

Log. Univers.

© 2020 The Author(s)

<https://doi.org/10.1007/s11787-020-00257-y>

Logica Universalis



# Bases for Structures and Theories I

Jeffrey Ketland

**Abstract.** Sometimes structures or theories are formulated with different sets of primitives and yet are definitionally equivalent. In a sense, the transformations between such equivalent formulations are rather like basis transformations in linear algebra or co-ordinate transformations in geometry. Here an analogous idea is investigated. Let a relational signature  $P = \{P_i\}_{i \in I_P}$  be given. For a set  $\Phi = \{\phi_i\}_{i \in I_\Phi}$  of  $L_P$ -formulas, we introduce a corresponding set  $Q = \{Q_i\}_{i \in I_\Phi}$  of new relation symbols and a set of explicit definitions of the  $Q_i$  in terms of the  $\phi_i$ . This is called a definition system, denoted  $d_\Phi$ . A definition system  $d_\Phi$  determines a *translation function*  $\tau_\Phi : L_Q \rightarrow L_P$ . Any  $L_P$ -structure  $A$  can be uniquely definitionally expanded to a model  $A^+ \models d_\Phi$ , called  $A + d_\Phi$ . The reduct  $A + d_\Phi$  to the  $Q$ -symbols is called the *definitional image*  $D_\Phi A$  of  $A$ . Likewise, a theory  $T$  in  $L_P$  may be extended a definitional extension  $T + d_\Phi$ ; the restriction of this extension  $T + d_\Phi$  to  $L_Q$  is called the *definitional image*  $D_\Phi T$  of  $T$ . If  $T_1$  and  $T_2$  are in disjoint signatures and  $T_1 + d_\Phi \equiv T_2 + d_\Theta$ , we say that  $T_1$  and  $T_2$  are *definitionally equivalent* (wrt the definition systems  $d_\Phi$  and  $d_\Theta$ ). Some results relating these notions are given, culminating in two characterization theorems for the definitional equivalence of structures and theories.

**Mathematics Subject Classification.** Primary 03C07; Secondary 03C95.

**Keywords.** Definitional equivalence, Theories, Definability.

## 1. Introduction

Sometimes theories are formulated with different sets of primitives and yet are definitionally equivalent. The *non-logical primitives* of a formalized language  $L$  are called its *signature*. There are many examples of theories (often involving formalized systems of arithmetic and set theory) formulated in very different signatures, which are nonetheless “equivalent”. To take a simple example, consider the theory  $T_1$  of a reflexive relation:

$$T_1 : \quad \forall x P(x, x),$$

expressed with a binary relation symbol  $P$ . Suppose we introduce a new binary relation symbol,  $Q$ , and give an explicit definition of it—call this definition  $d_1$ —in terms of  $P$  as follows:

$$d_1 : \quad \forall x \forall y (Q(x, y) \leftrightarrow (P(x, y) \wedge x \neq y)).$$

Call the extended theory  $T_1 + d_1$ . Then  $T_1 + d_1 \vdash \forall x \neg Q(x, x)$ .

Let  $T_2$  be the theory saying that  $Q$  is an irreflexive relation:

$$T_2 : \quad \forall x \neg Q(x, x).$$

Consider following “*inverse definition*” of  $P$  in terms of  $Q$ , call it  $d_2$ :

$$d_2 : \quad \forall x \forall y (P(x, y) \leftrightarrow (Q(x, y) \vee x = y)).$$

Call the extended theory  $T_2 + d_2$ . Then  $T_2 + d_2 \vdash \forall x P(x, x)$ . Moreover, we have:

$$T_1 + d_1 \vdash d_2 \quad \text{and} \quad T_2 + d_2 \vdash d_1.$$

These two theories,  $T_1$  and  $T_2$ , are in fact *equivalent* with respect to these definitions. That is,  $T_1 + d_1$  and  $T_2 + d_2$  are *logically equivalent*:

$$\boxed{T_1 + d_1 \equiv T_2 + d_2.}$$

We say that  $T_1$  and  $T_2$  have a “common definitional extension” and are therefore “*definitionally equivalent*” (see Definition 27 below).

As the reader may have guessed, this example derives from the standard mathematical method of passing between a non-strict preorder  $\leq$  and its corresponding strict preorder  $<$ : they may be defined in terms of each other. To put the above in the more standard notation:

$$\begin{aligned} & \{ \forall x (x \leq x), \forall x \forall y (x < y \leftrightarrow (x \leq y \wedge x \neq y)) \} \\ & \equiv \{ \forall x \neg (x < x), \forall x \forall y (x \leq y \leftrightarrow (x < y \vee x = y)) \}. \end{aligned}$$

In effect, the explicit definitions  $d_1$  and  $d_2$  are “mutual inverses”. This notion will play a major role later.

Moving on to increasingly sophisticated examples, there are equivalent reformulations of Peano arithmetic PA with different primitives from the usual signature  $\{0, S, +, \times\}$ . For example, one might take exponentiation as the single basic notion, along with certain axioms for exponentiation, along with explicit definitions for  $0, S, +$  and  $\times$  (along with induction). The result is equivalent to PA.

As is well-known, one can interpret a modification we shall call  $ZF_0$  of ZF set theory, obtained by removing the axiom of infinity, adding its negation, along with an additional axiom of “transitive containment”, into PA. Moreover, the inverse of this interpretation interprets PA into  $ZF_0$ .<sup>1</sup> So PA is *definitionally equivalent* to the theory of *finite sets*  $ZF_0$ .

---

<sup>1</sup>A detailed analysis and proof of this “mathematical folklore” result is given in Kaye and Wong [13], who show that there exist translation functions between PA and  $ZF_0$  which are mutual inverses (Theorem 20 of their paper). A fairly detailed discussion of the equivalence of Peano arithmetic and finite set theory  $ZF_0$  may also be found in the lecture notes Andr eka and N emeti [2], pp. 41–45.

A similar relationship holds between the theory of *formalized syntax* and PA. Suppose  $S_A$  is the theory of concatenation for strings from alphabet  $A$ , with  $|A| \geq 2$ , and with the appropriate induction principle. Then  $S_A$  is *definitionally equivalent* to PA.<sup>2</sup>

There are other examples—from mathematics, logic and philosophy of science.<sup>3</sup> To return to the broader point, the transformations between such equivalent formulations are rather like “basis transformations” in linear algebra and other parts of mathematics. In this paper and the follow-up, an analogous idea is investigated.<sup>4</sup>

## 2. Syntax, Structures and Theories

Throughout, everything we consider is 1-sorted, relational and first-order.

**Definition 1.** Let  $P = \{P_i\}_{i \in I}$  be a set and let  $a : P \rightarrow \mathbb{N}$ . The pair  $S = (P, a)$  is called a one-sorted *signature*, and  $a$  is called the *arity* function for  $S$ . The multiset  $t = (a(P_i) \mid P_i \in P)$  is called the *similarity type* of the signature  $S$ . If  $a(P_i) = 0$ , then  $P_i$  is called a *sentence letter* (or a *propositional atom*). The *alphabet* of  $S$  is  $P$ .

**Definition 2.**  $(P^c, a^c)$  is a *copy* of  $(P, a)$  iff the similarity types of  $(P, a)$  and  $(P^c, a^c)$  are the same.  $(P^c, a^c)$  is a *disjoint copy* of  $(P, a)$  when, in addition,  $P \cap P^c = \emptyset$ .

**Definition 3.**  $L_P$  is the first-order language over the signature  $S = (P, a)$ , where each symbol  $P_i$  is a primitive relation symbol of arity  $a(P_i)$ . We will sometimes call  $L_P$  “the  $P$ -language”.

---

<sup>2</sup>A weak theory of concatenation without the induction principle is now usually called TC (Grzegorzczak [8]), and the precise interpretability relationship of TC and Robinson arithmetic Q has recently been clarified. Although similar systems had been studied before (Quine [15]; Tarski et al. [17]), the undecidability of TC is demonstrated in Grzegorzczak [8]; in Grzegorzczak and Zdanowski [9], the essential undecidability of TC is demonstrated through an interpretation of TC into Robinson arithmetic Q. Subsequently, Visser and Sterken [18], Ganea [7] and Švejdar [16] demonstrated the interpretability of Q into TC.

<sup>3</sup>As another example, there is the theory CEM, of “classical extensional mereology”, in a language with basic binary relation symbol  $\preceq$  ( $x \preceq y$  means “ $x$  is part of  $y$ ”). There is a definitionally equivalent theory I shall call F (for “fusions”) with a basic binary operation symbol  $\oplus$  (where  $x \oplus y$  can be read “the fusion of  $x$  and  $y$ ”). The detailed formulation and proof of definitional equivalence are given in Ketland and Schindler [12].

<sup>4</sup>Often, we have translations/interpretations  $T_1 \rightarrow T_2$  and  $T_2 \rightarrow T_1$ , for theories which appear, on the surface, to be quite different. But it is not automatically true that  $T_1$  and  $T_2$  are definitionally equivalent: the translations involved must be *mutual inverses* of each other. (See our Theorem 2 below.) A valuable discussion of this point, and a criterion for it to hold, is Friedman and Visser [6], who show that when two theories are “bi-interpretable via identity-preserving interpretations”, then they are definitionally equivalent (Sect. 5 of their paper). Moreover, they give an example of two finitely axiomatized sequential theories that are bi-interpretable but not definitionally equivalent (Sect. 7 of their paper). An example of a pair of theories which “define each others’ models” but which are not definitionally equivalent is given in Andr eka et al. [1].

**Definition 4.** Let

$$\Phi = \{\phi_1, \dots\} = \{\phi_i\}_{i \in I_\Phi}$$

be a set of  $L_P$ -formulas, which will be called *defining formulas*. Given the  $\phi_i$ , a corresponding set

$$Q = \{Q_i\}_{i \in I_\Phi}$$

of new relation symbols  $Q_i$  is introduced, such that the arity of each  $Q_i$  matches the arity of its corresponding  $\phi_i$ . The new language  $L_Q$  will sometimes be called “the  $Q$ -language”. The combined language is then called  $L_{P,Q}$ .

