Definable Categorical Equivalence

Abstract

This paper aims to provide a new criterion of formal equivalence of theories that is suitable for being supplemented with preservation conditions concerning interpretation. I argue that both the internal structure of models and choices of morphisms between models are aspects of formalisms that are relevant when it comes to their interpretations. So an adequate criterion should take these aspects into account. The two most important criteria presently discussed in philosophy of science—generalised definitional equivalence (Morita equivalence) and categorical equivalence—are not optimal in this respect. Generalised definitional equivalence neglects choices of morphisms, whereas categorical equivalence neglects the internal structure of models. I put forward the new criterion of definable categorical equivalence which takes into account both aspects and connects the category-theoretic approach and the definability-theoretic approach to formal equivalence.
This paper 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 generalisations of definitional equivalence (Andráska, Madarász, and Németi, 2008; Barrett and Halvorson, 2016) and category-theoretic criteria (Halvorson, 2016; Halvorson and Tsementzis, 2017; Weatherall, 2016a). 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 and full equivalence is what ultimately matters (cf. Coffey, 2014). Whether two theories are fully equivalent depends on how their formalisms are empirically and ontologically interpreted. Different interpretations lead to different equivalence judgements.

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. 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, e.g., by Putnam (1983, pp. 38–39).² 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 paper addresses the first point (but work devoted to the second and third point 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—generalised defini-
tional 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 example, these results shed new light on GDE and answer 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. DCE might be of interest for anyone engage with 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 Generalised 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, \( \text{Mod}(T) \), and that its models are \( \Sigma(T) \)-structures of a signature \( \Sigma(T) \).\(^3\) A signature is simply a collection of non-logical 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.\(^4\)

\textbf{Definition 1.} \( T_1 \) is GD-equivalent to \( T_2 \) iff their classes of models \( \text{Mod}(T_1) \) and \( \text{Mod}(T_2) \) have stepwise GD-expansions \( E_1 \) and \( E_2 \) such that \( E_1 = E_2 \).

(For a definition of ‘GD-expansion’ see Appendix A.)

\(^3\)So models are taken to be structures in the model-theoretic sense (see Hodges, 1993).
\(^4\)For details about higher-order signatures see Hudetz (2017), Johnstone (2002, Section D4.1) or Awodey (1997, Chapter I, Section 1).
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 where theories are required to have the same primitive sorts.

For self-contained expositions of the criterion of GDE see Barrett and Halvorson (2016) or Andréka, Madarász, and Németi (2008). Note that Halvorson and Barrett call GDE ‘Morita equivalence’.

### 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 (cf. Halvorson, 2016; Weatherall, 2016a,b; 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, i.e. intended to play a representational role in possible applications.\(^5\) The idea is that only features invariant under the chosen morphisms count as meaningful.\(^6\) 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 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 non-meaningful (surplus or auxiliary) features of models can be of great practical usefulness and it may be hard and even undesirable to eliminate them.

---

\(^5\) Of course, choosing morphisms in not sufficient for interpreting a 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 ‘pre-interpreted formalisms’. They are prepared for being interpreted but not endowed with a particular interpretation.

\(^6\) I only assume that the choice of morphisms is coupled with interpretive decisions. This implies neither 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), nor to an interpretation first approach, according to which the interpretive identification of surplus structure is prior to the choice of morphisms.
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 (2016a) 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 $\mathcal{C}$ comprises a class of objects, $\text{Ob}(\mathcal{C})$, a class of morphisms (arrows) between these objects, $\text{Mor}(\mathcal{C})$, 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, on the other hand, are morphisms of functors and natural isomorphisms are invertible natural transformations. 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. $\mathcal{C}_1$ and $\mathcal{C}_2$ are equivalent categories iff there are functors $F : \mathcal{C}_1 \rightarrow \mathcal{C}_2$, $G : \mathcal{C}_2 \rightarrow \mathcal{C}_1$ such that their composition $GF$ is naturally isomorphic to $1_{\mathcal{C}_1}$ and $FG$ is naturally isomorphic to $1_{\mathcal{C}_2}$ (where $1_{\mathcal{C}_i}$ is the identity functor on $\mathcal{C}_i$ for $1 \leq i \leq 2$).

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 $\text{Cat}(T)$ denote its category of models.

---

7For details about these notions see any textbook on category theory, e.g. MacLane (1998) or Awodey (2006).
Definition 3. $T_1$ is C-equivalent to $T_2$ iff $\text{Cat}(T_1)$ and $\text{Cat}(T_2)$ are equivalent or dually equivalent categories.

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

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 upon 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 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 formalism 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 trivialisation 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. where every morphism is an isomorphism and no model has non-trivial automorphisms) are C-equivalent as soon as they have the same number of non-isomorphic 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.\(^8\)

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 \(\text{Cat}(T)\) having the models of \(T\) as objects. We use the same assumptions about \(\text{Mod}(T)\) as in the case of GDE (see Section 2.1). So models are assumed to be higher-order structures.\(^9\) Construing models of scientific theories as higher-order structures does not lead to a relevant loss of generality (see Hudetz, 2017). Models such 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

---

\(^8\)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.

