities. Second, it is a necessary condition for sameness of representational capacities that the models of each formalism have internal structure that is sufficient for recovering the relevant internal structure of the models of the other formalism.

It is hardly surprising that the fact that CE does not take into account the internal structure of models leads to a trivialization of CE in certain cases. It is easy to see that CE collapses to a mere cardinality comparison of model classes when the theories in question have "very few" morphisms. More precisely, any two theories with categories of models that are rigid groupoids (i.e., when every morphism is an isomorphism and no model has nontrivial automorphisms) are C-equivalent as soon as they have the same number of nonisomorphic models. We will see in section 4.1 that this does not hold for DCE.

3. Definable Categorical Equivalence. The upshot so far is this: both choices of morphisms and the internal structure of models are aspects of formalisms that are relevant when it comes to their possible interpretations. Hence, a criterion that is suitable for being supplemented with preservation conditions concerning interpretation should take these aspects into account. GDE and CE are not optimal in this respect because GDE does not take into account choices of morphisms and CE does not take into account the internal structure of models.[10]

In this section, I propose a criterion that is designed to take into account both morphisms and the internal structure of models. The basic idea is to strengthen CE by imposing a definability constraint on equivalence functors. For this reason, the new criterion is called 'definable categorical equivalence'. The definability constraint serves to capture the above-mentioned idea that, if theories are formally equivalent, the models in each formalism should have sufficient internal structure to recover the internal structure of the other formalism's models.

Before laying down the definition of DCE, some preliminary remarks are in order. We consider theories T that come with an associated category $\mathrm{Cat}(T)$ having the models of T as objects. We use the same assumptions about $\mathrm{Mod}(T)$ as in the case of GDE (see sec. 2.1). So models are assumed to be higher-order structures.[11] Construing models of scientific theories as higher-order structures does not lead to a relevant loss of generality (see Hudetz 2017). Models such

10. To be fair, GDE and CE were not designed to take these aspects into account. So their lack of doing so is by no means an objection against these criteria per se. It just speaks against using them for certain purposes.

11. One could easily lift the requirement that the objects are higher-order structures in the usual sense, i.e., structures consisting of a set (or element) for each symbol of a higher-order signature: they could also be higher-order models in other topoi than the topos of sets. But let us keep things simple and classical for now.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

as dynamical systems, differentiable manifolds, stochastic processes, causal graphs, or games (in normal or extensive form) can all be seen as higher-order structures. Moreover, theories with first-order structures as models are subsumed as a special case.

The essential idea of DCE does not depend on this particular choice of representing formalisms of theories. DCE could in principle also be explicated in a different framework in which we can rigorously talk about definability and models. I use the model-theoretic framework because it is very general, comparatively simple, and close to mathematical practice, and many philosophers are familiar with it. Moreover, it makes a comparison with GDE possible, which is also naturally explicated in this framework.

There is only a mild requirement concerning morphisms in $\mathrm{Cat}(T)$ since the type of structures used as models does not necessarily determine what the relevant notion of morphism is (see Awodey 1997, 18; Landry 2007). The only requirement is that $\Sigma(T)$-isomorphic models of T should also be isomorphic in $\mathrm{Cat}(T)$. The idea is that being $\Sigma(T)$-isomorphic is the strictest notion of isomorphism for $\Sigma(T)$-structures. The $\Sigma(T)$-isomorphisms preserve all features of structures expressible in terms of the symbols in $\Sigma(T)$. The notion of $\Sigma(T)$-isomorphism is the standard notion of isomorphism in model theory.[12] It is crucial that the above requirement allows that there are invertible morphisms in $\mathrm{Cat}(T)$, that is, $\mathrm{Cat}(T)$-isomorphisms, that are not $\Sigma(T)$-isomorphisms. For example, gauge transformations serving as extra isomorphisms in the category of models of a potential-based formulation of classical electromagnetism are isomorphisms of this kind (see Weatherall 2016). Moreover, $\mathrm{Cat}(T)$ may also contain noninvertible morphisms such as homomorphisms or embeddings.

I will now define DCE and then unpack the definition step by step.

Definition 4. T_1 is *DC-equivalent* to T_2 iff there are reconstruction functors

$$F: \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$$
$$G: \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$$

constituting a (dual) equivalence between the categories $\mathrm{Cat}(T_1)$ and $\mathrm{Cat}(T_2)$.

So in contrast to CE, the existence of arbitrary functors constituting a (dual) equivalence between categories of models is not sufficient for DCE. Roughly speaking, a reconstruction functor is a functor such that for every model $\mathcal{A}$ of the source theory, the corresponding model $F(\mathcal{A})$ of the target theory can be

12. For a definition of Σ-isomorphism, see any model theory textbook, e.g., Hodges (1993, 5).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

reconstructed up to isomorphism from $\mathcal{A}$. To reconstruct $F(\mathcal{A})$ from $\mathcal{A}$ is to define $F(\mathcal{A})$ on the basis of $\mathcal{A}$. The reconstruction has to be uniform in the sense that the same definitions are used for all models. More precisely:

Definition 5. A functor $F : \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$ is a *reconstruction functor* iff there is a reconstruction manual Φ for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$ such that for every object $\mathcal{A}$ in $\mathrm{Cat}(T_1)$, $F(\mathcal{A})$ is defined by Φ over $\mathcal{A}$, or a $\Sigma(T_2)$-isomorphic copy of $F(\mathcal{A})$ is defined by Φ over $\mathcal{A}$.

The rest of this section is devoted to clarifying what reconstruction manuals are and what it means that a structure is defined over another structure by a reconstruction manual.

Reconstruction manuals tell us how to reconstruct structures of one signature from structures of another signature. They can be viewed as functions mapping the nonlogical symbols of one signature to open formulas over the other signature. The free variables of these associated formulas need to be of the right type, of course. So a reconstruction manual presupposes an association between the types of the two signatures in question. In the general case of higher-order signatures, this can be made precise as follows.

Definition 6. Let Σ_1 and Σ_2 be higher-order signatures. Then α is an *association of Σ_1-types with Σ_2-types* iff α maps every type of Σ_1 to a type of Σ_2 such that:

1. α commutes with the product type constructor: $\alpha(\tau_1 \times \tau_2) = \alpha(\tau_1) \times \alpha(\tau_2)$.
2. α commutes with the power type constructor: $\alpha(\wp(\tau)) = \wp(\alpha(\tau))$.

Using this auxiliary notion, we can make precise what reconstruction manuals are.

Definition 7. Let Σ_1 and Σ_2 be higher-order signatures. Then Φ is a *reconstruction manual for Σ_1 in terms of Σ_2* iff there is an association α of Σ_1-types with Σ_2-types such that:

1. Φ maps every basic sort σ of Σ_1 to a Σ_2-formula Φ_σ with one free variable of type $\alpha(\sigma)$.
2. Φ maps every predicate P in Σ_1 with simple type $\wp(\tau)$ to a Σ_2-formula Φ_P with a free variable of type $\alpha(\tau)$.
3. Φ maps every function symbol f in Σ_1 with input-type τ_1 and output-type τ_2 to a Σ_2-formula Φ_f with two free variables of types $\alpha(\tau_1)$, $\alpha(\tau_2)$.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

4. Φ maps every individual constant c in Σ_1 of basic sort σ to a Σ_2-formula Φ_c with a free variable of type $\alpha(\sigma)$.

With the concept of reconstruction manuals at hand, we can make precise what it means that the models of one theory can be defined from corresponding models of another theory. Definability of models is explained in terms of definability of their components. The latter notion is understood as usual: suppose $\mathcal{A}$ is a Σ-structure and φ an open formula over Σ. Then a set X is *defined by* φ *over* $\mathcal{A}$ just in case X contains exactly those (tuples of) objects that satisfy φ in $\mathcal{A}$ (i.e., $X = \{\bar{a} : \mathcal{A} \vDash \varphi[\bar{a}]\}$). The set defined by φ in $\mathcal{A}$, that is, the extension of φ in $\mathcal{A}$, is denoted by '$\varphi^{\mathcal{A}}$'. Similarly, the extension of a symbol S of Σ in $\mathcal{A}$ is denoted by '$S^{\mathcal{A}}$'. Using this notation, then the next definition explicates when a model is defined over another model by means of a reconstruction manual.

Definition 8. Suppose $\mathcal{A}_1$ is a Σ_1-structure and $\mathcal{A}_2$ is a Σ_2-structure. Then $\mathcal{A}_2$ is *defined by* Φ *over* $\mathcal{A}_1$ iff Φ is a reconstruction manual for Σ_2 in terms of Σ_1 such that the following conditions hold:

1. for every symbol S of Σ_2 that is not an individual constant, $S^{\mathcal{A}_2}$ is defined by Φ_S over $\mathcal{A}_1$ (i.e., $S^{\mathcal{A}_2} = \Phi_S^{\mathcal{A}_1}$).
2. for every individual constant c of Σ_2, $\{c^{\mathcal{A}_2}\}$ is defined by Φ_c over $\mathcal{A}_1$ (i.e., $\{c^{\mathcal{A}_2}\} = \Phi_c^{\mathcal{A}_1}$).