A theory  $T$  in  $L$  is a set of  $L$ -sentences. When we require deductive closure, we say so. We use a deductive system such that, if  $\Delta \vdash \alpha$ , then  $\Delta \vdash \forall x \alpha$ , so long as  $x$  doesn’t appear free in any formulas in  $\Delta$ .<sup>5</sup>  $T \vdash \alpha$  means: there exists a derivation of  $\alpha$  from the axioms/rules of  $T$ . An  $L$ -structure  $A$  will always interpret all *variables*. So, we can always write  $A \models \alpha$ , even where  $\alpha$  has free variables, since  $A$  will assign a value  $x^A$  to each variable  $x \in \text{FV}(\alpha)$ .<sup>6</sup> The Completeness Theorem holds in the usual form:  $\Delta \vdash \alpha$  iff, for any  $A \models \Delta$ ,  $A \models \alpha$ . A theory  $T$  in  $L_P$  is said to be *deductively closed* iff, for all  $\alpha \in \text{Sent}(L_P)$ , if  $T \vdash \alpha$ , then  $\alpha \in T$ . The *deductive closure* of  $T$  (written  $\text{DedCl}(T)$ ) is  $\{\alpha \in \text{Sent}(L_P) \mid T \vdash \alpha\}$ .

So far as I can tell, nothing in this paper uses either proof theoretic methods or model theoretic methods beyond what is taught at intermediate logic.<sup>7</sup> We do introduce specific new terminology for the following notions:

---

|   |   |
|---|---|
| <i>Definition systems</i>                 | $d_\Phi$ .  |
| <i>Definitional expansions/extensions</i> | $A \mapsto A + d_\Phi$ and $T \mapsto T + d_\Phi$ .                                 |
| <i>Definitional image operator</i>        | $A \mapsto D_\Phi A$ and $T \mapsto D_\Phi T$ .                                     |
| <i>Definitional equivalence</i>           | $A \xleftrightarrow[\Theta]{\Phi} B$ and $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$ . |

---

**Definition 5.** A *structure*  $A$  for the language  $L_P$  specifies a non-empty domain,  $\text{dom}(A)$ ; and interprets each variable  $x$  of  $L_P$  as an element  $x^A \in \text{dom}(A)$ ; and interprets each  $n$ -ary relation symbol  $P_i$  as a  $n$ -ary relation  $(P_i)^A \subseteq (\text{dom}(A))^n$ .

**Definition 6.** Given signature  $P$ , let  $P^c$  be a disjoint copy of  $P$ . Let  $A$  be an  $L_P$ -structure. Then the *disjoint copy*  $A^c$  of  $A$  in  $L_{P^c}$  is defined by setting  $\text{dom}(A^c) = \text{dom}(A)$  and, for each  $P_i$ , setting  $(P_i^c)^{A^c} = (P_i)^A$ . Let  $T$  be an  $L_P$ -theory. Then the *disjoint copy*  $T^c$  of  $T$  in  $L_{P^c}$  is defined by replacing every occurrence of  $P_i$ , in any sentence in  $T$  by the new symbol  $P_i^c$ .

---

<sup>5</sup>The deductive system I usually have in mind is the Hilbert system set out in Machover [14] or in Enderton [5].

<sup>6</sup>For any variable  $x$ , the denotation of  $x$  in  $A$  is  $x^A$ . We define  $A_a^x$  to be the structure just like  $A$ , except that, for the variable  $x$ , we have  $x^{A_a^x} = a$ .

<sup>7</sup>The terminology and definitions we use largely follow those of Machover [14], Hodges [10], or Enderton [5].

We may next give inductive definitions of the *denotation* function  $t \mapsto t^A$  (specifying what any term  $t$  in  $L_P$  refers to in  $A$ ) and the *satisfaction* relation  $\models$  between  $A$  and  $L_P$ -formulas.

**Definition 7.** If  $A_1$  and  $A_2$  are  $L_P$ -structures, then an *isomorphism*

$$f : A_1 \rightarrow A_2$$

is a bijection from  $\text{dom}(A_1)$  to  $\text{dom}(A_2)$  satisfying the preservation condition that  $f[(P_i)^{A_1}] = (P_i)^{A_2}$ , for each relation symbol  $P_i$  in the signature  $P$ . This is written

$$A_1 \stackrel{f}{\cong} A_2$$

(or simply  $A_1 \cong A_2$  if the isomorphism is left implicit).

**Definition 8.** If  $A$  is an  $L_P$ -structure and  $T, T_1, T_2$  are sets of  $L_P$ -sentences:

---

|     |                       |    |  |
|-----|-----------------------|----|--|
| (1) | $\text{Th}_{L_P}(A)$  | := | the set of $L_P$ -sentences true in $A$ .        |
| (2) | $A \models T$         | := | for all $\alpha \in T$ , $A \models \alpha$ .    |
| (3) | $\text{Mod}_{L_P}(T)$ | := | the class of $L_P$ -structures $A \models T$     |
| (4) | $T_1 \equiv T_2$      | := | $\text{Mod}(T_1) = \text{Mod}(T_2)$ .            |
| (5) | $T_1 \vdash T_2$      | := | for all $\alpha \in T_2$ , $T_1 \vdash \alpha$ . |

---

$\text{Th}_{L_P}(A)$  is called  $\text{Th}(A)$  if it's clear from context what language is involved. Likewise,  $\text{Mod}_{L_P}(T)$  is called  $\text{Mod}(T)$  if it's clear from context what language is involved.

If  $T_1$  and  $T_2$  are deductively closed theories, then  $T_1 \vdash T_2$  iff  $T_2 \subseteq T_1$ . The Completeness Theorem tells us that  $T_1 \equiv T_2$  iff  $T_1 \vdash T_2$  and  $T_2 \vdash T_1$ .

**Definition 9.** Let  $A$  be an  $L_P$ -structure and  $A^+$  be an  $L_{P,Q}$ -structure. Then  $A^+$  is an *expansion* of  $A$  iff for all  $P_i$ ,  $(P_i)^{A^+} = (P_i)^A$ . This is equivalent to saying that  $A$  is a *reduct* of  $A^+$ . If  $A^+$  is an  $L_{P,Q}$ -structure, its reduct to  $P$  is denoted  $A^+ \upharpoonright_P$  (an  $L_P$ -structure) and its reduct to  $Q$  is denoted  $A^+ \upharpoonright_Q$  (an  $L_Q$ -structure) and we have:  $(P_i)^{A^+ \upharpoonright_P} = (P_i)^{A^+}$  and  $(Q_i)^{A^+ \upharpoonright_Q} = (Q_i)^{A^+}$ .

The central property of expansions is that the truth value of a formula in the smaller language  $L_P$  remains invariant as we pass from an  $L_P$ -structure to an expanded structure for  $L_{P,Q}$ : if  $\alpha \in L_P$  and an  $L_{P,Q}$ -structure  $A^+$  is an expansion of an  $L_P$ -structure  $A$ , then  $A^+ \models \alpha$  iff  $A \models \alpha$ .

**Definition 10.** A theory  $T^+$  is an *extension* of  $T$  iff  $T$  is a subset of  $T^+$ . Let signatures  $P, Q$ , and corresponding languages  $L_P, L_Q$  and  $L_{P,Q}$  be given. An extension  $T^+$  in  $L_{P,Q}$  of  $T$  in  $L_P$  is called a *conservative* extension of  $T$  with respect to  $L_P$ -formulas iff, for any  $L_P$ -formula  $\alpha$ ,

$$T^+ \vdash \alpha \Rightarrow T \vdash \alpha$$

### 3. Definition

**Definition 11.** Given the set  $\Phi = \{\phi_i\}_{i \in I}$  of  $L_P$ -formulas, we introduce a disjoint set  $Q = \{Q_i\}_{i \in I}$  of new relation symbols, with  $\text{card } Q = \text{card } \Phi$ , and with the arity of  $Q_i$  matching the arity of  $\phi_i$ , and let  $n_i$  be  $a(\phi_i)$ . The *definition system* over  $\Phi$ , which we write as,

$$d_\Phi$$

is the set of explicit definitions,

$$\forall x_1 \dots x_{n_i} (Q_i(x_1, \dots, x_{n_i}) \leftrightarrow \phi_i)$$

where  $\{x_1, \dots, x_{n_i}\} = \text{FV}(\phi_i)$ . These define the new symbols  $Q_i$  in terms of the  $L_P$ -formulas  $\phi_i$ . We shall sometimes write  $\forall \bar{x} (Q_i(\bar{x}) \leftrightarrow \phi_i)$  instead of  $\forall x_1 \dots x_n (Q_i(x_1, \dots, x_n) \leftrightarrow \phi_i)$ .<sup>8</sup>

**Definition 12.** If  $A$  is an  $L_P$ -structure and  $\phi(x_1, \dots, x_n)$  is an  $L_P$ -formula, then

$$\phi(A) := \{(a_1, \dots, a_n) \in (\text{dom}(A))^n \mid A \models \phi[a_1, \dots, a_n]\}$$

is the relation that  $\phi$  defines in  $A$ .

**Definition 13.** An  $L_{P,Q}$ -structure  $A^+$  is a *definitional expansion* of an  $L_P$ -structure  $A$  with respect to the definition system  $d_\Phi$  just if  $A^+$  is an expansion of  $A$  interpreting each  $Q_i$ , and  $A^+ \models d_\Phi$ . This ensures that  $(Q_i)^{A^+} = \phi_i(A^+)$ , for each  $Q_i$ .