\(^9\)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.
simple, 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 \( \text{Cat}(T) \) since the type of structures used as models does not necessarily determine what the relevant notion of morphism is (cf. Awodey, 1997, p. 18; Landry, 2007). The only requirement is that \( \Sigma(T) \)-isomorphic models of \( T \) should also be isomorphic in \( \text{Cat}(T) \). The idea is that being \( \Sigma(T) \)-isomorphic is the strictest notion of isomorphism for \( \Sigma(T) \)-structures. \( \Sigma(T) \)-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.\(^{10}\) It is crucial that the above requirement allows that there are invertible morphisms in \( \text{Cat}(T) \), i.e. \( \text{Cat}(T) \)-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 (cf. Weatherall, 2016a). Moreover, \( \text{Cat}(T) \) may also contain non-invertible 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

\[
F : \text{Cat}(T_1) \rightarrow \text{Cat}(T_2)
\]

\[
G : \text{Cat}(T_2) \rightarrow \text{Cat}(T_1)
\]

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

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

\(^{10}\)For a definition of \( \Sigma \)-isomorphism, see any model theory textbook, e.g. Hodges (1993, p. 5).
Reconstruction manuals tell us how to reconstruct structures of one signature from structures of another signature. They can be viewed as functions mapping the non-logical 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 $\Sigma_1$ and $\Sigma_2$ be higher-order signatures.

$\alpha$ is an **association of $\Sigma_1$-types with $\Sigma_2$-types** iff $\alpha$ maps every type of $\Sigma_1$ to a type of $\Sigma_2$ such that:

1. $\alpha$ commutes with the product type constructor: $\alpha(\tau_1 \times \tau_2) = \alpha(\tau_1) \times \alpha(\tau_2)$.
2. $\alpha$ 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 $\Sigma_1$ and $\Sigma_2$ be higher-order signatures.

$\Phi$ is a **reconstruction manual for $\Sigma_1$ in terms of $\Sigma_2$** iff there is an association $\alpha$ of $\Sigma_1$-types with $\Sigma_2$-types such that:

1. $\Phi$ maps every basic sort $\sigma$ of $\Sigma_1$ to a $\Sigma_2$-formula $\Phi_\sigma$ with one free variable of type $\alpha(\sigma)$.
2. $\Phi$ maps every predicate $P$ in $\Sigma_1$ with simple type $\wp(\tau)$ to a $\Sigma_2$-formula $\Phi_P$ with a free variable of type $\alpha(\tau)$.
3. $\Phi$ maps every function symbol $f$ in $\Sigma_1$ with input-type $\tau_1$ and output-type $\tau_2$ to a $\Sigma_2$-formula $\Phi_f$ with two free variables of types $\alpha(\tau_1), \alpha(\tau_2)$.
4. $\Phi$ maps every individual constant $c$ in $\Sigma_1$ of basic sort $\sigma$ to a $\Sigma_2$-formula $\Phi_c$ with a free variable of type $\alpha(\sigma)$.

With 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 $A$ is a $\Sigma$-structure and $\varphi$ an open formula over $\Sigma$. Then a set $X$ is **defined by $\varphi$ over $A$** just in case $X$ contains exactly those (tuples of) objects that satisfy $\varphi$ in $A$, i.e. $X = \{ \bar{a} : A \models \varphi[\bar{a}] \}$. The set defined by $\varphi$ in $A$, i.e. the extension of $\varphi$ in $A$, is denoted by $\langle \varphi \rangle^A$. Similarly, the extension of a symbol $S$ of $\Sigma$ in $A$ is denoted by $\langle S \rangle^A$. Using these notations, then 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 $\Sigma_1$-structure and $\mathcal{A}_2$ is a $\Sigma_2$-structure. $\mathcal{A}_2$ is defined by $\Phi$ over $\mathcal{A}_1$ iff $\Phi$ is a reconstruction manual for $\Sigma_2$ in terms of $\Sigma_1$ such that the following conditions hold:

1. for every symbol $S$ of $\Sigma_2$ that is not an individual constant, $S^{\mathcal{A}_2}$ is defined by $\Phi_S$ over $\mathcal{A}_1$ (i.e. $S^{\mathcal{A}_2} = \Phi_S^{\mathcal{A}_1}$).

2. for every individual constant $c$ of $\Sigma_2$, $\{c^{\mathcal{A}_2}\}$ is defined by $\Phi_c$ 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 liberal attitude 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 $C$. 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 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 paper. 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 paper.
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 close them.

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 Section 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: $T_1 := \{\exists x_1 \, xx = x\}$ and $T_2 := \{\exists x_2 \, yy = y\} \cup \{\forall x_2 \, (Q_0 y \rightarrow Q_i y) : i \in \mathbb{N}\}$ with signatures $\Sigma_1 := \{\sigma_1, P_0, P_1, \ldots\}$ and $\Sigma_2 := \{\sigma_2, Q_0, Q_1, \ldots\}$, 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. $T_2$ 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 non-isomorphic 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 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 CE is stronger than 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.\(^{11}\) 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

\(^{11}\)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.
general and more robust\textsuperscript{12} than the first one. 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 $A_1$ and $A_2$, there is an equivalence between the categories $C_1$ and $C_2$, where $C_i$ has $A_i$ as its sole object and the trivial identity automorphism of $A_i$ as its only arrow. In such cases, CE is trivially satisfied even if $A_1$ and $A_2$ are not definable from each other in any sense. In contrast, DCE is not necessarily satisfied as it explicitly takes into account definability.

\textbf{Proposition 2.} \textit{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 spacetime manifold with no non-trivial isometries (i.e. where spacetime 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, p. 13) that C-equivalence does not imply that the models of each theory are somehow 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.

\subsection*{4.2 Proving CE often yields a proof of DCE}

We have now seen that DCE is a non-trivial 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 \textit{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

\footnote{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 Andr\'eka and István Németi for sharing their proof.}
algebra and vice versa. Another example is Gelfand duality, where one defines
a compact Hausdorff space from any given unital commutative $C^*$-algebra $\mathcal{A}$
(namely the space of non-vanishing $C^*$-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 $\mathcal{X}$ (namely 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, p. 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 be-
tween 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 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 a particular
kind of auxiliary structure, namely coordinate systems. Although introducing co-
ordinates 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 sys-
tems. The reason is that the presence of a coordinate system allows 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 ($\text{MG}_n$) and $n$-dimensional Euclidean geometry ($\text{EG}_n$). The models of $\text{MG}_n$ and $\text{EG}_n$ 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 generalised inner products (i.e. non-degenerate symmetric bilinear forms). This is a slight simplification. Strictly speaking, Minkowski and Euclidean spaces should be treated as affine spaces—rather than vector spaces—equipped with generalised inner products. 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 (cf. Naber, 2012).