This completes the explication of DCE. Before investigating the properties of DCE, three remarks about possible modification are in order. First, using formulas of higher-order languages in reconstruction manuals reflects a liberal attitude about which kinds of logico-mathematical constructions are admissible to reconstruct models. This liberality ties in with scientific practice. Working scientists do not restrict themselves to first-order constructions. For example, defining states over a C^*-algebra is a higher-order construction: states are defined as certain *functions* from a C^*-algebra to its underlying field $\mathbb{C}$. However, using higher-order languages is not essential to the idea of DCE. So if one wishes to take a less liberal stance on admissible logical constructions for philosophical reasons, one can impose constraints on which kinds of formulas are allowed in reconstruction manuals (e.g., only first-order formulas). This will lead to stricter versions of DCE.

Second, to go beyond formal equivalence, one can impose admissibility constraints on reconstruction manuals and functors that take into consideration how formalisms are interpreted. A natural constraint would be to allow only reconstruction manuals that preserve meaningfulness in the sense that they map meaningful symbols of one theory to meaningful formulas of another theory. Also, constraints regarding empirical content or theoretical

https://doi.org/10.1086/701047 Published online by Cambridge University Press

content beyond the empirical can be brought in at this stage. However, putting forward concrete formulations of preservation conditions for reconstruction functors falls outside the scope of this article. But work devoted to this issue is already in progress.

Third, there are cases in which it is not necessary to reconstruct the entire internal structure of models. It may suffice to reconstruct only those parts of models that play a representational role. Such cases can be captured using the notion of *partial* reconstruction manuals. Roughly speaking, a partial reconstruction manual maps the meaningful formulas of one theory to meaningful formulas of another theory. Explicating DCE with partial reconstruction manuals raises questions about meaningfulness and interpretation that would lead us too far afield. Therefore, we will restrict our attention to the simpler case of full reconstruction manuals in this article.

4. Properties of the New Criterion. To understand the new criterion better, it is useful to examine in which ways it differs from CE and GDE and in which ways it is similar to these criteria.

4.1. DCE Is Strictly Stronger than CE. There are cases of CE in which equivalence functors trivially exist but mutual reconstruction of models is not possible along these functors. In particular, the criterion of DCE does not collapse to a mere cardinality comparison of model classes when it comes to theories with "very few" morphisms (described at the end of sec. 2.4). I will illustrate this point using two examples. First, let us consider a well-known simple toy example of a pathological C-equivalence due to Barrett and Halvorson (2016, theorem 5.7). They specify two single-sorted first-order axiom systems: $T_1 \coloneqq \{\exists_{\sigma_1} ! xx = x\}$ and $T_2 \coloneqq \{\exists_{\sigma_2} ! yy = y\} \cup \{\forall_{\sigma_2} y(Q_0 y \rightarrow Q_i y) : i \in \mathbb{N}\}$ with signatures $\Sigma_1 \coloneqq \{\sigma_1, P_0, P_1, \dots\}$ and $\Sigma_2 \coloneqq \{\sigma_2, Q_0, Q_1, \dots\}$, respectively. Elementary embeddings serve as morphisms in their categories of models. Barrett and Halvorson prove that T_1 and T_2 are C-equivalent but not GD-equivalent. Theory T_2 is stronger than T_1 in virtue of its additional axioms. The reason why T_1 and T_2 are C-equivalent is that they have the same number of nonisomorphic models and their categories of models are rigid groupoids. However, one can show that no functors constituting an equivalence between these categories are reconstruction functors. Hence:

Proposition 1. The theories T_1 and T_2 are not DC-equivalent.

Although this is only a toy example, it shows that, relative to CE, the definability condition built into DCE makes a difference and is therefore not redundant. This observation might seem trivial. Nevertheless, it is important to make in order to address an idea that has recently been put forward in the

https://doi.org/10.1086/701047 Published online by Cambridge University Press

literature. Feintzeig (2017) claims that an adjunction between two categories of models hints at a sense in which the models in one category are definable from the models in the other category. Since every C-equivalence is an adjunction, the same must hold for CE if Feintzeig's claim is right. This idea leads to the question whether CE entails that the models in one category are in some sense definable from the models in the other category. If this were so, then adding a definability condition to CE would be redundant. However, it is not so in general. It might be true in special cases in which all the internal structure of models is fully encoded by the morphisms between them.[13] But there are cases in which CE does not capture or even hint at mutual definability in any way. Here is an example that is more general and more robust than the first one.[14] There is no reasonable sense of definability according to which any two rigid structures whatsoever are definable from each other. Nevertheless, for any two rigid structures $\mathcal{A}_1$ and $\mathcal{A}_2$, there is an equivalence between the categories C_1 and C_2, where C_i has $\mathcal{A}_i$ as its sole object and the trivial identity automorphism of $\mathcal{A}_1$ as its only arrow. In such cases, CE is trivially satisfied even if $\mathcal{A}_1$ and $\mathcal{A}_2$ are not definable from each other in any sense. In contrast, DCE is not necessarily satisfied, as it explicitly takes into account definability.

> **Proposition 2.** There are theories with an up to isomorphism unique rigid model that are C-equivalent but not DC-equivalent.

For instance, take a simple theory describing a social network with an up to isomorphism unique rigid model (a finite graph) and a cosmological model that arises from general relativity by means of conditions singling out a unique space-time manifold with no nontrivial isometries (i.e., where space-time is highly asymmetric). Then these theories are trivially C-equivalent. However, they are not DC-equivalent because a model of general relativity with its rich internal structure is not definable from a finite graph.

Some authors take it as "a definite mark against categorical equivalence as a general standard for equivalence of theories" (Barrett 2018, 636) that C-equivalence does not imply that the models of each theory are somehow

13. In such cases, it would seem plausible that CE should be sufficient for DCE. Yet, it is not clear at all what it exactly means that morphisms fully encode the internal structure of models. However, to explicate this idea seems to be an interesting project.

14. The example by Barrett and Halvorson (2016) is not very robust in the following sense. Although T_1 and T_2 are not GD-equivalent in classical first-order logic, it turns out that they are GD-equivalent when infinitely long formulas with countable conjunctions and disjunctions are allowed in definitions. Thanks to Hajnal Andréka and István Németi for sharing their proof.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

definable in terms of the models of the other. From this perspective, DCE can be seen as an amendment of CE not subject to this concern.

4.2. Proving CE Often Yields a Proof of DCE. We have now seen that DCE is a nontrivial strengthening of CE, ruling out several instances of the latter. However, the definability constraint in the criterion of DCE does not constitute an extreme strengthening either. DCE is not very far from CE in the sense that from many proofs of C-equivalences occurring in practice one can extract proofs of DC-equivalences. The reason is that proofs of (dual) equivalences between concrete categories often establish more than the mere existence of (contravariant) equivalence functors. In many cases, one establishes how to define structures of one kind from structures of another kind and vice versa in order to specify functors witnessing the duality or equivalence in question. In such cases, theorems sometimes state a bit less than their proofs establish.

For example, when proving that the category of Boolean algebras is dual to the category of Stone spaces one defines a Stone space from a given Boolean algebra and vice versa. Another example is Gelfand duality, where one defines a compact Hausdorff space from any given unital commutative C^*-algebra $\mathcal{A}$ (i.e., the space of nonvanishing C^*-algebra homomorphisms from $\mathcal{A}$ to $\mathbb{C}$ with the spectral topology) and one defines a unital commutative C^*-algebra from any given compact Hausdorff space $\mathcal{X}$ (i.e., the algebra of continuous functions from $\mathcal{X}$ to $\mathbb{C}$). Roughly speaking, the definitions involved in specifying the functors constituting Gelfand duality capture a sense in which state spaces and algebras of observables of classical mechanical systems can be reconstructed from each other (Strocchi 2005, 15).

Actually, finding appropriate definitions is often a crucial step in proofs of (dual) equivalences between categories. It is what enables us to specify functors in the first place, and the fact that the so specified functors constitute a (dual) equivalence between the categories in question depends on the judicious choice of definitions. In this sense, the idea of reconstruction underlying DCE ties in with scientific practice.

It follows from the results in the previous section that not all instances of CE have proofs involving definitions in this way. However, I conjecture that many instances of CE that we encounter in practice do, especially those arising from reconstruction or representation theorems. This conjecture suggests to undertake case studies as a task for future research. For example, is the C-equivalence between formulations of Hamiltonian mechanics and Lagrangian mechanics (Barrett 2017) a case of DCE? Does the Gelfand-Naimark representation theorem (Gelfand and Naimark 1943) allow us to prove that the Hilbert space formalism and the C^*-algebra formalism of quantum mechanics are DC-equivalent? Do reconstruction theorems relating different approaches to

https://doi.org/10.1086/701047 Published online by Cambridge University Press

quantum field theories (e.g., Wightman 1956; Osterwalder and Schrader 1973, 1975) give rise to DC-equivalences?

4.3. DCE and GDE Differ in the Presence of Surplus Structure. A major difference between DCE and GDE is that DCE is not overly sensitive to the presence of auxiliary or surplus structure because it takes morphisms into account. This section is devoted to showing that. We put the focus on a particular kind of auxiliary structure, namely, coordinate systems. Although introducing coordinates is an insignificant extension of a formalism as far as representational capacities are concerned, it makes a big difference for GDE: models that are not GD-equivalent can become GD-equivalent when equipped with coordinate systems. The reason is that the presence of a coordinate system allows us to formulate new definitions that could not be formulated otherwise. These newly available definitions can be used to establish GDE even though the definientia involved are not meaningful formulas. This leads to pathological positive cases of GDE. As we will see, DCE is not beset with this problem.