Given any  $L_P$ -structure  $A$ , it is clear that there is a unique definitional expansion  $A^+ \models d_\Phi$ . We introduce the following notation for this expansion:

**Definition 14.**  $A + d_\Phi :=$  the unique definitional expansion  $A^+ \models d_\Phi$ .

We are going to treat the definitional expansion map

$$A \mapsto A + d_\Phi$$

as a unary operator  $+d_\Phi$  (indexed by  $\Phi$ ), taking us from  $L_P$ -structures to  $L_{P,Q}$ -structures. It is clear that it is well-defined (i.e., unique, given  $\Phi$ ). It also satisfies the following useful “right cancellation” law (this amounts, in essence, to taking a reduct):

**Lemma 1.** *If  $A_1 + d_\Phi \cong A_2 + d_\Phi$  then  $A_1 \cong A_2$ .*

**Definition 15.** A relation symbol  $P_i$  in the signature  $P$  is *explicitly definable* in  $T$  just if

$$T \vdash P_i(\bar{x}) \leftrightarrow \theta$$

for some formula  $\theta$  in the language of the subsignature  $P \setminus P_i$ . We say that  $\theta$  is a *defining formula* for  $P_i$  in  $T$ .

---

<sup>8</sup>As explained below,  $d_\Theta$  will be an abbreviation for the set of definitions of the form  $\forall \bar{x} (P_i(\bar{x}) \leftrightarrow \theta_i)$ , where the  $\theta_i$  are  $L_Q$ -formulas.

**Definition 16.** A relation symbol  $P_i$  in the signature  $P$  is *implicitly definable* in  $T$  in  $L_P$  just in case, given any pair of models  $A, B \models T$ , with  $\text{dom}(A) = \text{dom}(B)$  and which assign the same extension to all  $P_j$  *except*  $P_i$ , we have  $(P_i)^A = (P_i)^B$ .

Beth's Theorem states that a relation symbol  $P_i$  is implicitly definable in  $T$  iff  $P_i$  is explicitly definable in  $T$ .<sup>9</sup>

**Definition 17.** Given a definition system  $d_\Phi$ , *the definitional extension* of  $T$  wrt  $\Phi$  is  $T + d_\Phi$ . We say that  $T^+$  in  $L_{P,Q}$  is a *definitional extension* of  $T$  in  $L_P$  just if

$$T^+ \equiv T + d_\Phi,$$

for some definition system  $d_\Phi$ , where  $\Phi$  is some set of  $L_P$ -formulas.

So, each new relation symbol  $Q_i$  is explicitly defined in  $T + d_\Phi$  and the defining formula for  $Q_i(\bar{x})$  is simply  $\phi_i$ . Analogous to what we did with structures, we are going to treat the definitional extension map

$$T \mapsto T + d_\Phi$$

for theories as a unary operator  $+d_\Phi$  (indexed by  $\Phi$ ), taking us from  $L_P$ -theories to  $L_{P,Q}$ -theories. Again, it is well-defined (i.e., unique, given  $\Phi$ ) and satisfies an analogous "right cancellation" law for  $L_P$ -theories:

**Lemma 2.** *The following are straightforward consequences of the definitions:*

- (1) *If  $T_1 + d_\Phi \equiv T_2 + d_\Phi$  then  $T_1 \equiv T_2$ .*
- (2)  *$A \models T$  iff  $A + d_\Phi \models T + d_\Phi$ .*
- (3) *If  $B \models T + d_\Phi$ , then  $B = A + d_\Phi$ , for some  $A \models T$ .*

Before moving on to translations, we give three standard lemmas about definitional and conservative extensions (the converse of Lemma 3 is far from being true):

**Lemma 3.** *If  $T^+$  in  $L_{P,Q}$  is a definitional extension of  $T$  in  $L_P$ , then  $T^+$  is a conservative extension of  $T$  wrt  $L_P$ -formulas.*

**Lemma 4.**  *$T + d_\Phi$  is a conservative extension of  $T$  for  $L_P$ -formulas.*

**Lemma 5.** *Let  $T$  in  $L_P$  and  $T^+$  in  $L_{P,Q}$  be such that  $T \subseteq T^+$ . Suppose that, for any model  $A \models T$ , there is an expansion  $A^+ \models T^+$ . Then  $T^+$  is a conservative extension of  $T$  for  $L_P$ -formulas.*

---

<sup>9</sup>For a proof, based on Craig's interpolation lemma, see Boolos and Jeffrey [3], Ch. 25.

## 4. Translation

**Definition 18.** Let a definition system  $d_\Phi$  be given. Define the *translation*, induced by  $\Phi$

$$\tau_\Phi : L_Q \rightarrow L_P$$

as follows. For symbols  $Q_i$ , variables  $x, y, \bar{x}$ , and for  $L_Q$ -formulas  $\alpha, \alpha_1, \alpha_2$ : #

|     |                                   |    |  |
|-----|-----------------------------------|----|--|
| (1) | $\tau_\Phi(Q_i(\bar{x}))$         | := | $(\phi_i)'$                                  |
| (2) | $\tau_\Phi(x = y)$                | := | $(x = y)$                                    |
| (3) | $\tau_\Phi(\neg\alpha)$           | := | $\neg\tau_\Phi(\alpha)$                      |
| (4) | $\tau_\Phi(\alpha_1 \# \alpha_2)$ | := | $\tau_\Phi(\alpha_1) \# \tau_\Phi(\alpha_2)$ |
| (5) | $\tau_\Phi(\mathbf{q}x\alpha)$    | := | $\mathbf{q}x\tau_\Phi(\alpha)$ .             |

is any binary connective,  $\mathbf{q}$  is a quantifier and  $(\phi_i)'$  is the result of ensuring that the free variables appearing  $\phi_i$  are relabelled, to match those of  $Q_i(\bar{x})$ . We call  $\tau_\Phi$  the *translation* induced by  $\Phi$ . It maps from the new language  $L_Q$  back to the original language  $L_P$ .<sup>10</sup>

**Lemma 6.** For any  $\alpha, \beta \in L_Q$ , if  $\alpha \vdash \beta$  then  $\tau_\Phi(\alpha) \vdash \tau_\Phi(\beta)$ .

*Proof.* We will prove this using a lemma below. Suppose  $\tau_\Phi(\alpha) \not\vdash \tau_\Phi(\beta)$ . This gives us a model  $A \models \tau_\Phi(\alpha)$  and  $A \not\models \tau_\Phi(\beta)$ . By Lemma 15(1) below,  $D_\Phi A \models \alpha$  and  $D_\Phi A \not\models \beta$ . So,  $\alpha \not\vdash \beta$ .  $\square$

Lemma 6 is a general property of translations, but its converse is not true in general.

Corresponding to a translation  $\tau_\Phi : L_Q \rightarrow L_P$  is its “lift”  $\tau_\Phi^+ : L_{P,Q} \rightarrow L_P$  from the combined language  $L_{P,Q}$  down to  $L_P$ :

**Definition 19.** Let  $\Phi$  be given, along with definition system  $d_\Phi$ . Define the *lifted translation*  $\tau_\Phi^+$  induced by  $\Phi$

$$\tau_\Phi^+ : L_{P,Q} \rightarrow L_P$$

as follows. For symbols  $Q_i, P_i$ , variables  $x, y, \bar{x}$ : Along with the requirement

|     |                             |    |                |
|-----|-----------------------------|----|----------------|
| (1) | $\tau_\Phi^+(Q_i(\bar{x}))$ | := | $(\phi_i)'$    |
| (2) | $\tau_\Phi^+(x = y)$        | := | $(x = y)$      |
| (3) | $\tau_\Phi^+(P_i(\bar{x}))$ | := | $P_i(\bar{x})$ |

that  $\tau_\Phi^+$  commutes with the logical operators on the full language  $L_{P,Q}$ .

Thus, the translation  $\tau_\Phi$  is the *restriction* to  $L_Q$  of its lift,  $\tau_\Phi^+$ .

Note that because the translations we are interested in always act as the identity on equations, it is always the case that  $\vdash \alpha \leftrightarrow \tau_\Phi(\alpha)$  if  $\alpha$  is an equation. Thus, in inductive proofs establishing biconditionals of the form  $\alpha \leftrightarrow \tau_\Phi(\alpha)$ , we only need to check the condition holds for atomic formulas which are not identity formulas.

<sup>10</sup>The clauses (3)–(5) are usually read as saying “ $\tau_\Phi$  commutes with the logical operators”.



**Lemma 7.** *We have:*

---

|     |   |   |
|-----|---|---|
| (1) | $A + d_{\Phi} \models \alpha \leftrightarrow \tau_{\Phi}(\alpha)$   | <i>for any <math>\alpha \in L_Q</math>.</i>     |
| (2) | $A + d_{\Phi} \models \alpha \leftrightarrow \tau_{\Phi}^+(\alpha)$ | <i>for any <math>\alpha \in L_{P,Q}</math>.</i> |
| (3) | $T + d_{\Phi} \vdash \alpha \leftrightarrow \tau_{\Phi}(\alpha)$    | <i>for any <math>\alpha \in L_Q</math>.</i>     |
| (4) | $T + d_{\Phi} \vdash \alpha \leftrightarrow \tau_{\Phi}^+(\alpha)$  | <i>for any <math>\alpha \in L_{P,Q}</math>.</i> |

---

*Proof.* For (2), we reason by induction. Let  $\alpha$  be an atomic  $L_{P,Q}$ -formula. As noted above, for equations ( $x = y$ ), the translation  $\tau_{\Phi}(x = y)$  is trivially ( $x = y$ ): so  $\alpha \leftrightarrow \tau_{\Phi}(\alpha)$  always holds for equations. Suppose  $\alpha$  is an atomic formula of the form  $Q_i(\bar{x})$ . Then its translation  $\tau_{\Phi}^+(\alpha)$  is  $\phi_i$ . Since we have  $A + d_{\Phi} \models \forall \bar{x}(Q_i(\bar{x}) \leftrightarrow \phi_i)$ , we have:  $A + d_{\Phi} \models Q_i(\bar{x}) \leftrightarrow \tau_{\Phi}(Q_i(\bar{x}))$ . Instead let  $\alpha$  be  $P_i(\bar{x})$ . Then its translation  $\tau_{\Phi}^+(\alpha)$  is simply  $\alpha$ . So,  $A + d_{\Phi} \models \alpha \leftrightarrow \tau_{\Phi}^+(\alpha)$ . The other cases are shown by induction on the construction of  $\alpha$ .

For (1), the result follows from (2), by restricting to  $L_Q$ -formulas (since  $\tau_{\Phi}^+(\alpha) = \tau_{\Phi}(\alpha)$  for  $\alpha \in L_Q$ ).