So we take an $n$-dimensional Minkowski space $(V, \eta)$ to consist of an $n$-dimensional vector space $V = (V, \oplus, \odot)$ over the field $\mathbb{R}$ 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 $(V, \langle \cdot, \cdot \rangle)$ is an $n$-dimensional real vector space $V$ 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.

When endowed with a physical interpretation (cf. Malament, 2009, Section 3.2), 4-dimensional Minkowski geometry yields the theory of spacetime underlying special relativity. 4-dimensional Euclidean geometry, when physically interpreted, ascribes a very different structure to spacetime. But more than that: these geometries should count as inequivalent irrespective of concrete physical applications because their generalised inner products are very different from a 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. $\text{MG}_n$ and $\text{EG}_n$ are not DC-equivalent.
2. $\text{MG}_n$ and $\text{EG}_n$ 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 characterisation of fundamental spacetime structure or other

---

13 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 generalised inner product). The signature of Euclidean spaces only differs from that of Minkowski spaces in having a different symbol for the inner product.

14 For a detailed exposition of Minkowski and Euclidean spaces as metric affine spaces see Malament (2009, Section 2).
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 defence of the legitimacy of a coordinate-based approach for foundational work see Wallace (2016).

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

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

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 systems for an $n$-dimensional Euclidean space $\mathcal{E} = (V, \langle \cdot, \cdot \rangle)$ iff $f$ is an isometric isomorphism from $\mathcal{E}$ to the real coordinate space $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\text{std}})$ with the standard Euclidean inner product on $\mathbb{R}^n$, which is given by:

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

Now we can specify the models of Minkowski and Euclidean geometry extended with coordinate systems ($\mathcal{M}_n^+$ and $\mathcal{E}_n^+$). The models of $\mathcal{M}_n^+$ are those structures $(V, f^0, \ldots, f^{n-1}, \eta)$ such that $(V, \eta)$ is a model of $\mathcal{M}_n$ and $(f^0, \ldots, f^{n-1})$ is a Lorentz coordinate system for $(V, \eta)$. The models of $\mathcal{E}_n^+$ are those structures $(V, f^0, \ldots, f^{n-1}, \langle \cdot, \cdot \rangle)$ such that $(V, \langle \cdot, \cdot \rangle)$ is a model of $\mathcal{E}_n$ and $(f^0, \ldots, f^{n-1})$ is a Euclidean coordinate system for $(V, \langle \cdot, \cdot \rangle)$.\footnote{Note that for every model of $\mathcal{M}_n$, there are infinitely many models of $\mathcal{M}_n^+$. The same holds for $\mathcal{E}_n$ and $\mathcal{E}_n^+$.} With these specifications at hand, one can prove that:

**Proposition 4.** $\mathcal{M}_n^+$ and $\mathcal{E}_n^+$ are GD-equivalent (for $n \geq 2$).

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

$$\eta(u, v) = f^0(u) \cdot f^0(v) - \ldots - f^{n-1}(u) \cdot f^{n-1}(v).$$  \hspace{1cm} (3)$$
Now define a Euclidean inner product $\langle \cdot, \cdot \rangle$ over $V^+$:

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

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

Let us now see how a different analysis of the relationship between $\text{MG}^+_n$ and $\text{EG}^+_n$ 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 $\text{MG}^+_n$) and the model with the new coordinate system will represent exactly the same physical situation. This suggests to specify the categories of models of $\text{MG}^+_n$ and $\text{EG}^+_n$ in such a way that choices of coordinates may vary under morphisms. In particular, we should not take $\Sigma(\text{MG}^+_n)$-isomorphisms and $\Sigma(\text{EG}^+_n)$-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—i.e. 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 $\text{Cat}(\text{MG}^+_n)$ in such a way that $h : (V, f^0, \ldots, f^{n-1}, \eta) \rightarrow (V', f^0', \ldots, f'^{n-1}, \eta')$ is a morphism in $\text{Cat}(\text{MG}^+_n)$ iff $h$ is a bijection from $V$ to $V'$ that is an affine map\(^ {17} \) from $V$ to $V'$ 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. The morphisms in $\text{Cat}(\text{EG}^+_n)$ are specified analogously.

It follows that, the automorphism group of each model in $\text{Cat}(\text{MG}^+_n)$ is the Poincaré group, whereas the automorphism group of models in $\text{Cat}(\text{EG}^+_n)$ is the Euclidean group. The fact that the Poincaré group and the Euclidean group are not isomorphic implies that:

**Proposition 5.** $\text{MG}^+_n$ and $\text{EG}^+_n$ are not DC-equivalent (for $n \geq 2$).

We can say more when we consider which formulas count as meaningful in $\text{MG}^+_n$ and $\text{EG}^+_n$. 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 $\varphi$ over $\Sigma(\text{MG}^+_n)$ with $n$ free variables invariant under the morphisms of $\text{Cat}(\text{MG}^+_n)$ just in case for all morphisms $h : \mathcal{M} \rightarrow \mathcal{M}'$ in $\text{Cat}(\text{MG}^+_n)$, an $n$-tuple