Let us look at an example to illustrate these points. We compare n-dimensional Minkowski geometry (MG_n) and n-dimensional Euclidean geometry (EG_n). The models of MG_n and EG_n are n-dimensional Minkowski spaces and n-dimensional Euclidean spaces, respectively. Both Minkowski spaces and Euclidean spaces are here understood as real vector spaces equipped with generalized inner products (i.e., nondegenerate symmetric bilinear forms).[15] This is a slight simplification. Strictly speaking, Minkowski and Euclidean spaces should be treated as affine spaces—rather than vector spaces—equipped with generalized inner products.[16] But since this would unnecessarily complicate our presentation without affecting any results, the chosen simplification is unproblematic. It is also common in physics to simplify in this way (see Naber 2012).

So we take an n-dimensional Minkowski space $(\mathbf{V}, \eta)$ to consist of an n-dimensional vector space $\mathbf{V} = (V, \oplus, \odot)$ over the field $\mathbf{R}$ of real numbers together with a Minkowskian inner product $\eta : V \times V \to \mathbb{R}$ of metric signature $(1, n - 1)$. An n-dimensional Euclidean space $(\mathbf{V}, \langle \cdot, \cdot \rangle)$ is an n-dimensional real vector space $\mathbf{V}$ endowed with a Euclidean inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ of metric signature $(n, 0)$. Up to isomorphism, there is only one n-dimensional Minkowski space and only one n-dimensional Euclidean space.

15. So Minkowski spaces have a first-order signature comprising two basic sort symbols (for scalars and vectors) as well as various operation symbols (for the field operations, the vector space operations and the generalized inner product). The signature of Euclidean spaces only differs from that of Minkowski spaces in having a different symbol for the inner product.

16. For a detailed exposition of Minkowski and Euclidean spaces as metric affine spaces, see Malament (2009, sec. 2).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

When endowed with a physical interpretation (see Malament 2009, sec. 3.2), four-dimensional Minkowski geometry yields the theory of space-time underlying special relativity. Four-dimensional Euclidean geometry, when physically interpreted, ascribes a very different structure to space-time. But more than that these geometries should count as inequivalent irrespective of concrete physical applications because their generalized inner products are very different from a purely mathematical point of view. They encode very different notions of distance and angles. The criteria of GDE and DCE both capture this difference when the models of Euclidean and Minkowski geometry are not endowed with coordinate systems:

Proposition 3.

1. MG_n and EG_n are not DC-equivalent.
2. MG_n and EG_n are not GD-equivalent.

However, we will now see that if we endow their models with coordinate systems, the resulting versions of Minkowski and Euclidean geometry become GD-equivalent but not DC-equivalent. Endowing models with coordinate systems can hardly be regarded as an illegitimate step. Although coordinate systems may be dispensable for a characterization of fundamental space-time structure or other foundational issues, they still play an important role in practice. Usually, one cannot apply geometry empirically without choosing a reference frame and an associated coordinate system. For a defense of the legitimacy of a coordinate-based approach for foundational work, see Wallace (2016).

We need a bit of terminology to specify the coordinatized models. A *linear coordinate system* for an n-dimensional vector space $\mathbf{V}$ is an n-tuple of coordinate functions $(f^0, \ldots, f^{n-1})$, where $f^i : V \to \mathbb{R}$ maps every vector $v \in V$ to a real number serving as the ith coordinate of v $(0 \leq i < n)$, such that the function $f : V \to \mathbb{R}^n$ defined by $f(v) \coloneqq (f^0(v), \ldots, f^{n-1}(v))$ is a vector space isomorphism from $\mathbf{V}$ to the real coordinate space $\mathbf{R}^n$.[17] We say that a linear coordinate system $(f^0, \ldots, f^{n-1})$ for $\mathbf{V}$ is a *Lorentz coordinate system* for an n-dimensional Minkowski space $\mathcal{M} = (\mathbf{V}, \eta)$ iff f is an isometric isomorphism from $\mathcal{M}$ to the real coordinate space $(\mathbf{R}^n, \eta_{\text{std}})$ with the standard Minkowski inner product η_{std} on $\mathbf{R}^n$, which takes the following form in the standard coordinates x^i on $\mathbb{R}^n$ (mapping elements of $\mathbb{R}^n$ to their ith components):

$$\eta_{\text{std}}(u, v) = x^0(u) \cdot x^0(v) - \ldots - x^{n-1}(u) \cdot x^{n-1}(v). \tag{1}$$

17. Note that choosing a linear coordinate system for a vector space amounts to the same as choosing an ordered basis. Hence, coordinate systems are often simply defined as ordered bases.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

It immediately follows that all Lorentz coordinate systems for $\mathcal{M}$ are related to each other by Lorentz transformations. Analogously, $(f^0, \ldots, f^{n-1})$ is a *Euclidean coordinate system* for an n-dimensional Euclidean space $\mathcal{E} = (\mathbf{V}, \langle \cdot, \cdot \rangle)$ iff f is an isometric isomorphism from $\mathcal{E}$ to the real coordinate space $(\mathbf{R}^n, \langle \cdot, \cdot \rangle_{\text{std}})$ with the standard Euclidean inner product on $\mathbf{R}^n$, which is given by

$$\langle u, v \rangle_{\text{std}} = x^0(u) \cdot x^0(v) + \ldots + x^{n-1}(u) \cdot x^{n-1}(v). \tag{2}$$

Now we can specify the models of Minkowski and Euclidean geometry extended with coordinate systems (MG_n^+ and EG_n^+). The models of MG_n^+ are those structures $(\mathbf{V}, f^0, \ldots, f^{n-1}, \eta)$ such that $(\mathbf{V}, \eta)$ is a model of MG_n and $(f^0, \ldots, f^{n-1})$ is a Lorentz coordinate system for $(\mathbf{V}, \eta)$. The models of EG_n^+ are those structures $(\mathbf{V}, f^0, \ldots, f^{n-1}, \langle \cdot, \cdot \rangle)$ such that $(\mathbf{V}, \langle \cdot, \cdot \rangle)$ is a model of EG_n and $(f^0, \ldots, f^{n-1})$ is a Euclidean coordinate system for $(\mathbf{V}, \langle \cdot, \cdot \rangle)$.[18] With these specifications at hand, one can prove that

Proposition 4. MG_n^+ and EG_n^+ are GD-equivalent (for $n \geq 2$).

To get an impression of what is going on here, consider a model $\mathcal{M} = (\mathbf{V}^+, \eta)$ of MG_n^+, where $\mathbf{V}^+ = (\mathbf{V}, f^0, \ldots, f^{n-1})$. It follows that $\mathcal{M}$ is a definitional expansion of $\mathbf{V}^+$ because $(f^0, \ldots, f^{n-1})$ is a Lorentz coordinate system for $\mathcal{M}$, and hence $\mathcal{M}$ satisfies

$$\eta(u, v) = f^0(u) \cdot f^0(v) - \ldots - f^{n-1}(u) \cdot f^{n-1}(v). \tag{3}$$

Now define a Euclidean inner product $\langle \cdot, \cdot \rangle$ over $\mathbf{V}^+$:

$$\langle u, v \rangle = f^0(u) \cdot f^0(v) + \ldots + f^{n-1}(u) \cdot f^{n-1}(v). \tag{4}$$

This definition can be formulated in the first-order language of n-dimensional vector spaces with coordinate functions. Then $\mathcal{E} \coloneqq (\mathbf{V}^+, \langle \cdot, \cdot \rangle)$ is a model of EG_n^+ and also a definitional expansion of $\mathbf{V}^+$. Hence, $\mathcal{M}$ and $\mathcal{E}$ are definitional expansions of the same underlying structure $\mathbf{V}^+$ and, therefore, have a common definitional expansion $(\mathbf{V}^+, \eta, \langle \cdot, \cdot \rangle)$. Thus, $\mathcal{M}$ and $\mathcal{E}$ are definitionally equivalent.

Let us now see how a different analysis of the relationship between MG_n^+ and EG_n^+ can be achieved by making use of morphisms to wash out auxiliary structure. The crucial point is that choices of coordinates need not be preserved under morphisms. One can change coordinates (e.g., via a Lorentz transformation in the case of MG_n^+), and the model with the new coordinate system will represent exactly the same physical situation. This suggests spec-

18. Note that for every model of MG_n there are infinitely many models of MG_n^+. The same holds for EG_n and EG_n^+.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

ifying the categories of models of MG_n^+ and EG_n^+ in such a way that choices of coordinates may vary under morphisms. In particular, we should not take $\Sigma(MG_n^+)$-isomorphisms and $\Sigma(EG_n^+)$-isomorphisms as the respective morphisms because they preserve all the internal structure of models, including coordinate functions. Furthermore, if we would like to wash out surplus structure that is due to our simplification—that is, the designated origins in Euclidean and Minkowski spaces—now is the opportunity to do so by choosing morphisms that do not preserve origins. This is what we will do. We define $Cat(MG_n^+)$ in such a way that $h : (\mathbf{V}, f^0, ..., f^{n-1}, \eta) \rightarrow (\mathbf{V}', f'^0, ..., f'^{n-1}, \eta')$ is a morphism in $Cat(MG_n^+)$ iff h is a bijection from V to V' that is an affine map from $\mathbf{V}$ to $\mathbf{V}'$ preserving Minkowski inner products (i.e., $\eta(u, v) = \eta'(h(u), h(v))$ for all $u, v \in V$), and for technical reasons we also require that h maps each scalar to itself.[19] The morphisms in $Cat(EG_n^+)$ are specified analogously.