For (4), reasoning by induction, let  $\alpha$  be an atomic  $L_P, Q$ -formula. If  $\alpha$  is atomic, then the condition is trivial. Suppose  $\alpha$  has the form  $P_i(\bar{x})$ . Again, he condition is trivial, since  $\tau_{\Phi}^+(P_i(\bar{x}))$  is  $P_i(\bar{x})$ . Instead, suppose  $\alpha$  has the form  $Q_i(\bar{x})$ . Then its translation  $\tau_{\Phi}^+(\alpha)$  is  $\phi_i$ .  $Q_i(\bar{x}) \leftrightarrow \phi_i$  is a theorem of  $T + d_{\Phi}$ , by construction. The other cases are shown by induction on the construction of  $\alpha$ .

For (3), the result follows from (4), by restricting to  $L_Q$ -formulas. □

**Definition 20.** Let  $\tau_{\Phi} : L_Q \rightarrow L_P$  be the translation induced by  $d_{\Phi}$ . If  $T_2$  is a theory in  $L_Q$ , then the *image* of  $T_2$  under  $\tau_{\Phi}$  is the set of  $L_P$ -sentences:

$$\tau_{\Phi}[T_2] := \{\tau_{\Phi}(\alpha) \in L_P \mid \alpha \in T_2\}$$

If  $T_1$  is a theory in  $L_P$ , then the *pre-image* of  $T_1$  under  $\tau_{\Phi}$  is the set of  $L_Q$ -sentences:

$$(\tau_{\Phi})^{-1}[T_1] := \{\beta \in L_Q \mid \tau_{\Phi}(\beta) \in T_1\}$$

Similarly, if  $\Theta = \{\theta_i\}_{i \in I_P}$  is a set of  $L_Q$ -formulas and  $d_{\Theta}$  is the corresponding definition system over  $\Theta$  (for the primitives  $P_i$  of  $L_P$ ), we can define a translation

$$\tau_{\Theta} : L_P \rightarrow L_Q$$

by requiring that  $\tau_{\Theta}$  commute with the logical operators and, for atomic  $L_P$ -formulas: Likewise, we can also define the lifted translation  $\tau_{\Theta}^+ : L_{P,Q} \rightarrow L_Q$ .

---

|      |                                  |               |
|------|----------------------------------|---------------|
| (i)  | $\tau_{\Theta}(P_i(\bar{x})) :=$ | $(\theta_i)'$ |
| (ii) | $\tau_{\Theta}(x = y) :=$        | $(x = y)$     |

---

**Definition 21.** Let  $\tau : L_Q \rightarrow L_P$  be a translation. Let  $T_1$  be a theory in  $L_P$  and  $T_2$  be a theory in  $L_Q$ . Then we say:

- (1)  $\tau$  interprets  $T_2$  into  $T_1$  iff  $T_1 \vdash \tau[T_2]$ .
- (2)  $\tau$  faithfully interprets  $T_2$  into  $T_1$  iff  $T_1 \equiv \tau[T_2]$ .

One may compare the condition  $T_1 \equiv \tau[T_2]$  with Visser's definition of faithful interpretability:

We write  $K : U \triangleleft_{\text{faith}} V$  for:  $K$  is a faithful interpretation of  $U$  in  $V$ .

This means that: for all  $U$ -sentences  $A$ , we have:  $U \vdash A$  iff  $V \vdash A^{\tau_K}$ .

(Visser [19], p. 6).

Thus  $\tau_\Phi : T_2 \triangleleft_{\text{faith}} T_1$  holds iff, for all  $\alpha \in L_{T_2}$ , we have:  $T_2 \vdash \alpha$  iff  $T_1 \vdash \tau_\Phi(\alpha)$ .

Thus,  $\tau_\Phi : T_2 \triangleleft_{\text{faith}} T_1$  iff  $T_1 \equiv \tau_\Phi[T_2]$ . This establishes:

**Lemma 8.**  $\tau_\Phi$  faithfully interprets  $T_2$  into  $T_1$  iff, for all  $\alpha \in L_Q$ , we have:  $T_2 \vdash \alpha$  iff  $T_1 \vdash \tau_\Phi(\alpha)$ .

**Definition 22.** Let  $\tau_\Phi : L_Q \rightarrow L_P$  and  $\tau_\Theta : L_P \rightarrow L_Q$  be translations induced by  $d_\Phi$  and  $d_\Theta$ . Let  $T_1$  be an  $L_P$  theory. Let  $T_2$  be an  $L_Q$  theory. Then  $\tau_\Theta$  is an *right inverse* of  $\tau_\Phi$  in  $T_1$  iff, for any  $\alpha \in L_P$ ,

$$T_1 \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$$

We write this more suggestively as:

$$(\tau_\Phi \tau_\Theta = 1)_{T_1}$$

And  $\tau_\Theta$  is an *left inverse* of  $\tau_\Phi$  in  $T_2$  iff, for any  $\beta \in L_Q$ ,

$$T_2 \vdash \beta \leftrightarrow \tau_\Theta(\tau_\Phi(\beta))$$

Likewise, we write this more suggestively as:

$$(\tau_\Theta \tau_\Phi = 1)_{T_2}$$

The following two lemmas are easy to prove, and yet hold to the key to much that follows. Both lemmas use “invertibility conditions”, of the form:

$$A + d_\Phi \models d_\Theta$$

$$T + d_\Phi \vdash d_\Theta.$$

As we see later, these conditions express a very strong constraint on the set  $\Phi$  of defining  $L_P$ -formulas involved—the property of being a “*representation basis*” for  $A$  (or  $T$ ) with inverse  $\Theta$ .

**Lemma 9.** *Suppose that  $A + d_\Phi \models d_\Theta$ . Then:*

---

|     |  |   |
|-----|--|---|
| (1) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Phi(\alpha)$              | <i>for <math>\alpha \in L_Q</math>.</i>     |
| (2) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Phi^+(\alpha)$            | <i>for <math>\alpha \in L_{P,Q}</math>.</i> |
| (3) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Theta(\alpha)$            | <i>for <math>\alpha \in L_P</math>.</i>     |
| (4) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ | <i>for <math>\alpha \in L_Q</math>.</i>     |
| (5) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ | <i>for <math>\alpha \in L_P</math>.</i>     |
| (6) | $A + d_\Phi \models \tau_\Phi(\theta_i) \leftrightarrow \theta_i$ .        |   |
| (7) | $A + d_\Phi \models \tau_\Theta(\phi_i) \leftrightarrow \phi_i$ .          |   |
| (8) | $A \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$          | <i>for <math>\alpha \in L_P</math>.</i>     |
| (9) | $A + d_\Phi \models \alpha \leftrightarrow \tau_\Theta^+(\alpha)$          | <i>for <math>\alpha \in L_{P,Q}</math>.</i> |

---

*Proof.* Claims (1) and (2) are already established in Lemma 7(1,2) and do not need the side condition. (They are included for convenience of reference.)

For (3), the proof is analogous to the proof of Lemma 7(1), but using the fact that  $A + d_\Phi \models d_\Theta$ . Reasoning by induction, let  $\alpha$  be an atomic  $L_P$ -sentence, say  $P_i(\bar{x})$ . Then its translation  $\tau_\Theta(\alpha)$  is  $\theta_i$ . And  $A + d_\Phi \models \forall \bar{x}(P_i(\bar{x}) \leftrightarrow \theta_i)$ . So  $A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \theta_i$ . That is,  $A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \tau_\Theta(P_i(\bar{x}))$ . The other cases are shown by induction on the construction of  $\alpha$ .

For (4), we already have that  $A + d_\Phi \models \alpha \leftrightarrow \tau_\Phi(\alpha)$ , for any  $\alpha \in L_Q$ . But  $\tau_\Phi(\alpha)$  is an  $L_P$ -formula. So, by (3), we have:  $A + d_\Phi \models \tau_\Phi(\alpha) \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ . So,  $A + d_\Phi \models \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ , as required.

For (5), the reasoning is analogous to that for (4). And (6) and (7) are merely applications of (1) and (3).

For (8), using (5), we have  $A + d_\Phi \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ , for any  $\alpha \in L_P$ . But  $A + d_\Phi$  is an expansion of the  $L_P$ -structure  $A$ , and  $\alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$  is an  $L_P$ -formula. Thus,  $A \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ .

For (9), the proof is analogous to the proof of (2), but is applied to the “lift”  $\tau_\Theta^+ : L_{P,Q} \rightarrow L_Q$  of  $\tau_\Theta$ . Reasoning by induction, let  $\alpha$  be an atomic  $L_P$ -sentence, say  $P_i(\bar{x})$ . Then its translation  $\tau_\Theta^+(\alpha)$  is  $\theta_i$ . And  $A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \theta_i$ , since  $A + d_\Phi \models d_\Theta$ . Instead, let  $\alpha$  be an atomic  $L_Q$ -sentence, say  $Q_i(\bar{x})$ . Then its translation  $\tau_\Theta^+(\alpha)$  is  $Q_i(\bar{x})$ . Trivially,  $A + d_\Phi \models Q_i(\bar{x}) \leftrightarrow Q_i(\bar{x})$ . The other cases are shown by induction on the construction of  $\alpha$ .  $\square$

The following lemma, and the corresponding proofs, is a near repetition of the previous one, except that it deals with theories:

**Lemma 10.** *Suppose that  $T + d_\Phi \vdash d_\Theta$ . Then:*

---

|     |   |                                   |
|-----|---|-----------------------------------|
| (1) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Phi(\alpha)$              | <i>for</i> $\alpha \in L_Q$ .     |
| (2) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Phi^+(\alpha)$            | <i>for</i> $\alpha \in L_{P,Q}$ . |
| (3) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Theta(\alpha)$            | <i>for</i> $\alpha \in L_P$ .     |
| (4) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ | <i>for</i> $\alpha \in L_Q$ .     |
| (5) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ | <i>for</i> $\alpha \in L_P$ .     |
| (6) | $T + d_\Phi \vdash \tau_\Phi(\theta_i) \leftrightarrow \theta_i$ .        |                                   |
| (7) | $T + d_\Phi \vdash \tau_\Theta(\phi_i) \leftrightarrow \phi_i$ .          |                                   |
| (8) | $T \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$          | <i>for</i> $\alpha \in L_P$ .     |
| (9) | $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Theta^+(\alpha)$          | <i>for</i> $\alpha \in L_{P,Q}$ . |