\(^ {17}h\) is an affine map from a vector space $V = (V, \oplus, \odot)$ to another $V' = (V', \oplus', \odot')$ iff there is a vector space isomorphism $H$ from $V$ to $V'$ and a $w \in V$ such that for all $v \in V$: $h(v) = H(v) \odot' w$.\)
(a_1, \ldots, a_n) satisfies \( \varphi \) in \( \mathcal{M} \) iff \((h(a_1), \ldots, h(a_n))\) satisfies \( \varphi \) in \( \mathcal{M}' \). It follows that the formula

\[
z = f^0(u) \cdot f^0(v) + \ldots + f^{n-1}(u) \cdot f^{n-1}(v)
\]

(5)

is not invariant under the morphisms of \( \text{Cat}(\text{MG}^+_n) \). Hence, formula (5) is not meaningful in \( \text{MG}^+_n \). So although formula (5) defines a Euclidean inner product over any given model of \( \text{MG}^+_n \), this inner product will not count as meaningful in \( \text{MG}^+_n \). Thus, any reconstruction manual for \( \Sigma(\text{EG}^+_n) \) in terms of \( \Sigma(\text{MG}^+_n) \) yielding (5) as a translation of \( \langle u, v \rangle = z \) does not preserve meaningfulness. An analogous conclusions holds for reconstruction manuals for \( \Sigma(\text{MG}^+_n) \) in terms of \( \Sigma(\text{EG}^+_n) \).

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 generalisation 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 generalised 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 \( \text{Cat}(T_i) \) are precisely elementary embeddings \((1 \leq i \leq 2)\). If \( T_1 \) and \( T_2 \) 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 also not be much wider

\[^{18}\text{In contrast, the formula}
\]

\[
z = f^0(u) \cdot f^0(v) - \ldots - f^{n-1}(u) \cdot f^{n-1}(v)
\]

is invariant under the morphisms of \( \text{Cat}(\text{MG}^+_n) \) although coordinate functions vary from model to model. This formula is meaningful in \( \text{MG}^+_n \). So a formula may be meaningful even if the terms and function symbols occurring in it are not meaningful.
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 paper.

**Theorem 2.** Suppose $T_1$ and $T_2$ are theories with disjoint signatures and suppose the isomorphisms in $\text{Cat}(T_i)$ 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 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, i.e. $E_1|\Sigma(T_1) \cup \Sigma(T_2) = E_2|\Sigma(T_1) \cup \Sigma(T_2)$.

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 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 $F : \text{Cat}(T_1) \to \text{Cat}(T_2)$, $G : \text{Cat}(T_2) \to \text{Cat}(T_1)$ and reconstruction manuals $\Phi$ for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$ and $\Psi$ 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 $\Sigma(T_i)$-formulas defining a $\Sigma(T_i)$-isomorphism from $\mathcal{A}_1$ to $GF(\mathcal{A}_1)$ for all $\mathcal{A}_1 \in \text{Mod}(T_1)$ 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 $\mathcal{A}_2 \in \text{Mod}(T_2)$ such that

3. these isomorphisms are compatible with each other and with $\Phi$ and $\Psi$.

$^{19}E|\Sigma := \{A|\Sigma : A \in E\}$, where $A|\Sigma$ is the reduct of $A$ to $\Sigma$ (cf. Hodges, 1993, p. 9).
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 B.8, Appendix B.

**Theorem 3.** Suppose $T_1$ and $T_2$ are theories with disjoint signatures and suppose the isomorphisms in $\text{Cat}(T_i)$ 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.

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 (cf. 2018, p. 14). The question is whether there is a special property $P$ of functors such that the existence of a C-equivalence that is constituted by functors having property $P$ 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 (2016a) examples, different potential-based models $A_1, A_2$ that are related by a gauge transformation will be mapped to the same model $A$ by the composite auto-equivalence functor. However, the gauge transformations from $A_1$ to $A$ and from $A_2$ to $A$ are not induced by uniformly defined bijections in the way the many-sorted Ahlbrandt-Ziegler condition would require (cf. Definition B.8, Appendix 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 Conclusions

Let me sum up the core results of this paper. 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 where 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 research 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 pre-interpreted 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 notion of reconstruction functor by allowing partial reconstruction manuals, in which non-meaningful 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 $\Sigma$ and $\Sigma^+$ are higher-order signatures with $\Sigma \subset \Sigma^+$ and $C$ and $E$ are classes of $\Sigma$-structures and $\Sigma^+$-structures, respectively.

**Definition A.1.** $E$ is a GD-expansion of $C$ to $\Sigma^+$ by $\Delta$ iff $\Delta$ is a set of explicit definitions and sort-definitions of the symbols in $\Sigma^+ \setminus \Sigma$ in terms of $\Sigma$ such that

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

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

**Definition A.2.** $E$ is a stepwise GD-expansion of $C$ to $\Sigma^+$ 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 $E_n = E$ and
3. $E_i$ is a GD-expansion of $E_{i-1}$ to $\Sigma_i$ by $\Delta_i$ (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 $\Sigma_2$-isomorphism, there is only one model $A^*_2$ of $T_2$ such that the predicate $Q_0$ has a non-empty extension. Suppose for reductio that there are reconstruction functors $F : \text{Cat} (T_1) \to \text{Cat} (T_2)$ and $G : \text{Cat} (T_2) \to \text{Cat} (T_1)$ constituting a (dual) equivalence. Then for every $A_1 \vDash T_1$ there is an $A_2 \vDash T_2$ such that $A_2$ is $\Sigma_2$-isomorphic to $F(A_1)$ and $A_2$ is definable from $A_1$ by some reconstruction manual $\Phi$. It follows that

(*) there can only be one model $A^*_1 \vDash T_1$ up to $\Sigma_1$-isomorphism such that $\Phi_{Q_0}^{A^*_1} \neq \emptyset$, namely $A^*_1 \cong G(A^*_2)$.