It follows that the automorphism group of each model in $Cat(MG_n^+)$ is the Poincaré group, whereas the automorphism group of models in $Cat(EG_n^+)$ is the Euclidean group. The fact that the Poincaré group and the Euclidean group are not isomorphic implies that

Proposition 5. MG_n^+ and EG_n^+ are not DC-equivalent (for $n \geq 2$).

We can say more when we consider which formulas count as meaningful in MG_n^+ and EG_n^+. On a usual interpretation of these formalisms, formulas with free vector variables are only meaningful when they are invariant under morphisms. We call a formula φ over $\Sigma(MG_n^+)$ with n free variables *invariant under the morphisms of* $Cat(MG_n^+)$ just in case for all morphisms $h : \mathcal{M} \rightarrow \mathcal{M}'$ in $Cat(MG_n^+)$, an n-tuple $(a_1, ..., a_n)$ satisfies φ in $\mathcal{M}$ iff $(h(a_1), ..., h(a_n))$ satisfies φ in $\mathcal{M}'$. It follows that the formula

$$z = f^0(u) \cdot f^0(v) + ... + f^{n-1}(u) \cdot f^{n-1}(v) \tag{5}$$

is not invariant under the morphisms of $Cat(MG_n^+)$.[20] Hence, formula (5) is not meaningful in MG_n^+. So although formula (5) defines a Euclidean inner product over any given model of MG_n^+, this inner product will not count as meaningful in MG_n^+. Thus, any reconstruction manual for $\Sigma(EG_n^+)$ in terms of $\Sigma(MG_n^+)$ yielding (5) as a translation of $\langle u, v \rangle = z$ does not preserve mean-

19. Bijection h is an affine map from a vector space $\mathbf{V} = (V, \oplus, \odot)$ to another $\mathbf{V}' = (V', \oplus', \odot')$ iff there is a vector space isomorphism H from $\mathbf{V}$ to $\mathbf{V}'$ and $w \in V$ such that, for all $v \in V$, $h(v) = H(v) \oplus' w$.

20. In contrast, the formula $z = f^0(u) \cdot f^0(v) - ... - f^{n-1}(u) \cdot f^{n-1}(v)$ is invariant under the morphisms of $Cat(MG_n^+)$, although coordinate functions vary from model to model. This formula is meaningful in MG_n^+. So a formula may be meaningful even if the terms and function symbols occurring in it are not meaningful.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

ingfulness. An analogous conclusion holds for reconstruction manuals for $\Sigma(\mathrm{MG}_n^+)$ in terms of $\Sigma(\mathrm{EG}_n^+)$.

4.4. DCE Is Close to GDE in the Absence of Surplus Structure. We have seen that DCE and GDE sometimes coincide and sometimes do not coincide. This section shows under which conditions the two criteria coincide.

First, one can show that DCE is a generalization of GDE for theories of a certain kind. Barrett and Halvorson (2016) have shown that GD-equivalent first-order theories with elementary embeddings as morphisms are C-equivalent. Their proof can be generalized for the case of theories with higher-order structures as models. Moreover, one can show that the equivalence functors constructed from a GD-equivalence are indeed reconstruction functors (which is not entirely trivial). So we get the following result.

> **Theorem 1.** Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the morphisms in $\mathrm{Cat}(T_i)$ are precisely elementary embeddings $(1 \leq i \leq 2)$. If T_1 and T_2 are GD-equivalent, then T_1 and T_2 are DC-equivalent.

Let us briefly pause here and reflect on what the choice of elementary embeddings as morphisms means for a theory. Elementary embeddings are injective mappings between models that preserve all relations definable in the theory's language. This choice of morphisms reflects the idea that all definable features of models may be taken to be meaningful. (The same holds for $\Sigma(T)$-isomorphisms, which form a subclass of elementary embeddings.) So taking precisely elementary embeddings (or precisely $\Sigma(T)$-isomorphisms) as the morphisms of a theory T means viewing the models of T as not exhibiting any auxiliary or surplus structure.

Under this reading, theorem 1 partially captures the idea that morphisms need to be taken into account only in the presence of auxiliary or surplus structure. This idea furthermore suggests that DCE should not be much wider than GDE when it comes to theories without auxiliary or surplus structure. The following theorem captures a sense in which this is indeed the case. This theorem is one of the main technical results of this article.

> **Theorem 2.** Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the isomorphisms in $\mathrm{Cat}(T_i)$ are precisely the $\Sigma(T_i)$-isomorphisms $(1 \leq i \leq 2)$. If T_1 and T_2 are DC-equivalent, then T_1 and T_2 are GD-equivalent up to external auxiliaries.

The qualification 'up to external auxiliaries' concerns auxiliary vocabulary introduced in GD-expansions rather than auxiliary vocabulary of the theories themselves. Every GD-expansion to a signature with new sorts also

https://doi.org/10.1086/701047 Published online by Cambridge University Press

introduces auxiliary function symbols relating the new sorts to old sorts. GDE up to external auxiliaries is like GDE except that it does not require such "external" auxiliary vocabulary to be recovered. More precisely, T_1 and T_2 are *GD-equivalent up to external auxiliaries* iff they have stepwise GD-expansions E_1 and E_2 that coincide insofar as one ignores external auxiliary vocabulary on both sides; that is, $E_1|_{\Sigma(T_1)\cup\Sigma(T_2)} = E_2|_{\Sigma(T_1)\cup\Sigma(T_2)}$.[21]

So for theories with standard model-theoretic isomorphisms, DCE differs from GDE only in not requiring that auxiliary symbols used for reconstructing the sorts of one theory from the sorts of another theory have to be recovered as well. Hence, DCE is slightly more liberal than GDE in these cases. This raises the question whether and when we should regard recovering external auxiliaries as necessary. I am not aware of any published arguments concerning this question. To shed light on what the requirement of recovering external auxiliaries amounts to, it is useful to find a corresponding strengthening of DCE. It turns out that to obtain GDE simpliciter instead of GDE up to auxiliaries, one only needs to strengthen DCE by adding an extra definability condition regarding isomorphisms between models and their inner models given by composition of reconstruction functors. Call such isomorphisms 'inner isomorphisms'.

Definition 9. T_1 and T_2 are *DC-equivalent with definable inner isomorphisms* iff there are functors $F : \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$, $G : \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$ and reconstruction manuals Φ for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$ and Ψ for $\Sigma(T_1)$ in terms of $\Sigma(T_2)$ establishing a DCE between T_1 and T_2 such that

1. there is a family of $\Sigma(T_1)$-formulas defining a $\Sigma(T_1)$-isomorphism from $\mathcal{A}_1$ to $GF(\mathcal{A}_1)$ for all $\mathcal{A}_1 \in \mathrm{Mod}(T_1)$ and
2. there is a family of $\Sigma(T_2)$-formulas defining a $\Sigma(T_2)$-isomorphism from $\mathcal{A}_2$ to $FG(\mathcal{A}_2)$ for all $\mathcal{A}_2 \in \mathrm{Mod}(T_2)$ such that
3. these isomorphisms are compatible with each other and with Φ and Ψ.

Call 1–3 the 'many-sorted Ahlbrandt-Ziegler condition'. Its underlying idea is closely related to the concept of bi-interpretability due to Ahlbrandt and Ziegler (1986). For a more detailed formulation of this condition, see definition B8, appendix B.

Theorem 3. Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the isomorphisms in $\mathrm{Cat}(T_i)$ are precisely the $\Sigma(T_i)$-isomorphisms ($1 \leq i \leq 2$). If T_1 and T_2 are DC-equivalent with definable inner isomorphisms, then T_1 and T_2 are GD-equivalent.

21. Note that $E|_\Sigma := \{\mathcal{A}|_\Sigma : \mathcal{A} \in E\}$, where $\mathcal{A}|_\Sigma$ is the reduct of $\mathcal{A}$ to Σ (cf. Hodges 1993, 9).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

This shows that the requirement to recover external auxiliaries contained in GDE is closely related to the requirement that inner isomorphisms are definable in a certain way. Moreover, this theorem answers a question posed by Barrett (2018, 637). The question is whether there is a special property $\mathfrak{P}$ of functors such that the existence of a C-equivalence that is constituted by functors having property $\mathfrak{P}$ entails GDE (given the theories in question have precisely elementary embeddings as morphisms). The property of being a reconstruction functor comes close. But to achieve GDE, a monadic property of functors is unlikely to suffice. However, a binary relation between functors works, namely, the relation of being reconstruction functors that jointly satisfy the Ahlbrandt-Ziegler condition.