---

*Proof.* Essentially, a repetition of the proofs for Lemma 9. □

**Lemma 11.** *Suppose  $T_1$  is an  $L_P$ -theory and  $T_2$  is an  $L_Q$ -theory. Then:*

- 
- |     |  |
|-----|--|
| (1) | <i>If</i> $(\tau_\Phi \tau_\Theta = 1)_{T_1}$ <i>then</i> $T_1 + d_\Phi \vdash d_\Theta$ . |
| (2) | <i>If</i> $(\tau_\Theta \tau_\Phi = 1)_{T_2}$ <i>then</i> $T_2 + d_\Theta \vdash d_\Phi$ . |
- 

*Proof.* For (1), we assume  $\tau_\Theta$  is a right inverse of  $\tau_\Phi$  in  $T_1$ . I.e., for any  $\alpha \in L_P$ ,  $T_1 \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ . Thus,  $T_1 \vdash P_i(\bar{x}) \leftrightarrow \tau_\Phi(\tau_\Theta(P_i(\bar{x})))$ . Since  $\tau_\Theta(P_i(\bar{x}))$  is  $\theta_i$ , we have  $T_1 \vdash P_i(\bar{x}) \leftrightarrow \tau_\Phi(\theta_i)$ . Taking the definitional extension,  $T_1 + d_\Phi \vdash P_i(\bar{x}) \leftrightarrow \tau_\Phi(\theta_i)$ . Now, for any  $L_Q$ -formula  $\beta$ , we have  $T_1 + d_\Phi \vdash \beta \leftrightarrow \tau_\Phi(\beta)$  from Lemma 7(2). So, since  $\theta_i \in L_Q$ ,  $T_1 + d_\Phi \vdash \theta_i \leftrightarrow \tau_\Phi(\theta_i)$ . Thus,  $T_1 + d_\Phi \vdash P_i(\bar{x}) \leftrightarrow \theta_i$ , as required.

We obtain (2) by relabelling everything ( $T_2$  is now a theory in  $L_Q$ ). □

Lemma 11 says that if  $\tau_\Theta$  is a right-inverse of  $\tau_\Phi$ , relative to  $T$ , then every relation symbol  $P_i$  from the original language can be explicitly defined from the  $\theta_i$ . In a sense, the original definition system,  $d_\Phi$  is a kind of *inverse* of  $d_\Theta$ .

## 5. Definitional Images

**Definition 23.** Let  $A$  be an  $L_P$ -structure. Then the  $L_Q$ -structure  $D_\Phi A$  is defined by:

$$D_\Phi A := (A + d_\Phi) \upharpoonright_{L_Q}$$

$D_\Phi A$  is called the *definitional image* of  $A$  with respect to  $\Phi$ .

Immediately, we see that the following three conditions provide an equivalent characterization of  $D_\Phi A$ :

- 
- |     |  |
|-----|--|
| (1) | $\text{dom}(D_\Phi A) = \text{dom}(A)$                       |
| (2) | $x^{D_\Phi A} = x^A$ , for any variable $x$                  |
| (3) | $(Q_i)^{D_\Phi A} = \phi_i(A)$ , for each new symbol $Q_i$ . |
-

**Lemma 12.** For any  $\alpha \in L_Q$ :  $\alpha(D_\Phi A) = \alpha(A + d_\Phi)$ .

Although  $A + d_\Phi$  is by construction a definitional expansion of  $A$ , it is by no means automatically true that  $A + d_\Phi$  is a definitional expansion of  $D_\Phi A$ . The requirement for this to hold is that each primitive relation  $(P_i)^A$  be definable in  $D_\Phi A$ , by some formula, say  $\theta_i$ .

Turning to theories, we introduce analogous concepts:

**Definition 24.** The *definitional image* of  $T$ , with respect to  $\Phi$ , is the restriction of the deductive closure of  $T + d_\Phi$  to the new language  $L_Q$ . The definitional image of  $T$  with respect to  $\Phi$  is denoted  $D_\Phi T$ . That is,

$$D_\Phi T := \text{DedCl}(T + d_\Phi) \upharpoonright_{L_Q} = \{\beta \in L_Q \mid T + d_\Phi \vdash \beta\}$$

The definitional image  $D_\Phi T$  of a theory  $T$  in  $L_P$  is, essentially, the *pre-image*  $(\tau_\Phi)^{-1}[T]$  of  $T$  under the translation  $\tau_\Phi$  induced by  $\Phi$ :

**Lemma 13.**  $(\tau_\Phi)^{-1}[T] \equiv D_\Phi T$ .

*Proof.* Suppose  $D_\Phi T \vdash \beta$ , for  $\beta \in L_Q$ . Thus,  $T + d_\Phi \vdash \beta$ . But  $T + d_\Phi \vdash \beta \leftrightarrow \tau_\Phi(\beta)$ , by Lemma 10(1). So,  $T + d_\Phi \vdash \tau_\Phi(\beta)$ . Since  $T + d_\Phi$  conservatively extends  $T$  for  $L_P$ -formulas,  $T \vdash \tau_\Phi(\beta)$ . Hence,  $\beta \in (\tau_\Phi)^{-1}[T]$ . And so  $(\tau_\Phi)^{-1}[T] \vdash \beta$ . Conversely, suppose  $\beta \in (\tau_\Phi)^{-1}[T]$ . So,  $T \vdash \tau_\Phi(\beta)$ . So,  $T + d_\Phi \vdash \tau_\Phi(\beta)$ . But  $T + d_\Phi \vdash \beta \leftrightarrow \tau_\Phi(\beta)$ , by Lemma 10(1) again. Thus,  $D_\Phi T \vdash \beta$ .  $\square$

**Definition 25.** Let  $A$  be an  $L_P$ -structure,  $B$  an  $L_Q$ -structure,  $T_1$  an  $L_P$ -theory and  $T_2$  an  $L_Q$ -theory. Then we say:

- 
- (1)  $\Phi$  defines  $B$  in  $A$  iff  $B \cong D_\Phi A$ .
  - (2)  $\Phi$  proof-theoretically defines  $T_2$  in  $T_1$  iff  $T_2 \equiv D_\Phi T_1$ .
  - (3)  $\Phi$  model-theoretically defines  $T_2$  in  $T_1$  iff  $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$ .
- 

The second of these, (2), amounts to saying that  $T_2 \vdash D_\Phi T_1$  and  $D_\Phi T_1 \vdash T_2$ . The third is equivalent to saying that

$$\text{Mod}(T_2) = (\text{Mod}(T_1 + d_\Phi)) \upharpoonright_{L_Q} .$$

Note that the restriction  $\upharpoonright_{L_Q}$  is taken *after* the models are extracted from the definitional extension  $T_1 + d_\Phi$ . If the restriction is taken first, we get the rather different set  $\text{Mod}((\text{DedCl}(T_1 + d_\Phi)) \upharpoonright_{L_Q})$  of models: i.e.,  $\text{Mod}(D_\Phi T_1)$ . Indeed, this is generally a superset of  $D_\Phi[\text{Mod}(T_1)]$ . As Lemma 15(3) will show, we have:  $D_\Phi[\text{Mod}(T)] \subseteq \text{Mod}(D_\Phi T)$ .

## 6. Some Book-Keeping Lemmas

We next provide several groups of “book-keeping” lemmas about translations and definitional images. The first, Lemma 14, primarily concerns theories. The second group, Lemma 15, concerns semantics and models. The third group (in

particular, Lemma 16) establishes five calculationaly useful equivalences for the “definition invertibility condition” on structures:

$$\boxed{A + d_\Phi \models d_\Theta.}$$

The fourth group establishes some analogous results for theories, including the main equivalence (Lemma 19) for the “definition invertibility condition” on theories:

$$\boxed{T + d_\Phi \models d_\Theta.}$$

**Lemma 14.** *Each of the following is true:*

- 
- |     |   |
|-----|---|
| (1) | $T + d_\Phi$ is a conservative extension of $D_\Phi T$ for $L_Q$ -formulas.                                   |
| (2) | $D_\Phi T_1 \vdash T_2$ iff $T_1 \vdash \tau_\Phi[T_2]$ .   |
| (3) | $\tau_\Phi$ interprets $T_2$ into $T_1$ iff $D_\Phi T_1 \vdash T_2$ .   |
| (4) | $\tau_\Phi$ interprets $D_\Phi T$ into $T$ .  |
| (5) | $D_\Phi T = \{\beta \in L_Q \mid T \vdash \tau_\Phi(\beta)\}$ .   |
| (6) | $D_\Phi T \vdash \beta$ iff $T \vdash \tau_\Phi(\beta)$ (for any $\beta \in L_Q$ ).                           |
| (7) | $T_2 \equiv D_\Phi T_1$ iff, for all $\beta \in L_Q$ : $T_2 \vdash \beta$ iff $T_1 \vdash \tau_\Phi(\beta)$ . |
| (8) | $T_2 \equiv D_\Phi T_1$ iff $\tau_\Phi$ faithfully interprets $T_2$ into $T_1$ .                              |
- 

*Proof.* To establish (1), note that  $D_\Phi T$  is simply the restriction of the deductive closure of  $T + d_\Phi$  to  $L_Q$ -sentences. So if  $T + d_\Phi \vdash \alpha$ , where  $\alpha \in L_Q$ , then  $D_\Phi T \vdash \alpha$ , as required. Statement (2) simply expresses the relationship between images and pre-images. Statement (3) is an immediate corollary of (2). Statement (4) is an immediate corollary of (3).