The formula $\Phi_{Q_0}$ contains only finitely many of the predicates $P_0, P_1, \ldots$ of $\Sigma_1$. Therefore, we get many models $A$ of $T_1$ such that $A$ is not $\Sigma_1$-isomorphic to $G(A^*_2)$ but $\Phi_{Q_0}^{A} \neq \emptyset$. We achieve this by varying the extensions of those $P_i$ not contained in the formula $\Phi_{Q_0}$, while using the same extension as in $G(A^*_2)$ for the predicates contained in $\Phi_{Q_0}$. Thus we get a contradiction to (*). □
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 spacetime manifold without non-trivial 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 $A$ be a structure such that every basic type has a finite extension. Note that if $\tau_1$ and $\tau_2$ are types with finite extensions, then so are the product type $\tau_1 \times \tau_2$ and the power types $\varphi(\tau_1)$. Thus, every type has a finite extension. Therefore, every $\Sigma(A)$-formula can only have a finite extension under $A$. Hence, no infinite set is definable over $A$. \hfill \square

Proposition 3. 

(1) $MG_n$ and $EG_n$ are not DC-equivalent.

(2) $MG_n$ and $EG_n$ are not GD-equivalent.

Proof. The $\Sigma(MG_n)$-automorphism group of the models of $MG_n$ is not isomorphic to the $\Sigma(EG_n)$-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. \hfill \square

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

Proof. Let $VEC^+$ be the set of all structures $(V, f^0, \ldots, f^{n-1})$ such that $V$ is an $n$-dimensional vector space and $(f^0, \ldots, f^{n-1})$ is a linear coordinate system for $V$. It suffices to show that both $Mod(MG_n^+)$ and $Mod(EG_n^+)$ are GD-expansions of $VEC^+$. Let $M = (V^+, \eta)$ be a model of $MG_n^+$, where $V^+ = (V, f^0, \ldots, f^{n-1})$. Then $(f^0, \ldots, f^{n-1})$ is a Lorentz coordinate system for $M$. This implies that $M$ 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 $V^+ \in VEC^+$. Conversely, if $V^+ \in VEC^+$, one can define a Minkowskian inner product $\eta$ over $V^+$ using the equation above. Then $(V^+, \eta) \in Mod(MG_n^+)$. Hence, $Mod(MG_n^+$) is a GD-expansion of $VEC^+$. Analogously, one shows that $Mod(EG_n^+)$ is a GD-expansion of $VEC^+$. \hfill \square

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 $\text{Cat}(MG_n^+)$ and $\text{Cat}(EG_n^+)$ because the Poincaré group is not isomorphic to the Euclidean group. \hfill \square
Definition B.1. A stepwise GD-expansion $E$ of $C$ to $\Sigma^+$ by $(\Delta_1, \ldots, \Delta_n)$ uses only subsorts iff all basic sorts of $\Sigma^+$ are introduced as subsorts of types of $\Sigma$ by $(\Delta_1, \ldots, \Delta_n)$.

Lemma B.1. If $T_1$ and $T_2$ are GD-equivalent, there are stepwise GD-expansions $E_1$ and $E_2$ of $\text{Mod}(T_1)$ and $\text{Mod}(T_2)$ 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 $\text{Cat}(T_i)$ 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 B.1, $\text{Mod}(T_1)$ and $\text{Mod}(T_2)$ have stepwise GD-expansions $E_1$ and $E_2$ to some signature $\Sigma^+$ by $(\Delta^1_1, \ldots, \Delta^1_n)$ and $(\Delta^2_1, \ldots, \Delta^2_m)$ using only subsorts such that $E_1 = E_2$. This gives rise to reconstruction manuals. Let $\Phi$ map every symbol $S$ of $\Sigma(T_2)$ to the open $\Sigma(T_1)$-formula used to define $S$ in $(\Delta^1_1, \ldots, \Delta^1_n)$. Then $\Phi$ is a reconstruction manual for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$. Analogously, we get a reconstruction manual $\Psi$ for $\Sigma(T_1)$ in terms of $\Sigma(T_2)$.

Let $\text{Cat}(E_i)$ 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 $P_1 : \text{Cat}(E_1) \to \text{Cat}(T_1)$ and $P_2 : \text{Cat}(E_2) \to \text{Cat}(T_2)$ are equivalence functors, as described in Barrett and Halvorson (2016, Section 5). By the axiom of choice, there are functors $P^{-1}_1$ and $P^{-1}_2$ essentially inverse to $P_1$ and $P_2$. Since $\text{Cat}(E_1) = \text{Cat}(E_2)$, we can define functors $F := P_2P^{-1}_1 : \text{Cat}(T_1) \to \text{Cat}(T_2)$ and $G := P_1P^{-1}_2 : \text{Cat}(T_2) \to \text{Cat}(T_1)$. $F$ and $G$ constitute an equivalence of categories between. It is a matter of routine to check that $F$ and $G$ are reconstruction functors w.r.t. the reconstruction manuals $\Phi$ and $\Psi$.

In order to prove Theorem 2, we need several lemmata and definitions.

Lemma B.2. Suppose $\alpha$ and $\alpha'$ both are associations of $\Sigma_1$-types with $\Sigma_2$-types satisfying (1) in Definition 7 with respect to the same reconstruction manual $\Phi$ for $\Sigma_1$ in terms of $\Sigma_2$. Then for every type $\tau$ of $\Sigma_1$ : $\alpha(\tau) = \alpha'(\tau)$.

So given a type $\tau$ of $\Sigma_1$, a reconstruction manual $\Phi$ for $\Sigma_1$ in terms of $\Sigma_2$ associates a unique type $\Phi(\tau)$ of $\Sigma_2$ with $\tau$.