There are good reasons not to include the many-sorted Ahlbrandt-Ziegler condition as a regular component of the criterion of DCE. When strengthened with this extra condition, it would not be able to capture formal equivalences between theories with categories of models that do not (only) contain standard model-theoretic isomorphisms. For instance, in Weatherall's (2016) examples, different potential-based models $\mathcal{A}_1$, $\mathcal{A}_2$ that are related by a gauge transformation will be mapped to the same model $\mathcal{A}$ by the composite auto-equivalence functor. However, the gauge transformations from $\mathcal{A}_1$ to $\mathcal{A}$ and from $\mathcal{A}_2$ to $\mathcal{A}$ are not induced by uniformly defined bijections in the way the many-sorted Ahlbrandt-Ziegler condition would require (cf. definition B8, app. B). So if we seek to capture equivalences between theories that do not (only) have standard model-theoretic isomorphisms, we should not impose the many-sorted Ahlbrandt-Ziegler condition as necessary.

5. Conclusion. I have argued that choices of morphisms and the internal structure of models both matter when it comes to possible interpretations of a formalism. I have pointed out that GDE does not take into account morphisms and CE does not take into account the internal structure of models. I have presented a new criterion, namely, DCE, which connects the category-theoretic approach and the definability-theoretic approach to formal equivalence. DCE is a strengthening of CE with a definability constraint on equivalence functors. It is also closely related to GDE in cases in which morphisms preserve all definable features of models. DCE takes into account the internal structure of models as well as choices of morphisms. Thereby, it allows a more nuanced analysis of scientific theories.

An important task for future work is to supplement DCE with preservation conditions concerning interpretation by imposing constraints on which reconstruction manuals and functors are admissible. One can specify preservation conditions corresponding to different levels of interpretation: (1) preservation of meaningfulness for preinterpreted formalisms, (2) preservation of empirical content for empirically interpreted formalisms, and (3) preservation of theoretical content for formalisms that are endowed with a full-blown interpretation going beyond the empirical. Furthermore, one can weaken the

https://doi.org/10.1086/701047 Published online by Cambridge University Press

notion of reconstruction functor by allowing partial reconstruction manuals, in which nonmeaningful features of models do not need to be reconstructed. Working out these ideas in detail and exploring their consequences seems to be a promising direction for future research.

Appendix A

GD-Expansions

Suppose Σ and Σ^+ are higher-order signatures with $\Sigma \subset \Sigma^+$ and C and E are classes of Σ-structures and Σ^+-structures, respectively.

Definition A1. E is a *GD-expansion of C to Σ^+ by Δ* iff Δ is a set of explicit definitions and sort-definitions of the symbols in $\Sigma^+ \setminus \Sigma$ in terms of Σ such that

1. all $\mathcal{A} \in C$ satisfy the admissibility conditions of the definitions in Δ;
2. all $\mathcal{A}^+ \in E$ satisfy Δ;
3. and $C = \{\mathcal{A}^+|_\Sigma : \mathcal{A}^+ \in E\}$.

For the form of sort-definitions, see Barrett and Halvorson (2016, sec. 4.1).

Definition A2. E is a *stepwise GD-expansion of C to Σ^+ by* $(\Delta_1, \ldots, \Delta_n)$ iff there is a sequence $(E_0, E_1, \ldots, E_n)$ of classes of structures of signatures $\Sigma_0 \subset \Sigma_1 \subset \ldots \subset \Sigma_n$ such that

1. $\Sigma_0 = \Sigma$ and $\Sigma_n = \Sigma^+$;
2. $E_0 = C$ and $E_n = E$;
3. and E_i is a GD-expansion of E_{i-1} to Σ_i by Δ_i (for $1 \leq i \leq n$).

Appendix B

Proofs

Proposition 1. The theories T_1 and T_2 are not DC-equivalent.

Proof. Note that, up to Σ_2-isomorphism, there is only one model $\mathcal{A}_2^*$ of T_2 such that the predicate Q_0 has a nonempty extension. Suppose for reductio that there are reconstruction functors $F : \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$ and

https://doi.org/10.1086/701047 Published online by Cambridge University Press

$G : \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$ constituting a (dual) equivalence. Then for every $\mathcal{A}_1 \vDash T_1$ there is an $\mathcal{A}_2 \vDash T_2$ such that $\mathcal{A}_2$ is Σ_2-isomorphic to $F(\mathcal{A}_1)$ and $\mathcal{A}_2$ is definable from $\mathcal{A}_1$ by some reconstruction manual Φ. It follows that

(*) there can only be one model $\mathcal{A}_1^* \vDash T_1$ up to Σ_1-isomorphism such that $\Phi_{Q_0}^{\mathcal{A}_1^*} \neq \varnothing$; namely, $\mathcal{A}_1^* \cong G(\mathcal{A}_2^*)$.

The formula Φ_{Q_0} contains only finitely many of the predicates $P_0, P_1, \ldots$ of Σ_1. Therefore, we get many models $\mathcal{A}$ of T_1 such that $\mathcal{A}$ is not Σ_1-isomorphic to $G(\mathcal{A}_2^*)$ but $\Phi_{Q_0}^{\mathcal{A}} \neq \varnothing$. We achieve this by varying the extensions of those P_i not contained in the formula Φ_{Q_0}, while using the same extension as in $G(\mathcal{A}_2^*)$ for the predicates contained in Φ_{Q_0}. Thus, we get a contradiction to (*). QED

Proposition 2. There are theories with an up to isomorphism unique rigid model that are C-equivalent but not DC-equivalent.

Proof. Let T_1 be a theory with an up to isomorphism unique *finite* rigid model (e.g., an asymmetric graph) and let T_2 be a theory with an up to isomorphism unique *infinite* rigid model (e.g., a space-time manifold without nontrivial isometries). Then T_1 and T_2 are trivially C-equivalent. But they are not DC-equivalent. It is sufficient to show that no infinite structure is definable over a finite structure. Let $\mathcal{A}$ be a structure such that every basic type has a finite extension. Note that if τ_1 and τ_2 are types with finite extensions, then so are the product type $\tau_1 \times \tau_2$ and the power types $\wp(\tau_i)$. Thus, every type has a finite extension. Therefore, every $\Sigma(\mathcal{A})$-formula can only have a finite extension under $\mathcal{A}$. Hence, no infinite set is definable over $\mathcal{A}$. QED

Proposition 3.

1. MG_n and EG_n are not DC-equivalent.
2. MG_n and EG_n are not GD-equivalent.

Proof. The $\Sigma(\mathrm{MG}_n)$-automorphism group of the models of MG_n is not isomorphic to the $\Sigma(\mathrm{EG}_n)$-automorphism group of the models of EG_n. Since GDE preserves automorphism groups, MG_n and EG_n cannot be GD-equivalent. For analogous reasons, they cannot be DC-equivalent either. QED

Proposition 4. MG_n^+ and EG_n^+ are GD-equivalent (for $n \geq 2$).

Proof. Let VEC^+ be the set of all structures $(\mathbf{V}, f^0, \ldots, f^{n-1})$ such that $\mathbf{V}$ is an n-dimensional vector space and $(f^0, \ldots, f^{n-1})$ is a linear coordinate

https://doi.org/10.1086/701047 Published online by Cambridge University Press

system for $\mathbf{V}$. It suffices to show that both $\mathrm{Mod}(\mathrm{MG}_n^+)$ and $\mathrm{Mod}(\mathrm{EG}_n^+)$ are GD-expansions of VEC^+. Let $\mathcal{M} = (\mathbf{V}^+, \eta)$ be a model of MG_n^+, where $\mathbf{V}^+ = (\mathbf{V}, f^0, \ldots, f^{n-1})$. Then $(f^0, \ldots, f^{n-1})$ is a Lorentz coordinate system for $\mathcal{M}$. This implies that $\mathcal{M}$ satisfies the definition

$$\eta(u, v) = f^0(u) \cdot f^0(v) - \ldots - f^{n-1}(u) \cdot f^{n-1}(v).$$

Moreover, it follows that $\mathbf{V}^+ \in \mathrm{VEC}^+$. Conversely, if $\mathbf{V}^+ \in \mathrm{VEC}^+$, one can define a Minkowskian inner product η over $\mathbf{V}^+$ using the equation above. Then $(\mathbf{V}^+, \eta) \in \mathrm{Mod}(\mathrm{MG}_n^+)$. Hence, $\mathrm{Mod}(\mathrm{MG}_n^+)$ is a GD-expansion of VEC^+. Analogously, one shows that $\mathrm{Mod}(\mathrm{EG}_n^+)$ is a GD-expansion of VEC^+. QED

Proposition 5. MG_n^+ and EG_n^+ are not DC-equivalent (for $n \geq 2$).

Proof. There cannot be reconstruction functors establishing a (dual) equivalence between $\mathrm{Cat}(\mathrm{MG}_n^+)$ and $\mathrm{Cat}(\mathrm{EG}_n^+)$ because the Poincaré group is not isomorphic to the Euclidean group. QED

Definition B1. A stepwise GD-expansion E of C to Σ^+ by $(\Delta_1, \ldots, \Delta_n)$ *uses only subsorts* iff all basic sorts of Σ^+ are introduced as subsorts of types of Σ by $(\Delta_1, \ldots, \Delta_n)$.

Lemma B1. If T_1 and T_2 are GD-equivalent, there are stepwise GD-expansions E_1 and E_2 of $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$ using only subsorts such that $E_1 = E_2$.