For (5), note that  $D_\Phi T = (\tau_\Phi)^{-1}[T]$ , from Lemma 13 above. Thus,  $D_\Phi T = \{\beta \in L_Q \mid T \vdash \tau_\Phi(\beta)\}$ . For (6), let  $\beta \in L_Q$ . Then, from (5), we have  $D_\Phi T \vdash \beta$  iff  $T \vdash \tau_\Phi(\beta)$ , as required.

For (7), suppose first that  $T_2 \equiv D_\Phi T_1$ . So, for all  $\beta \in L_Q$ ,  $T_2 \vdash \beta$  iff  $D_\Phi T_1 \vdash \beta$ . But from (6),  $D_\Phi T_1 \vdash \beta$  iff  $T_1 \vdash \tau_\Phi(\beta)$ . So, for all  $\beta \in L_Q$ ,  $T_2 \vdash \beta$  iff  $T_1 \vdash \tau_\Phi(\beta)$ , as required.

Instead suppose that for all  $\beta \in L_Q$ , we have:  $T_2 \vdash \beta$  iff  $T_1 \vdash \tau_\Phi(\beta)$ . Then reason as follows: So,  $T_2 \equiv D_\Phi T_1$ , as required.

---

|                    |   |
|--------------------|---|
| $T_2 \vdash \beta$ | iff $T_1 \vdash \tau_\Phi(\beta)$ (from the hypothesis)   |
|                    | iff $T_1 + d_\Phi \vdash \beta$ (as $T_1 + d_\Phi \vdash \beta \leftrightarrow \tau_\Phi(\beta)$ ; Lemma 10(1)) |
|                    | iff $D_\Phi T_1 \vdash \beta$ (from the definition of $D_\Phi T_1$ ; i.e., Definition 24).                      |

---

(8) is an immediate corollary of (7) and Lemma 8. □

**Lemma 15.** *Each of the following holds:*

- 
- |     |  |
|-----|--|
| (1) | For any $\alpha \in L_Q$ : $D_\Phi A \models \alpha$ iff $A \models \tau_\Phi(\alpha)$ .   |
| (2) | If $A \models T$ , then $D_\Phi A \models D_\Phi T$ .                                      |
| (3) | $D_\Phi[\text{Mod}(T)] \subseteq \text{Mod}(D_\Phi T)$ .                                   |
| (4) | Let $T_2$ be a theory in $L_Q$ . If $A + d_\Phi \models T_2$ then $D_\Phi A \models T_2$ . |
- 

*Proof.* (1): By induction. Suppose  $\alpha = Q_i(\bar{x})$ . Thus,

$$\begin{aligned}
 D_\Phi A \models \alpha &\Leftrightarrow ((x_1)^{D_\Phi A}, \dots, (x_n)^{D_\Phi A}) \in (Q_i)^{D_\Phi A} \\
 &\Leftrightarrow ((x_1)^A, \dots, (x_n)^A) \in \phi_i(A) \\
 &\Leftrightarrow A \models \phi_i \\
 &\Leftrightarrow A \models \tau_\Phi(Q_i(\bar{x})) \\
 &\Leftrightarrow A \models \tau_\Phi(\alpha).
 \end{aligned}$$

The case of equality, and the induction steps are routine.

(2): Let  $A \models T$  and  $D_\Phi T \vdash \alpha$ , for some  $\alpha \in L_Q$ . We want to show  $D_\Phi A \models \alpha$ . We do this using the translation  $\tau_\Phi$ . Since  $D_\Phi T \vdash \alpha$ , it follows, by Lemma 14(4), that  $T \vdash \tau_\Phi(\alpha)$ . And so,  $A \models \tau_\Phi(\alpha)$ . From (1),  $D_\Phi A \models \alpha$  iff  $A \models \tau_\Phi(\alpha)$ . So,  $D_\Phi A \models \alpha$ , as required.

(3): Let  $B \in D_\Phi[\text{Mod}(T)]$ . So, there is some  $A \models T$  with  $B \cong D_\Phi A$ . Since  $A \models T$ , we have  $D_\Phi A \models D_\Phi T$ , by (2). So,  $B \in \text{Mod}(D_\Phi T)$ .

(4): Suppose that for any  $\beta$  such that  $T_2 \vdash \beta$ , we have  $A + d_\Phi \models \beta$ . We want to show that for any  $\beta$  such that  $T_2 \vdash \beta$ , we have  $D_\Phi A \models \beta$ . So, let  $T_2 \vdash \beta$  and  $A + d_\Phi \models \beta$ . By Lemma 7(1),  $A + d_\Phi \models \tau_\Phi(\beta)$ . So,  $A \models \tau_\Phi(\beta)$ . By (1),  $D_\Phi A \models \beta$ .  $\square$

Neither the converse of Lemma 15(2) nor the converse of the inclusion in Lemma 15(3) is true.

**Lemma 16.** *The following are equivalent:*

- 
- |     |   |
|-----|---|
| (1) | $A + d_\Phi \models d_\Theta$ .   |
| (2) | For all $P_i$ , for all $Q_j$ : $(P_i)^{A+d_\Phi} = \theta_i(A + d_\Phi)$ & $(Q_j)^{A+d_\Phi} = \phi_j(A + d_\Phi)$ . |
| (3) | $A + d_\Phi \cong D_\Phi A + d_\Theta$ .  |
| (4) | $D_\Theta D_\Phi A = A$ .   |
| (5) | $A \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ (with $\alpha \in L_P$ )                            |
- 

*Proof.* For (1)  $\Leftrightarrow$  (2). This is simply unwinding the definitions.

For (1)  $\Rightarrow$  (3), suppose  $A + d_\Phi \models d_\Theta$ . Then, from (2), we have, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = \theta_i(A + d_\Phi) \ \& \ (Q_j)^{A+d_\Phi} = \phi_j(A + d_\Phi).$$


---

So, since  $\theta_i(A + d_\Phi) = \theta_i(D_\Phi A + d_\Theta) = (P_i)^{D_\Phi A + d_\Theta}$ , we have, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = (P_i)^{D_\Phi A+d_\Theta} \ \&\mathcal{L} \ (Q_j)^{A+d_\Phi} = \phi_j(A+d_\Phi).$$


---

So, since  $\phi_j(A+d_\Phi) = (Q_j)^{D_\Phi A+d_\Theta}$ , we have, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = (P_i)^{D_\Phi A+d_\Theta} \ \&\mathcal{L} \ (Q_j)^{A+d_\Phi} = (Q_j)^{D_\Phi A+d_\Theta}.$$


---

So,  $A+d_\Phi \cong D_\Phi A+d_\Theta$ , as required.

For (3)  $\Rightarrow$  (1). This reverses the reasoning (1)  $\Rightarrow$  (3). Suppose  $A+d_\Phi \cong D_\Phi A+d_\Theta$ . That is, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = (P_i)^{D_\Phi A+d_\Theta} \ \&\mathcal{L} \ (Q_j)^{A+d_\Phi} = (Q_j)^{D_\Phi A+d_\Theta}.$$


---

Now  $(Q_j)^{D_\Phi A+d_\Theta} = (Q_j)^{D_\Phi A} = \phi_j(A) = \phi_j(A+d_\Phi)$ . So, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = (P_i)^{D_\Phi A+d_\Theta} \ \&\mathcal{L} \ (Q_j)^{A+d_\Phi} = \phi_j(A+d_\Theta).$$


---

But trivially,  $D_\Phi A+d_\Theta \models d_\Theta$ . So,  $D_\Phi A+d_\Theta \models P_i(\bar{x}) \leftrightarrow \theta_i$ , for each  $P_i$ . Thus,

$$(P_i)^{D_\Phi A+d_\Theta} = \theta_i(D_\Phi A+d_\Theta).$$

But we also have

$$D_\Phi A+d_\Theta \cong A+d_\Theta$$

Therefore,

$$(P_i)^{D_\Phi A+d_\Theta} = \theta_i(A+d_\Theta)$$

Thus, for all  $P_i$ , for all  $Q_j$ ,

---


$$(P_i)^{A+d_\Phi} = \theta_i(A+d_\Theta) \ \&\mathcal{L} \ (Q_j)^{A+d_\Phi} = \phi_j(A+d_\Theta).$$


---

And therefore,  $A+d_\Phi \models d_\Theta$ , as required.

For (1)  $\Rightarrow$  (4). Suppose  $A+d_\Phi \models d_\Theta$ . So, from (3) above, we have:

$$A+d_\Phi \cong D_\Phi A+d_\Theta$$

Therefore,  $D_\Phi A+d_\Theta \models d_\Phi$ . A proof analogous to that of (3) above gives us that if  $B+d_\Theta \models d_\Phi$  then  $B+d_\Theta \cong D_\Theta B+d_\Phi$ , and so, if  $D_\Phi A+d_\Theta \models d_\Phi$  then  $D_\Phi A+d_\Theta \cong D_\Theta D_\Phi A+d_\Phi$ . And so, we may conclude,

$$D_\Phi A+d_\Theta \cong D_\Theta D_\Phi A+d_\Phi$$

And thus,

$$A+d_\Phi \cong D_\Theta D_\Phi A+d_\Phi$$

And right cancellation gives,

$$A \cong D_\Theta D_\Phi A$$



But since the map  $D_\Phi$  leaves the domain invariant, we have  $A = D_\Theta D_\Phi A$ .

For (4)  $\Rightarrow$  (5), suppose  $A = D_\Theta D_\Phi A$ . We have, applying Lemma 15(1) twice, for any  $\alpha \in L_P$ ,  $\beta \in L_Q$ :

$$\begin{aligned} A \models \tau_\Phi(\beta) &\text{ iff } D_\Phi A \models \beta \\ D_\Phi A \models \tau_\Theta(\alpha) &\text{ iff } D_\Theta D_\Phi A \models \alpha \end{aligned}$$

So, for  $\alpha \in L_P$ ,

$$A \models \tau_\Phi(\tau_\Theta(\alpha)) \text{ iff } D_\Theta D_\Phi A \models \alpha$$

But  $A = D_\Theta D_\Phi A$ , and so,

$$A \models \tau_\Phi(\tau_\Theta(\alpha)) \text{ iff } A \models \alpha$$

And thus,

$$A \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$$

as required.