Definition B.2. Let $\Phi$ be a reconstruction manual for $\Sigma^*$ in terms of $\Sigma$ and let $T$ be a theory with signature $\Sigma$. Then we say that $\Phi$ is formally admissible w.r.t. $T_2$ iff for every model $\mathcal{A}$ of $T$:

1. for every basic sort $\sigma$ of $\Sigma^*$, $\Phi_\sigma$ defines a non-empty set over $\mathcal{A}$.
2. for every function symbol $f$ of $\Sigma^*$, $\Phi_f$ defines a function over $\mathcal{A}$.
3. for every individual constant $c$ of $\Sigma^*$, $\Phi_c$ defines a singleton over $\mathcal{A}$.
From now on suppose that (a) $T$ and $T^*$ have disjoint signatures $\Sigma$ and $\Sigma^*$ and (b) $\Phi$ is a reconstruction manual for $\Sigma^*$ in terms of $\Sigma$ that is formally admissible w.r.t. $T$.

**Definition B.3.** $\Sigma^+$ is a $\Phi$-extension of $\Sigma$ by external auxiliaries iff $\Sigma^+$ contains the symbols of $\Sigma$ as well as $\Sigma^*$ and

1. for every basic sort $\sigma$ in $\Sigma^*$, $\Sigma^+$ contains exactly one new function symbol $i_\sigma$ with $\sigma$ as input-type and $\Phi(\sigma)$ as output-type.
2. for every type $\tau$ of $\Sigma^*$, $\Sigma^+$ contains exactly one new function symbol $j_\tau$ with $\tau$ as input-type and $\Phi(\tau)$ as output-type.

From now on let $\Sigma^+$ be a $\Phi$-extension of $\Sigma$ by external auxiliaries.

**Definition B.4.** Let $A \in \text{Mod}(T)$. Then we call the $\Sigma^+$-structure $A^+$ with the following properties the canonical expansion of $A$ to $\Sigma^+$ induced by $\Phi$:

1. $A^+|\Sigma = A$.
2. For every symbol $S$ in $\Sigma^*$ that is not an individual constant: $S^{A^+} = \Phi_S^A$.
3. For every individual constant $c$ in $\Sigma^*$: $\{c^{A^+}\} = \Phi_c^A$.
4. For every basic sort $\sigma$ in $\Sigma^*$: $i_\sigma^{A^+}$ is the identity on $\Phi_\sigma^A$.
5. For every type $\tau$ over $\Sigma^*$: $j_\tau^{A^+}$ is the identity on $D_\tau$, where $D_\tau$ is defined as follows:
   
   (a) if $\tau$ is a basic sort of $\Sigma^*$, then $D_\tau = \Phi_\tau^A$.
   
   (b) if $\tau = \tau_1 \times \tau_2$ is a product type of $\Sigma^*$, then $D_\tau = D_{\tau_1} \times D_{\tau_2}$.
   
   (c) if $\tau = \varphi(\tau')$ is a power type of $\Sigma^*$, then $D_\tau = \mathcal{P}(D_{\tau'})$.

**Definition B.5.** The canonical expansion of $\text{Mod}(T)$ to $\Sigma^+$ induced by $\Phi$ is the isomorphic closure of $\{A^+ : A \in \text{Mod}(T)\}$, where $A^+$ is the canonical expansion of $A$ to $\Sigma^+$ induced by $\Phi$ for $A \in \text{Mod}(T)$.

**Notation.** In the following it is useful to write variables with their type as superscript (e.g. $x^{\tau}$) to avoid confusions.

**Lemma B.3.** The canonical expansion of $\text{Mod}(T)$ to $\Sigma^+$ induced by $\Phi$ is a stepwise GD-expansion of $\text{Mod}(T)$ to $\Sigma^+$ by $(\Delta_\Phi^{\text{sorts}}, \Delta_\Phi^{\text{inj}}, \Delta_\Phi^{\text{const}})$, where

1. $\Delta_\Phi^{\text{sorts}}$ consists of subsort-definitions for all basic sorts $\sigma$ of $\Sigma^*$:

   \[
   \forall x^{\Phi(\sigma)}(\Phi_\sigma(x^{\Phi(\sigma)}) \leftrightarrow \exists y^\sigma \ i_\sigma(y^\sigma) = x^{\Phi(\sigma)}) \land \\
   \forall y_1^\sigma \forall y_2^\sigma(i_\sigma(y_1^\sigma) = i_\sigma(y_2^\sigma) \rightarrow y_1^\sigma = y_2^\sigma),
   \]

   (23)
(2) $\Delta^{\text{inj}}_\Phi$ consists of definitions of the injection symbols $j_\tau$ for types $\tau$ of $\Sigma^*$:

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

where the definiens $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is an $\Sigma \cup \Sigma(\Delta^{\text{sorts}}_\Phi)$-formula for every type $\tau$ of $\Sigma^*$, recursively defined as follows:

(a) if $\tau$ is a basic sort of $\Sigma^*$, 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 $\Sigma^*$, then $\varphi_\tau(x^\tau, z^{\Phi(\tau)})$ is the formula

$$\exists x^{\tau_1} \exists x^{\tau_2} (x^\tau = \langle x^{\tau_1}, x^{\tau_2} \rangle \land \exists z^{\Phi(\tau_1)} \exists z^{\Phi(\tau_2)} (\varphi_{\tau_1}(x^{\tau_1}, z^{\Phi(\tau_1)}) \land \varphi_{\tau_2}(x^{\tau_2}, z^{\Phi(\tau_2)}) \land z^{\Phi(\tau)} = \langle z^{\Phi(\tau_1)}, z^{\Phi(\tau_2)} \rangle)).$$

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

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

(3) $\Delta^{\text{const}}_\Phi$ consists of the following definitions:

(a) $\forall x^\tau (P(x^\tau) \leftrightarrow \Phi_\tau(j_\tau(x^\tau)))$, for every predicate $P$ in $\Sigma^*$ of type $\tau$.