Theorem 1. Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the morphisms in $\mathrm{Cat}(T_i)$ are precisely elementary embeddings $(1 \leq i \leq 2)$. If T_1 and T_2 are GD-equivalent, then T_1 and T_2 are DC-equivalent.

Proof. By lemma B1, $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$ have stepwise GD-expansions E_1 and E_2 to some signature Σ^+ by $(\Delta_1^1, \ldots, \Delta_n^1)$ and $(\Delta_1^2, \ldots, \Delta_m^2)$ using only subsorts such that $E_1 = E_2$. This gives rise to reconstruction manuals. Let Φ map every symbol S of $\Sigma(T_2)$ to the open $\Sigma(T_1)$-formula used to define S in $(\Delta_1^1, \ldots, \Delta_n^1)$. Then Φ is a reconstruction manual for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$. Analogously, we get a reconstruction manual Ψ for $\Sigma(T_1)$ in terms of $\Sigma(T_2)$.

Let $\mathrm{Cat}(E_i)$ be the category having the structures of E_i as objects and elementary embeddings between them as morphisms $(1 \leq i \leq 2)$. Then the reduct functors $P_1 : \mathrm{Cat}(E_1) \to \mathrm{Cat}(T_1)$ and $P_2 : \mathrm{Cat}(E_2) \to \mathrm{Cat}(T_2)$ are equivalence functors, as described in Barrett and Halvorson (2016, sec. 5). By the axiom of choice, there are functors P_1^{-1} and P_2^{-1} essentially inverse to P_1 and P_2. Since $\mathrm{Cat}(E_1) = \mathrm{Cat}(E_2)$, we can define functors $F \coloneqq P_2 P_1^{-1} :$

https://doi.org/10.1086/701047 Published online by Cambridge University Press

$\mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$ and $G \coloneqq P_1 P_2^{-1} : \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$. Functors F and G constitute an equivalence of categories. It is a matter of routine to check that F and G are reconstruction functors with regard to the reconstruction manuals Φ and Ψ. QED

In order to prove theorem 2, we need several lemmas and definitions.

Lemma B2. Suppose α and α' both are associations of Σ_1-types with Σ_2-types satisfying 1 in definition 7 with respect to the same reconstruction manual Φ for Σ_1 in terms of Σ_2. Then, for every type τ of $\Sigma_1 : \alpha(\tau) = \alpha'(\tau)$.

So given a type τ of Σ_1, a reconstruction manual Φ for Σ_1 in terms of Σ_2 associates a unique type $\Phi(\tau)$ of Σ_2 with τ.

Definition B2. Let Φ be a reconstruction manual for Σ^* in terms of Σ, and let T be a theory with signature Σ. Then we say that Φ is *formally admissible with regard to* T_2 iff for every model $\mathcal{A}$ of T

1. for every basic sort σ of Σ^*, Φ_σ defines a nonempty set over $\mathcal{A}$.
2. for every function symbol f of Σ^*, Φ_f defines a function over $\mathcal{A}$.
3. for every individual constant c of Σ^*, Φ_c defines a singleton over $\mathcal{A}$.

From now on suppose that (*a*) T and T^* have disjoint signatures Σ and Σ^* and (*b*) Φ is a reconstruction manual for Σ^* in terms of Σ that is formally admissible with regard to T.

Definition B3. Σ^+ is a Φ-*extension of* Σ *by external auxiliaries* iff Σ^+ contains the symbols of Σ as well as Σ^* and

1. for every basic sort σ in Σ^*, Σ^+ contains exactly one new function symbol i_σ with σ as input type and $\Phi(\sigma)$ as output type.
2. for every type τ of Σ^*, Σ^+ contains exactly one new function symbol j_τ with τ as input type and $\Phi(\tau)$ as output type.

From now on let Σ^+ be a Φ-extension of Σ by external auxiliaries.

Definition B4. Let $\mathcal{A} \in \mathrm{Mod}(T)$. Then we call the Σ^+-structure $\mathcal{A}^+$ with the following properties the *canonical expansion of* $\mathcal{A}$ *to* Σ^+ *induced by* Φ:

1. $\mathcal{A}^+|_\Sigma = \mathcal{A}$.
2. For every symbol S in Σ^* that is not an individual constant: $S^{\mathcal{A}^+} = \Phi_S^{\mathcal{A}}$.
3. For every individual constant c in $\Sigma^* : \{c^{\mathcal{A}^+}\} = \Phi_c^{\mathcal{A}}$.
4. For every basic sort σ in $\Sigma^* : i_\sigma^{\mathcal{A}^+}$ is the identity on $\Phi_\sigma^{\mathcal{A}}$.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

5. For every type τ over Σ^* : $j_\tau^{\mathcal{A}^+}$ is the identity on D_τ, where D_τ is defined as follows:

   a) if τ is a basic sort of Σ^*, then $D_\tau = \Phi_\tau^{\mathcal{A}}$.
   b) if $\tau = \tau_1 \times \tau_2$ is a product type of Σ^*, then $D_\tau = D_{\tau_1} \times D_{\tau_2}$.
   c) if $\tau = \wp(\tau')$ is a power type of Σ^*, then $D_\tau = P(D_{\tau'})$.

Definition B5. The *canonical expansion of Mod(T)* to Σ^+ *induced* by Φ is the isomorphic closure of $\{\mathcal{A}^+ : \mathcal{A} \in \mathrm{Mod}(T)\}$, where $\mathcal{A}^+$ is the canonical expansion of $\mathcal{A}$ to Σ^+ induced by Φ for $\mathcal{A} \in \mathrm{Mod}(T)$.

Henceforth, it is useful to write variables with their type as superscript (e.g., 'x^τ') to avoid confusion.

Lemma B3. The canonical expansion of $\mathrm{Mod}(T)$ to Σ^+ induced by Φ is a stepwise GD-expansion of $\mathrm{Mod}(T)$ to Σ^+ by $(\Delta_\Phi^{\mathrm{sorts}}, \Delta_\Phi^{\mathrm{inj}}, \Delta_\Phi^{\mathrm{const}})$, where

1. $\Delta_\Phi^{\mathrm{sorts}}$ consists of subsort-definitions for all basic sorts σ of Σ^*:

$$\forall\, x^{\Phi(\sigma)}\left(\Phi_\sigma\left(x^{\Phi(\sigma)}\right) \leftrightarrow \exists\, y^\sigma i_\sigma(y^\sigma) = x^{\Phi(\sigma)}\right) \wedge$$
$$\forall\, y_1^\sigma\, \forall\, y_2^\sigma (i_\sigma\,(y_1^\sigma) = i_\sigma\,(y_2^\sigma) \rightarrow y_1^\sigma = y_2^\sigma),$$

2. $\Delta_\Phi^{\mathrm{inj}}$ consists of definitions of the injection symbols j_τ for types τ of Σ^*:

$$\forall\, x^\tau\, \forall\, z^{\Phi(\tau)}\left(j_\tau(x^\tau) = z^{\Phi(\tau)} \leftrightarrow \varphi_\tau\left(x^\tau, z^{\Phi(\tau)}\right)\right),$$

   where the definiens $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is a $\Sigma \cup \Sigma(\Delta_\Phi^{\mathrm{sorts}})$-formula for every type τ of Σ^*, recursively defined as follows:

   a) if τ is a basic sort of Σ^*, then $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is the formula $z^{\Phi(\tau)} = i_\tau(x^\tau)$.
   b) if $\tau = \tau_1 \times \tau_2$ is a product type of Σ^*, then $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is the formula

$$\exists\, x_1^{\tau_1}\, \exists\, x_2^{\tau_2}\left(x^\tau = \langle x_1^{\tau_1}, x_2^{\tau_2}\rangle \wedge \exists\, z_1^{\Phi(\tau_1)}\, \exists\, z_2^{\Phi(\tau_2)}\left(\varphi_{\tau_1}\left(x_1^{\tau_1}, z_1^{\Phi(\tau_1)}\right)\right.\right.$$
$$\left.\left.\wedge\, \varphi_{\tau_2}(x_2^{\tau_2}, z_2^{\Phi(\tau_2)}) \wedge z^{\Phi(\tau)} = \langle z_1^{\Phi(\tau_1)}, z_2^{\Phi(\tau_2)}\rangle\right)\right).$$

   c) if $\tau = \wp(\tau')$ is a power type of Σ^*, then $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is the formula

$$z^{\Phi(\tau)} = \left\{u^{\Phi(\tau')} : \exists\, v^{\tau'}\left(x^\tau(v^{\tau'}) \wedge \varphi_{\tau'}\left(v^{\tau'}, u^{\Phi(\tau')}\right)\right)\right\}.$$

https://doi.org/10.1086/701047 Published online by Cambridge University Press

3. $\Delta_{\Phi}^{\text{const}}$ consists of the following definitions:

   a) $\forall\, x^{\tau}(P(x^{\tau}) \leftrightarrow \Phi_P(j_{\tau}(x^{\tau})))$, for every predicate P in Σ^* of type τ.
   b) $\forall\, x^{\tau_1}\, \forall\, y^{\tau_2}(f(x^{\tau_1}) = y^{\tau_2} \leftrightarrow \Phi_f(j_{\tau_1}(x^{\tau_1}), j_{\tau_2}(y^{\tau_2})))$, for every function symbol f in Σ^* with input-type τ_1 and output-type τ_2.
   c) $\forall\, x^{\sigma}(c = x^{\sigma} \leftrightarrow \Phi_c(j_{\tau}(x^{\tau})))$, for every individual constant c of basic sort σ in Σ^*.