For (5)  $\Rightarrow$  (1), assume  $A \models \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ . So,

$$A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \tau_\Phi(\tau_\Theta(P_i(\bar{x})))$$

But  $\tau_\Theta(P_i(\bar{x})) = \theta_i$ . So,

$$A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \tau_\Phi(\theta_i)$$

And  $\theta_i \in L_Q$ , and so, from Lemma 7(1),

$$A + d_\Phi \models P_i(\bar{x}) \leftrightarrow \theta_i$$

and thus,

$$A + d_\Phi \models d_\Theta$$

as required. □

Turning next to theories:

**Lemma 17.**  $D_\Phi T + d_\Theta \vdash \alpha \leftrightarrow \tau_\Theta(\alpha)$ , for all  $\alpha \in L_P$ .

*Proof.* We reason by induction on the construction of  $\alpha$ . Let  $\alpha = P_i(\bar{x})$ . Then we have:  $D_\Phi T + d_\Theta \vdash P_i(\bar{x}) \leftrightarrow \theta_i$ . But  $\tau_\Theta(P_i(\bar{x})) = \theta_i$ . And thus,  $D_\Phi T + d_\Theta \vdash P_i(\bar{x}) \leftrightarrow \tau_\Theta(P_i(\bar{x}))$ . The equality case and compound cases proceed routinely. (Notice this is analogous to Lemma 10(1).) □

**Lemma 18.** Suppose that  $T + d_\Phi \vdash d_\Theta$ . Then, for any  $\alpha \in L_Q$ :

---

|     |   |
|-----|---|
| (1) | $D_\Phi T \vdash \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ |
| (2) | $D_\Phi T + d_\Theta \vdash \alpha \leftrightarrow \tau_\Phi(\alpha)$ . |

---

*Proof.* For (1), we have, for any  $\beta \in L_Q$ , if  $T + d_\Phi \vdash \beta$ , then  $D_\Phi T \vdash \beta$ , since  $D_\Phi T$  is simply the restriction of the set of theorems of  $T + d_\Phi$  to  $L_Q$ . By Lemma 10(4), we have  $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ , for any  $\alpha \in L_Q$ . Now  $\alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$  is also in  $L_Q$ . Thus,  $D_\Phi T \vdash \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ .

For (2), from Lemma 17, we have, for any  $\alpha \in L_Q$ :  $D_\Phi T + d_\Theta \vdash \tau_\Phi(\alpha) \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ , since  $\tau_\Phi(\alpha) \in L_P$ . By (1), we have  $D_\Phi T + d_\Theta \vdash \alpha \leftrightarrow \tau_\Theta(\tau_\Phi(\alpha))$ . Hence,  $D_\Phi T + d_\Theta \vdash \alpha \leftrightarrow \tau_\Phi(\alpha)$ , as required.  $\square$

The next lemma is the most important result needed for Theorem 2 given in Sect. 8:

**Lemma 19.** *The following are equivalent*

---

|     |   |
|-----|---|
| (1) | $T + d_\Phi \vdash d_\Theta$ .            |
| (2) | $T + d_\Phi \equiv D_\Phi T + d_\Theta$ . |

---

*Proof.* (1)  $\Rightarrow$  (2). Suppose  $T + d_\Phi \vdash d_\Theta$ . We want to show

---

|     |   |
|-----|---|
| (a) | $T + d_\Phi \vdash D_\Phi T + d_\Theta$ . |
| (b) | $D_\Phi T + d_\Theta \vdash T + d_\Phi$ . |

---

First, for (a), let  $A + d_\Phi \models T + d_\Phi$ . So, since  $T + d_\Phi \vdash d_\Theta$ , we have  $A + d_\Phi \models d_\Theta$ . So, by Lemma 16(3), we have  $A + d_\Phi \cong D_\Phi A + d_\Theta$ . Since  $A \models T$ , we have  $D_\Phi A \models D_\Phi T$ . So,  $D_\Phi A + d_\Theta \models D_\Phi T + d_\Theta$ . Thus,  $A + d_\Phi \models D_\Phi T + d_\Theta$ . And therefore, since  $A$  was arbitrary,  $T + d_\Phi \vdash D_\Phi T + d_\Theta$ , as required.

For (b), we want to show that  $D_\Phi T + d_\Theta \vdash T + d_\Phi$ . That is, for for any  $L_{P,Q}$ -formula  $\alpha$ ,

If  $T + d_\Phi \vdash \alpha$  then  $D_\Phi T + d_\Theta \vdash \alpha$ .

First, note that we may relabel Lemma 9(2) in terms of some  $L_Q$ -structure  $B$  and definition system  $d_\Theta$ , rather than  $A$  and  $d_\Phi$ , to obtain: for any  $\alpha \in L_{P,Q}$ ,

---

|     |   |
|-----|---|
| (i) | $B + d_\Theta \models \alpha \leftrightarrow \tau_\Theta^+(\alpha)$ . |
|-----|---|

---

For a contradiction, suppose we have some  $\alpha \in L_{P,Q}$  such that

---

|       |   |
|-------|---|
| (ii)  | $T + d_\Phi \vdash \alpha$                |
| (iii) | $D_\Phi T + d_\Theta \not\vdash \alpha$ . |

---

Since  $T + d_\Phi \vdash d_\Theta$ , we have, from Lemma 10(9), that  $T + d_\Phi \vdash \alpha \leftrightarrow \tau_\Theta^+(\alpha)$  and since  $T + d_\Phi \vdash \alpha$ , we have:

---

|      |  |
|------|--|
| (iv) | $T + d_\Phi \vdash \tau_\Theta^+(\alpha).$ |
|------|--|

---

From (iii), there exists an  $L_Q$ -structure  $B \models D_\Phi T$  with  $B + d_\Theta \not\models \alpha$ . So, from (i),  $B + d_\Theta \not\models \tau_\Theta^+(\alpha)$ . And since  $\tau_\Theta^+(\alpha) \in L_Q$ , we have:

---

|     |  |
|-----|--|
| (v) | $B \not\models \tau_\Theta^+(\alpha).$ |
|-----|--|

---

Recall that  $D_\Phi T = \{\beta \in L_Q \mid T + d_\Phi \vdash \beta\}$ . So, since  $B \models D_\Phi T$ , it follows that, for any  $\beta \in L_Q$ , if  $T + d_\Phi \vdash \beta$ , then  $B \models \beta$ . Thus, if  $T + d_\Phi \vdash \tau_\Theta^+(\alpha)$ , then  $B \models \tau_\Theta^+(\alpha)$ . So, from (iv), we infer:

---

|      |                                    |
|------|------------------------------------|
| (vi) | $B \models \tau_\Theta^+(\alpha).$ |
|------|------------------------------------|

---

Contradiction.

(2)  $\Rightarrow$  (1). Let  $T + d_\Phi \equiv D_\Phi T + d_\Theta$ . Then  $T + d_\Phi \vdash D_\Phi T + d_\Theta$ . Thus,  $T + d_\Phi \vdash d_\Theta$ , as required.  $\square$

**Lemma 20.**  $D_\Theta D_\Phi T \vdash \alpha$  iff  $T \vdash \tau_\Phi(\tau_\Theta(\alpha))$ , for  $\alpha \in L_P$ .

*Proof.* If we examine the definitions of  $D_\Phi T$  and  $D_\Theta T_2$  (where  $T_2$  is in  $L_Q$ ), we get

$$\begin{aligned} D_\Phi T &= \{\alpha \in L_Q \mid T \vdash \tau_\Phi(\alpha)\} \\ D_\Theta T_2 &= \{\alpha \in L_P \mid T_2 \vdash \tau_\Theta(\alpha)\} \end{aligned}$$

Together, these imply that  $D_\Theta D_\Phi T \vdash \alpha$  iff  $T \vdash \tau_\Phi(\tau_\Theta(\alpha))$ .  $\square$

**Lemma 21.** Suppose  $T + d_\Phi \vdash d_\Theta$ . Then

---

|     |                                      |
|-----|--------------------------------------|
| (1) | $D_\Phi T + d_\Theta \vdash d_\Phi.$ |
| (2) | $D_\Theta D_\Phi T \equiv T.$        |

---

*Proof.* For (1), suppose  $T + d_\Phi \vdash d_\Theta$ . By Lemma 18(2), for any  $\alpha \in L_Q$ ,

$$D_\Phi T + d_\Theta \vdash \alpha \leftrightarrow \tau_\Phi(\alpha)$$

Thus,

$$D_\Phi T + d_\Theta \vdash Q_i(\bar{x}) \leftrightarrow \tau_\Phi(Q(\bar{x}))$$

Thus,

$$D_\Phi T + d_\Theta \vdash Q_i(\bar{x}) \leftrightarrow \phi_i$$

as required.

For (2), suppose  $T + d_\Phi \vdash d_\Theta$ . We want to show that, for any  $\alpha \in L_P$ , we have:  $D_\Theta D_\Phi T \vdash \alpha$  iff  $T \vdash \alpha$ . By Lemma 20, we have

$$D_\Theta D_\Phi T \vdash \alpha \text{ iff } T \vdash \tau_\Phi(\tau_\Theta(\alpha))$$

Since  $T + d_{\Phi} \vdash d_{\Theta}$ , by Lemma 10(8), we have

$$T \vdash \alpha \leftrightarrow \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

Hence,

$$T \vdash \alpha \text{ iff } T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

And therefore,

$$D_{\Theta}D_{\Phi}T \vdash \alpha \text{ iff } T \vdash \alpha$$

as required.  $\square$

**Lemma 22.** *The following are equivalent:*

---

|     |   |
|-----|---|
| (1) | $T \equiv D_{\Theta}D_{\Phi}T.$   |
| (2) | $T \vdash \alpha \text{ iff } T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha)), \text{ with } \alpha \in L_P.$ |

---

*Proof.* (1)  $\Rightarrow$  (2). From the previous lemma,  $D_{\Theta}D_{\Phi}T \vdash \alpha$  iff  $T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$ . So, if  $T \equiv D_{\Theta}D_{\Phi}T$ , we infer that  $T \vdash \alpha$  iff  $T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$ , for  $\alpha \in L_P$ .