(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 $\Sigma^*$ with input-type $\tau_1$ and output-type $\tau_2$.

(c) $\forall x^\sigma (c = x^\sigma \leftrightarrow \Phi_c(j_\tau(x^\tau)))$, for every individual constant $c$ of basic sort $\sigma$ in $\Sigma^*$.

Lemma B.4. Suppose the following: $H : \text{Cat}(T) \to \text{Cat}(T^*)$ is a reconstruction functor with reconstruction manual $\Phi$; the isomorphisms in $\text{Cat}(T^*)$ are precisely the $\Sigma^*$-isomorphisms; $A$ is a model of $T$; $A^+$ is the canonical expansion of $A$ to $\Sigma^+$ induced by $\Phi$. Then $H(A)$ is $\Sigma^*$-isomorphic to $A^+|_{\Sigma^*}$.

Lemma B.5. Suppose $E$ is the canonical expansion of $\text{Mod}(T)$ to $\Sigma^+$ induced by $\Phi$, $A \in E$ and $(A|_\Sigma)^+$ is the canonical expansion of $A|_{\Sigma}$ to $\Sigma^+$ induced by $\Phi$. Then $A$ is $\Sigma^*$-isomorphic to $(A|_{\Sigma})^+$.

Theorem 2. Suppose $T_1$ and $T_2$ are theories with disjoint signatures and suppose the isomorphisms in $\text{Cat}(T_i)$ 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.

Proof. Suppose $T_1$ and $T_2$ are DC-equivalent. Let $\Sigma_1$ and $\Sigma_2$ be their signatures. Then there are reconstruction functors $F : \text{Cat}(T_1) \to \text{Cat}(T_2)$ and $G : \text{Cat}(T_2) \to \text{Cat}(T_1)$ constituting a (dual) equivalence accompanied by reconstruction manuals $\Phi$ for $\Sigma_2$ over $\Sigma_1$ and $\Psi$ for $\Sigma_1$ over $\Sigma_2$. It follows that $\Phi$ and $\Psi$ are formally admissible w.r.t. $T_1$ and $T_2$, respectively. Let $\Sigma_1^+$ and $\Sigma_2^+$ be $\Phi$- and $\Psi$-extensions of $\Sigma_1$ and $\Sigma_2$ by external auxiliaries, respectively. Now let $E_1$ and $E_2$
be the canonical expansions of \(\text{Mod}(T_1)\) and \(\text{Mod}(T_2)\) to \(\Sigma_1^+\) and \(\Sigma_2^+\) induced by \(\Phi\) and \(\Psi\). We know from Lemma B.3 that \(E_1\) and \(E_2\) are stepwise GD-expansions of \(\text{Mod}(T_1)\) and \(\text{Mod}(T_2)\), respectively. We show that \(E_1|_{\Sigma_1 \cup \Sigma_2} = E_2|_{\Sigma_1 \cup \Sigma_2}\).

Suppose \(A \in E_1|_{\Sigma_1 \cup \Sigma_2}\). Then there is a \(B \in E_1\) such that \(A = B|_{\Sigma_1 \cup \Sigma_2}\). Moreover, \(A|_{\Sigma_1} \in \text{Mod}(T_1)\). Let \((A|_{\Sigma_1})^+\) be the canonical expansion of \(A|_{\Sigma_1}\) to \(\Sigma_1^+\) induced by \(\Phi\). Check that the following statements hold.

1. \(B\) is \(\Sigma_1^+\)-isomorphic to \((A|_{\Sigma_1})^+\) (by Lemma B.5 and \(B|_{\Sigma_1} = A|_{\Sigma_1}\)).
2. \(F(A|_{\Sigma_1})\) is \(\Sigma_2\)-isomorphic to \((A|_{\Sigma_1})^+|_{\Sigma_2}\) (Lemma B.4).
3. \(A|_{\Sigma_2}\) is \(\Sigma_2\)-isomorphic to \(F(A|_{\Sigma_1})\) (by 1, 2 and \(B|_{\Sigma_2} = A|_{\Sigma_2}\)).
4. \(A|_{\Sigma_1}\) is \(\Sigma_1\)-isomorphic to \(\text{GF}(A|_{\Sigma_1})\) (because \(F\) and \(G\) constitute a (dual) equivalence and isomorphisms in \(\text{Cat}(T_1)\) are \(\Sigma(T_i)\)-isomorphisms).
5. \(\text{GF}(A|_{\Sigma_1})\) is \(\Sigma_1\)-isomorphic to \(F(A|_{\Sigma_1})^+|_{\Sigma_1}\), where \(F(A|_{\Sigma_1})^+\) is the canonical expansion of \(F(A|_{\Sigma_1})\) induced by \(\Psi\) (Lemma B.4).
6. \(A|_{\Sigma_1}\) is \(\Sigma_1\)-isomorphic to \(F(A|_{\Sigma_1})^+|_{\Sigma_1}\) (by 4 and 5).
7. \(A\) is \(\Sigma_1 \cup \Sigma_2\)-isomorphic to \(F(A|_{\Sigma_1})^+|_{\Sigma_1 \cup \Sigma_2}\) (by 3 and 6).