Lemma B4. Suppose that $H : \mathrm{Cat}(T) \to \mathrm{Cat}(T^*)$ is a reconstruction functor with reconstruction manual Φ, the isomorphisms in $\mathrm{Cat}(T^*)$ are precisely the Σ^*-isomorphisms, $\mathcal{A}$ is a model of T, and $\mathcal{A}^+$ is the canonical expansion of $\mathcal{A}$ to Σ^+ induced by Φ. Then $H(\mathcal{A})$ is Σ^*-isomorphic to $\mathcal{A}^+|_{\Sigma^*}$.

Lemma B5. Suppose E is the canonical expansion of $\mathrm{Mod}(T)$ to Σ^+ induced by Φ, $\mathcal{A} \in E$, and $(\mathcal{A}|_{\Sigma})^+$ is the canonical expansion of $\mathcal{A}|_{\Sigma}$ to Σ^+ induced by Φ. Then $\mathcal{A}$ is Σ^+-isomorphic to $(\mathcal{A}|_{\Sigma})^+$.

Theorem 2. Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the isomorphisms in $\mathrm{Cat}(T_i)$ are precisely the $\Sigma(T_i)$-isomorphisms $(1 \le i \le 2)$. If T_1 and T_2 are DC-equivalent, then T_1 and T_2 are GD-equivalent up to external auxiliaries.

Proof. Suppose T_1 and T_2 are DC-equivalent. Let Σ_1 and Σ_2 be their signatures. Then there are reconstruction functors $F : \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$ and $G : \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$ constituting a (dual) equivalence accompanied by reconstruction manuals Φ for Σ_2 over Σ_1 and Ψ for Σ_1 over Σ_2. It follows that Φ and Ψ are formally admissible with regard to T_1 and T_2, respectively. Let Σ_1^+ and Σ_2^+ be Φ- and Ψ-extensions of Σ_1 and Σ_2 by external auxiliaries, respectively. Now let E_1 and E_2 be the canonical expansions of $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$ to Σ_1^+ and Σ_2^+ induced by Φ and Ψ. We know from lemma 3 that E_1 and E_2 are stepwise GD-expansions of $\mathrm{Mod}(T_1)$ and $\mathrm{Mod}(T_2)$, respectively. We show that $E_1|_{\Sigma_1 \cup \Sigma_2} = E_2|_{\Sigma_1 \cup \Sigma_2}$.Suppose $\mathcal{A} \in E_1|_{\Sigma_1 \cup \Sigma_2}$. Then there is a $\mathcal{B} \in E_1$ such that $\mathcal{A} = \mathcal{B}|_{\Sigma_1 \cup \Sigma_2}$. Moreover, $\mathcal{A}|_{\Sigma_1} \in \mathrm{Mod}(T_1)$. Let $(\mathcal{A}|_{\Sigma_1})^+$ be the canonical expansion of $\mathcal{A}|_{\Sigma_1}$ to Σ_1^+ induced by Φ. Check that the following statements hold.

1. $\mathcal{B}$ is Σ_1^+-isomorphic to $(\mathcal{A}|_{\Sigma_1})^+$ (by lemma B5 and $\mathcal{B}|_{\Sigma_1} = \mathcal{A}|_{\Sigma_1}$).
2. $F(\mathcal{A}|_{\Sigma_1})$ is Σ_2-isomorphic to $(\mathcal{A}|_{\Sigma_1})^+|_{\Sigma_2}$ (lemma B4).
3. $\mathcal{A}|_{\Sigma_2}$ is Σ_2-isomorphic to $F(\mathcal{A}|_{\Sigma_1})$ (by 1, 2 and $\mathcal{B}|_{\Sigma_2} = \mathcal{A}|_{\Sigma_2}$).
4. $\mathcal{A}|_{\Sigma_1}$ is Σ_1-isomorphic to $GF(\mathcal{A}|_{\Sigma_1})$ (because F and G constitute a (dual) equivalence and isomorphisms in $\mathrm{Cat}(T_i)$ are $\Sigma(T_i)$-isomorphisms).

https://doi.org/10.1086/701047 Published online by Cambridge University Press

5. $GF(\mathcal{A}|_{\Sigma_1})$ is Σ_1-isomorphic to $F(\mathcal{A}|_{\Sigma_1})^+|_{\Sigma_1}$, where $F(\mathcal{A}|_{\Sigma_1})^+$ is the canonical expansion of $F(\mathcal{A}|_{\Sigma_1})$ induced by Ψ (lemma B4).
6. $\mathcal{A}|_{\Sigma_1}$ is Σ_1-isomorphic to $F(\mathcal{A}|_{\Sigma_1})^+|_{\Sigma_1}$ (by 4 and 5).
7. $\mathcal{A}$ is $\Sigma_1 \cup \Sigma_2$-isomorphic to $F(\mathcal{A}|_{\Sigma_1})^+|_{\Sigma_1 \cup \Sigma_2}$ (by 3 and 6).

Since $F(\mathcal{A}|_{\Sigma_1}) \in \mathrm{Mod}(T_2)$, its canonical expansion $F(\mathcal{A}|_{\Sigma_1})^+$ by Ψ is in E_2. This implies $F(\mathcal{A}|_{\Sigma_1})^+|_{\Sigma_1 \cup \Sigma_2} \in E_2|_{\Sigma_1 \cup \Sigma_2}$. $E_2|_{\Sigma_1 \cup \Sigma_2}$ is closed under $\Sigma_1 \cup \Sigma_2$-isomorphism because E_2 is closed under Σ_2^+-isomorphism. Hence, $\mathcal{A} \in E_2|_{\Sigma_1 \cup \Sigma_2}$. The proof for the other direction is analogous. QED

Definition B6. We say that χ *is a family of formulas defining inner isomorphisms of* T_1*-models with regard to* F, Φ *and* G, Ψ iff

1. χ maps every basic sort σ_1 of $\Sigma(T_1)$ to a $\Sigma(T_1)$-formula $\chi_{\sigma_1}(x^{\sigma_1}, y^{\Phi\Psi(\sigma_1)})$.
2. for every basic sort σ_1 of $\Sigma(T_1)$ and every $\mathcal{A}_1 \in \mathrm{Mod}(T_1)$, $\chi_{\sigma_1}^{\mathcal{A}_1}$ is a bijection from $\sigma_1^{\mathcal{A}_1}$ to $\sigma_1^{GF(\mathcal{A}_1)}$; and these bijections constitute a $\Sigma(T_1)$-isomorphism from $\mathcal{A}_1$ to $GF(\mathcal{A}_1)$.

Definition B7. Let $\Sigma(T_1)^+$ be a Φ-extension of $\Sigma(T_1)$ by external auxiliaries, let E_1 be the canonical expansion of $\mathrm{Mod}(T_1)$ to $\Sigma(T_1)^+$ induced by Φ and let χ be a family of formulas defining inner isomorphisms of T_1-models with regard to F, Φ and G, Ψ. Then the χ*-expansion of Mod*(T_1) to $\Sigma(T_1)^+ \cup \Sigma(T_2)^+$ *induced by* Φ *and* Ψ is the stepwise GD-expansion of E_1 to $\Sigma(T_1)^+ \cup \Sigma(T_2)^+$ by $(\Delta_\Psi^i, \Delta_\Psi^{inj})$ where Δ_Ψ^{inj} is as in lemma B3 and Δ_Ψ^i comprises definitions of the following form for all basic sorts σ_1 of $\Sigma(T_1)$:

$$\forall\, x^{\sigma_1}\ \forall\, y^{\Psi(\sigma_1)}\left(i_{\sigma_1}(x^{\sigma_1}) = y^{\Psi(\sigma_1)} \leftrightarrow \chi_{\sigma_1}\left(x^{\sigma_1}, j_{\Psi(\sigma_1)}\left(y^{\Psi(\sigma_1)}\right)\right)\right) \qquad (\mathrm{Def}_{i_{\sigma_1}})$$

Definition B8. T_1 and T_2 are *DC-equivalent with definable inner isomorphisms* iff there are functors $F : \mathrm{Cat}(T_1) \to \mathrm{Cat}(T_2)$, $G : \mathrm{Cat}(T_2) \to \mathrm{Cat}(T_1)$ and reconstruction manuals Φ for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$ and Ψ for $\Sigma(T_1)$ in terms of $\Sigma(T_2)$ establishing a DCE between T_1 and T_2 such that the many-sorted Ahlbrandt-Ziegler condition holds:

1. there is a family of formulas χ defining inner isomorphisms of T_1-models with regard to F, Φ and G, Ψ and
2. there is a family of formulas $\tilde{\chi}$ defining inner isomorphisms of T_2-models with regard to G, Ψ and F, Φ such that
3. these isomorphisms are compatible with each other and with Φ and Ψ in the following sense:

https://doi.org/10.1086/701047 Published online by Cambridge University Press

a) for every model $\mathcal{B}_1$ of the χ-expansion of $\mathrm{Mod}(T_1)$ to $\Sigma(T_1)^+ \cup \Sigma(T_2)^+$ induced by Φ and Ψ and all basic sorts σ_2 of $\Sigma(T_2)$: $j^{\mathcal{B}_1}_{\Phi(\sigma_2)} \circ i^{\mathcal{B}_1}_{\sigma_2} = \tilde{\chi}^{\mathcal{B}_1}_{\sigma_2}$.
b) for every model $\mathcal{B}_2$ of the $\tilde{\chi}$-expansion of $\mathrm{Mod}(T_2)$ to $\Sigma(T_2)^+ \cup \Sigma(T_1)^+$ induced by Φ and Ψ and all basic sorts σ_1 of $\Sigma(T_1)$: $j^{\mathcal{B}_2}_{\Psi(\sigma_1)} \circ i^{\mathcal{B}_2}_{\sigma_1} = \chi^{\mathcal{B}_2}_{\sigma_1}$.

Theorem 3. Suppose T_1 and T_2 are theories with disjoint signatures, and suppose the isomorphisms in $\mathrm{Cat}(T_i)$ are precisely the $\Sigma(T_i)$-isomorphisms ($1 \leq i \leq 2$). If T_1 and T_2 are DC-equivalent with definable inner isomorphisms, then T_1 and T_2 are GD-equivalent.

Proof. Let Σ_1, Σ_2 be the signatures of T_1, T_2. There are functors F, G, reconstruction manuals Φ, Ψ, and stepwise GD-expansions E_1, E_2 to extended signatures Σ_1^+, Σ_2^+ as in the proof of theorem 2, and we know that $E_1|_{\Sigma_1 \cup \Sigma_2} = E_2|_{\Sigma_1 \cup \Sigma_2}$. Moreover, there are families of formulas χ and $\tilde{\chi}$ defining inner isomorphisms of T_1-models and T_2-models, respectively.

Note that, for every basic sort σ_1 and every type τ_1 of Σ_1, the auxiliary function symbols i_{σ_1} and j_{τ_1} are in Σ_2^+ but not in Σ_1^+. Analogously, for every basic sort σ_2 and every type τ_2 of Σ_2, the auxiliary function symbols i_{σ_2} and j_{τ_2} are in Σ_1^+ but not in Σ_2^+. Moreover, that is the only difference between Σ_1^+ and Σ_2^+.

Let E_1^+ be the χ-expansion of $\mathrm{Mod}(T_1)$ to $\Sigma_1^+ \cup \Sigma_2^+$ induced by Φ and Ψ, and let E_2^+ be the $\tilde{\chi}$-expansion of $\mathrm{Mod}(T_2)$ to the same signature. By definition, E_i^+ is a stepwise GD-expansion of E_i ($1 \leq i \leq 2$). We have to show that $E_1^+ = E_2^+$. It is sufficient if $E_2^+ \vDash \mathrm{Def}\, i_{\sigma_1}$ for σ_1 in Σ_1 and $E_1^+ \vDash \mathrm{Def}\, i_{\sigma_2}$ for σ_2 in Σ_2 because the definitions of j_{τ_1} and j_{τ_2} are the same for E_1^+ and E_2^+. That $E_2^+ \vDash \mathrm{Def}\, i_{\sigma_1}$ and $E_1^+ \vDash \mathrm{Def}\, i_{\sigma_2}$ follows directly from the many-sorted Ahlbrandt-Ziegler condition. QED

REFERENCES

Ahlbrandt, G., and M. Ziegler. 1986. "Quasi-Finitely Axiomatizable Totally Categorical Theories." *Annals of Pure and Applied Logic* 30:63–83.
Andréka, H., J. Madarász, and I. Németi. 2008. "Defining New Universes in Many-Sorted Logic." Unpublished manuscript, Renyi Institute of Mathematics, Hungarian Academy of Sciences.
Andréka, H., and I. Németi. 2014. "Comparing Theories: The Dynamics of Changing Vocabulary." In *Johan van Benthem on Logic and Information Dynamics*, ed. S. Smets and A. Baltag, 143–72. Cham: Springer.
Awodey, S. 1997. "Logic in Topoi: Functorial Semantics for Higher-Order Logic." PhD diss., University of Chicago.
———. 2006. *Category Theory*. Oxford: Oxford University Press.
Barrett, T. 2017. "Equivalent and Inequivalent Formulations of Classical Mechanics." *British Journal for the Philosophy of Science*, forthcoming.
———. 2018. "What Do Symmetries Tell Us about Structure?" *Philosophy of Science* 85 (4): 617–39.

https://doi.org/10.1086/701047 Published online by Cambridge University Press

Barrett, T., and H. Halvorson. 2016. "Morita Equivalence." *Review of Symbolic Logic* 9 (3): 556–82.
———. 2017. "From Geometry to Conceptual Relativity." *Erkenntnis* 82:1043–63.
Barrett, T., S. Rosenstock, and J. Weatherall. 2015. "On Einstein Algebras and Relativistic Spacetimes." *Studies in History and Philosophy of Modern Physics* B 52:309–16.
Coffey, K. 2014. "Theoretical Equivalence as Interpretive Equivalence." *British Journal for the Philosophy of Science* 65 (4): 821–44.
Feintzeig, B. 2017. "Deduction and Definability in Infinite Statistical Systems." *Synthese*, forthcoming.
Gelfand, I., and M. Naimark. 1943. "On the Imbedding of Normed Rings into the Ring of Operators in Hilbert Space." *Matematiceskij sbornik* 54 (2): 197–217.
Halvorson, H. 2016. "Scientific Theories." In *Oxford Handbook of the Philosophy of Science*. Oxford: Oxford University Press.
Halvorson, H., and D. Tsementzis. 2017. "Categories of Scientific Theories." In *Categories for the Working Philosopher*, ed. E. Landry. Oxford: Oxford University Press.
Hodges, W. 1993. *Model Theory*. Cambridge: Cambridge University Press.
Hudetz, L. 2015. "Linear Structures, Causal Sets and Topology." *Studies in History and Philosophy of Modern Physics* 52:294–308.
———. 2017. "The Semantic View of Theories and Higher-Order Languages." *Synthese*, forthcoming.
Ismael, J., and B. C. van Fraassen. 2003. "Symmetry as a Guide to Superfluous Theoretical Structure." In *Symmetries in Physics: Philosophical Reflections*, ed. K. Brading and E. Castellani, 371–92. Cambridge: Cambridge University Press.
Johnstone, P. 2002. *Sketches of an Elephant: A Topos Theory Compendium*. Vols. 1 and 2. Oxford: Oxford University Press.
Landry, E. 2007. "Shared Structure Need Not Be Shared Set-Structure." *Synthese* 158 (1): 1–17.
Lefever, K., and G. Székely. 2018. "Comparing Classical and Relativistic Kinematics in First-Order Logic." *Logique et Analyse* 61 (241): 57–117.
MacLane, S. 1998. *Categories for the Working Mathematician*. 2nd ed. New York: Springer.
Malament, D. 2009. "Geometry and Spacetime." Lecture notes, University of California, Irvine.
Naber, G. L. 2012. *The Geometry of Minkowski Spacetime: An Introduction to the Mathematics of the Special Theory of Relativity*. New York: Springer.
Nguyen, J., N. J. Teh, and L. Wells. 2017. "Why Surplus Structure Is Not Superfluous." *British Journal for the Philosophy of Science*, forthcoming.
Osterwalder, K., and R. Schrader. 1973. "Axioms for Euclidean Green's Functions." *Communications in Mathematical Physics* 31 (2): 83–112.
———. 1975. "Axioms for Euclidean Green's Functions." Pt. 2. *Communications in Mathematical Physics* 42 (3): 281–305.
Putnam, H. 1983. "Equivalence." In *Philosophical Papers*, vol. 3, *Realism and Reason*, ed. H. Putnam, 26–45. Cambridge: Cambridge University Press.
Rosenstock, S., and J. O. Weatherall. 2016. "A Categorical Equivalence between Generalized Holonomy Maps on a Connected Manifold and Principal Connections on Bundles over That Manifold." *Journal of Mathematical Physics* 57:102902.
Strocchi, F. 2005. *An Introduction to the Mathematical Structure of Quantum Mechanics: A Short Course for Mathematicians*. Singapore: World Scientific.
Teh, N., and D. Tsementzis. 2017. "Theoretical Equivalence in Classical Mechanics and Its Relationship to Duality." *Studies in History and Philosophy of Science* B 59:44–54.
Tsementzis, D. 2017. "A Syntactic Characterization of Morita Equivalence." *Journal of Symbolic Logic* 82 (4): 1181–98.
Wallace, D. 2016. "Who's Afraid of Coordinate Systems? An Essay on the Representation of Spacetime Structure." *Studies in the History and Philosophy of Modern Physics*, forthcoming.
Weatherall, J. O. 2016. "Are Newtonian Gravitation and Geometrized Newtonian Gravitation Theoretically Equivalent?" *Erkenntnis* 81 (5): 1073–91.
———. 2017. "Categories and the Foundations of Classical Field Theories." In *Categories for the Working Philosopher*, ed. E. Landry. Oxford: Oxford University Press.
Wightman, A. S. 1956. "Quantum Field Theory in Terms of Vacuum Expectation Values." *Physical Review* 101:860–66.

https://doi.org/10.1086/701047 Published online by Cambridge University Press