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

**Definition B.6.** We say that \(\chi\) is a family of formulas defining inner isomorphisms of \(T_1\)-models w.r.t. \(F, \Phi\) and \(G, \Psi\) iff

1. \(\chi\) maps every basic sort \(\sigma_1\) of \(\Sigma(T_1)\) to a \(\Sigma(T_1)\)-formula \(\chi_{\sigma_1}(x^{\sigma_1}, y^{\Phi(\sigma_1)})\).
2. for every basic sort \(\sigma_1\) of \(\Sigma(T_1)\) and every \(A_1 \in \text{Mod}(T_1)\), \(\chi_{\sigma_1}^A\) is a bijection from \(\sigma_1^A\) to \(\sigma_1^{\text{GF}(A_1)}\); and these bijections constitute a \(\Sigma(T_i)\)-isomorphism from \(A_1\) to \(\text{GF}(A_1)\).

**Definition B.7.** Let \(\Sigma(T_i)^+\) be a \(\Phi\)-extension of \(\Sigma(T_i)\) by external auxiliaries, let \(E_1\) be the canonical expansion of \(\text{Mod}(T_1)\) to \(\Sigma(T_i)^+\) induced by \(\Phi\) and let \(\chi\) be a family of formulas defining inner isomorphisms of \(T_1\)-models w.r.t. \(F, \Phi\) and \(G, \Psi\). Then the \(\chi\)-expansion of \(\text{Mod}(T_1)\) to \(\Sigma(T_i)^+ \cup \Sigma(T_2)^+\) induced by \(\Phi\) and \(\Psi\) is the stepwise GD-expansion of \(E_1\) to \(\Sigma(T_i)^+ \cup \Sigma(T_2)^+\) by \((\Delta_\psi, \Delta^{\text{inj}}_\psi)\), where \(\Delta^{\text{inj}}_\psi\) is as in Lemma B.3 and \(\Delta_\psi\) comprises definitions of the following form for all basic sorts \(\sigma_1\) of \(\Sigma(T_i)\):

\[
\forall x^{\sigma_1} \forall y^{\Psi(\sigma_1)} (i_{\sigma_1}(x^{\sigma_1}) = y^{\Phi(\sigma_1)} \iff \chi_{\sigma_1}(x^{\sigma_1}, j_{\Psi(\sigma_1)}(y^{\Psi(\sigma_1)})))
\]

(Def-\(i_{\sigma_1}\))

25
Definition B.8. $T_1$ and $T_2$ are DC-equivalent with definable inner isomorphisms iff there are functors $F : \text{Cat}(T_1) \to \text{Cat}(T_2)$, $G : \text{Cat}(T_2) \to \text{Cat}(T_1)$ and reconstruction manuals $\Phi$ for $\Sigma(T_2)$ in terms of $\Sigma(T_1)$ and $\Psi$ 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 $\chi$ defining inner isomorphisms of $T_1$-models w.r.t. $F, \Phi$ and $G, \Psi$ and
2. there is a family of formulas $\tilde{\chi}$ defining inner isomorphisms of $T_2$-models w.r.t. $G, \Psi$ and $F, \Phi$ such that
3. these isomorphisms are compatible with each other and with $\Phi$ and $\Psi$ in the following sense:
   1. for every model $\mathcal{B}_1$ of the $\chi$-expansion of $\text{Mod}(T_1)$ to $\Sigma(T_1)^+ \cup \Sigma(T_2)^+$ induced by $\Phi$ and $\Psi$ and all basic sorts $\sigma_2$ of $\Sigma(T_2)$: $j^{\mathcal{B}_1}_{\Phi(\sigma_2)} \circ i^{\mathcal{B}_1}_{\sigma_2} = \chi^{\mathcal{B}_1}_{\sigma_2}$.
   2. for every model $\mathcal{B}_2$ of the $\tilde{\chi}$-expansion of $\text{Mod}(T_2)$ to $\Sigma(T_2)^+ \cup \Sigma(T_1)^+$ induced by $\Phi$ and $\Psi$ and all basic sorts $\sigma_1$ 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 $\text{Cat}(T_i)$ are precisely the $\Sigma_i$-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 $\Sigma_1, \Sigma_2$ be the signatures of $T_1, T_2$. There are functors $F, G$, reconstruction manuals $\Phi, \Psi$ and stepwise GD-expansions $E_1, E_2$ to extended signatures $\Sigma_1^+, \Sigma_2^+$ 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 $\chi$ and $\tilde{\chi}$ defining inner isomorphisms of $T_1$-models and $T_2$-models, respectively.

Note that, for every basic sort $\sigma_1$ and every type $\tau_1$ of $\Sigma_1$, the auxiliary function symbols $i_{\sigma_1}$ and $j_{\tau_1}$ are in $\Sigma_2^+$ but not in $\Sigma_1^+$. Analogously, for every basic sort $\sigma_2$ and every type $\tau_2$ of $\Sigma_2$, the auxiliary function symbols $i_{\sigma_2}$ and $j_{\tau_2}$ are in $\Sigma_1^+$ but not in $\Sigma_2^+$. Moreover, that is the only difference between $\Sigma_1^+$ and $\Sigma_2^+$.

Let $E_1^+$ be the $\chi$-expansion of $\text{Mod}(T_1)$ to $\Sigma_1^+ \cup \Sigma_2^+$ induced by $\Phi$ and $\Psi$ and let $E_2^+$ be the $\tilde{\chi}$-expansion of $\text{Mod}(T_2)$ 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 $E_2^+ \models \text{Def-}i_{\sigma_1}$ for $\sigma_1$ in $\Sigma_1$ and $E_1^+ \models \text{Def-}i_{\sigma_2}$ for $\sigma_2$ in $\Sigma_2$ because the definitions of $j_{\tau_1}$ and $j_{\tau_1}$ are the same for $E_1^+$ and $E_2^+$. That $E_2^+ \models \text{Def-}i_{\sigma_1}$ and $E_1^+ \models \text{Def-}i_{\sigma_2}$ follows directly from the many-sorted Ahlbrandt-Ziegler condition.
References