(2)  $\Rightarrow$  (1). Suppose, for  $\alpha \in L_P$ ,

$$T \vdash \alpha \text{ iff } T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

We already have

$$D_{\Theta}D_{\Phi}T \vdash \alpha \text{ iff } T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

So, for  $\alpha \in L_P$

$$D_{\Theta}D_{\Phi}T \vdash \alpha \text{ iff } T \vdash \alpha$$

Thus,  $D_{\Theta}D_{\Phi}T \equiv T$ .  $\square$

Notice that, in the case of structures, we have:

$$D_{\Theta}D_{\Phi}A = A \text{ then } A \models \alpha \leftrightarrow \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

However, unlike that case, for *theories*, we only have established:

$$\text{If } T \equiv D_{\Theta}D_{\Phi}T \text{ then } T \vdash \alpha \text{ iff } T \vdash \tau_{\Phi}(\tau_{\Theta}(\alpha))$$

So, we have not established the provability of the biconditional  $\alpha \leftrightarrow \tau_{\Phi}(\tau_{\Theta}(\alpha))$  inside  $T$  itself.<sup>11</sup>

---

<sup>11</sup>That said, I do not have a counterexample. It is conceivable that the example given in Andr eka et al. [1] might yield such a counter-example.

## 7. Definitional Equivalence

We next explain what it means for structures and theories to be definitionally equivalent.<sup>12</sup> Intuitively, a pair of structures are definitionally equivalent when they have a common definitional expansion. And a pair of theories are definitionally equivalent when they have a common definitional extension.

Throughout the next two definitions,  $A$  is an  $L_P$ -structure,  $\Phi = \{\phi_i\}_{i \in I_1}$  is a set of  $L_P$ -formulas;  $Q$  is the new disjoint signature corresponding to  $\Phi$ ; and  $B$  is an  $L_Q$ -structure. Similarly,  $\Theta = \{\theta_i\}_{i \in I_2}$  is a set of  $L_Q$ -formulas.  $d_\Phi$  is the definition system of the  $Q_i$  primitives in terms of the  $\phi_i$ , and  $d_\Theta$  is the definition system the  $P_i$  primitives in terms of the  $\theta_i$ . Similarly,  $T_1$  is an  $L_P$ -theory and  $T_2$  is an  $L_Q$ -theory. Please note that the definitions of these notions given below assume *disjoint signatures*.

**Definition 26.** Structures  $A$  and  $B$  are *definitionally equivalent* wrt  $d_\Phi$  and  $d_\Theta$  iff

$$A + d_\Phi \cong B + d_\Theta.$$

If this is so, we write:

$$A \xleftrightarrow[\Theta]{\Phi} B$$

**Definition 27.** Theories  $T_1$  and  $T_2$  are *definitionally equivalent* wrt  $d_\Phi$  and  $d_\Theta$  iff

$$T_1 + d_\Phi \equiv T_2 + d_\Theta.$$

To express this, we write:

$$T_1 \xleftrightarrow[\Theta]{\Phi} T_2$$

These definitions require that the signatures of  $A$  and  $B$  (or  $T_1$  and  $T_2$ ) be disjoint. But is not  $A$  obviously definitionally equivalent to itself? Is not a theory definitionally equivalent to itself? Well, one can always arrange for a pair of structures  $A$  and  $B$  in overlapping signatures to be reformulated as *copies*  $A^c$  and  $B^c$  in entirely disjoint signatures (see Definitions 2, 6 above). If  $A^c$  is a disjoint copy of  $A$  in  $L_{P^c}$ , then clearly  $A^c$  is definitionally equivalent to  $A$  with respect to the trivial definition systems: That is,

---

|               |   |
|---------------|---|
| $d_P^{P^c} :$ | $\forall \bar{x}(P_i^c(\bar{x}) \leftrightarrow P_i(\bar{x}))$  |
| $d_{P^c}^P :$ | $\forall \bar{x}(P_i(\bar{x}) \leftrightarrow P_i^c(\bar{x})).$ |

---

$$A + d_P^{P^c} \cong A^c + d_{P^c}^P$$

A similar copying procedure can be adopted for theories too. If we have a theory  $T$  in  $L_P$ , and  $T^c$  is a disjoint copy of  $T$  in  $L_{P^c}$ , then clearly  $T^c$  is definitionally equivalent to  $T$  with respect to the definitions:

$$T + d_P^{P^c} \equiv T^c + d_{P^c}^P$$

---

<sup>12</sup>The concept appears to have first articulated by de Bouvère [4] and Kanger [11].

So, one can give a more general definition of definitional equivalence by first applying this copying procedure to both structures, and then applying the definition above.<sup>13</sup>

### 8. Main Results

We finally give two main results, which characterize definitional equivalence.

**Theorem 1.** *The following are equivalent:*

---

|     |  |
|-----|--|
| (1) | $A \xleftarrow[\ominus]{\Phi} B$               |
| (2) | $D_{\Phi}A \cong B$ and $D_{\ominus}B \cong A$ |

---

*Proof.* (1)  $\Rightarrow$  (2). Let  $A \xleftarrow[\ominus]{\Phi} B$ . Thus,  $A + d_{\Phi} \cong B + d_{\ominus}$ . So,

$$\begin{aligned} A + d_{\Phi} &\models d_{\ominus} \\ B + d_{\ominus} &\models d_{\Phi} \end{aligned}$$

Thus, by Lemma 16(3) (switching labels in the second case), we have:

$$\begin{aligned} A + d_{\Phi} &\cong D_{\Phi}A + d_{\ominus} \\ B + d_{\ominus} &\cong D_{\ominus}B + d_{\Phi} \end{aligned}$$

Thus,

$$\begin{aligned} B + d_{\ominus} &\cong D_{\Phi}A + d_{\ominus} \\ A + d_{\Phi} &\cong D_{\ominus}B + d_{\Phi} \end{aligned}$$

And by right cancellation,

$$\begin{aligned} B &\cong D_{\Phi}A \\ A &\cong D_{\ominus}B \end{aligned}$$

(2)  $\Rightarrow$  (1). Let  $D_{\Phi}A \cong B$  and  $D_{\ominus}B \cong A$ . Since we have two isomorphisms, we can arrange, without loss of generality, for  $\text{dom}(D_{\Phi}A) = \text{dom}(B)$ , and  $\text{dom}(D_{\ominus}B) = \text{dom}(A)$ . We want to show First, we work out the extensions

---

|     |  |
|-----|--|
| (a) | $(P_i)^{A+d_{\Phi}} = (P_i)^{B+d_{\ominus}}$ |
| (b) | $(Q_i)^{A+d_{\Phi}} = (Q_i)^{B+d_{\ominus}}$ |

---

of  $P_i$  and  $Q_i$  in  $A + d_{\Phi}$ . For  $Q_i$ , we have  $(Q_i)^{A+d_{\Phi}} = \phi_i(A)$ . Using the assumption  $D_{\ominus}B = A$ , we have  $(P_i)^{A+d_{\Phi}} = (P_i)^{D_{\ominus}B+d_{\Phi}}$ . So,  $(P_i)^{A+d_{\Phi}} = \theta_i(B)$ . And using the assumption  $D_{\Phi}A = B$ , we have  $(P_i)^{A+d_{\Phi}} = \theta_i(D_{\Phi}A)$ . Summarizing:

Next working out the extensions in  $B + d_{\ominus}$ , and using the assumption  $D_{\Phi}A = B$ , we have  $(P_i)^{B+d_{\ominus}} = (P_i)^{D_{\Phi}A+d_{\ominus}}$  and  $(Q_i)^{B+d_{\ominus}} = (Q_i)^{D_{\Phi}A+d_{\ominus}}$ . I.e., And (i) and (iii) imply (a), while (ii) and (iv) imply (b). □

---

<sup>13</sup>See Andr eka and N emeti [2] for a related but slightly different procedure.

|       |  |
|-------|--|
| (i)   | $(P_i)^{A+d_\Phi} = \theta_i(D_\Phi A).$   |
| (ii)  | $(Q_i)^{A+d_\Phi} = \phi_i(A).$            |
| (iii) | $(P_i)^{B+d_\Theta} = \theta_i(D_\Phi A).$ |
| (iv)  | $(Q_i)^{B+d_\Theta} = \phi_i(A).$          |

**Theorem 2.** *The following are equivalent:*

|     |  |
|-----|--|
| (1) | $T_1 \xleftrightarrow[\Theta]{\Phi} T_2.$                        |
| (2) | $(\tau_\Phi \tau_\Theta = 1)_{T_1}$ and $T_2 \equiv D_\Phi T_1.$ |

*Proof.* For (1)  $\Rightarrow$  (2), suppose  $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$ . Thus,  $T_1 + d_\Phi \equiv T_2 + d_\Theta$ . Thus,  $T_1 + d_\Phi \vdash d_\Theta$ .

Then, from Lemma 10(8), we have  $T_1 \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$ , for  $\alpha \in L_P$ . I.e.,  $(\tau_\Phi \tau_\Theta = 1)_{T_1}$ . And secondly, since  $T_1 + d_\Phi \vdash d_\Theta$ , we have  $T_1 + d_\Phi \equiv D_\Phi T_1 + d_\Theta$ , by Lemma 19. So,  $T_2 + d_\Theta \equiv D_\Phi T_1 + d_\Theta$ , and by right cancellation,  $T_2 \equiv D_\Phi T_1$ .

For (2)  $\Rightarrow$  (1), suppose  $(\tau_\Phi \tau_\Theta = 1)_{T_1}$  and  $T_2 \equiv D_\Phi T_1$ . From Lemma 11(1), we may conclude that  $T_1 + d_\Phi \vdash d_\Theta$ , and from this, we may conclude that  $T_1 + d_\Phi \equiv D_\Phi T_1 + d_\Theta$ , via Lemma 19. But since  $T_2 \equiv D_\Phi T_1$ , we may conclude that  $T_1 + d_\Phi \equiv T_2 + d_\Theta$ , as required.  $\square$

Indeed, from  $T_1 + d_\Phi \equiv T_2 + d_\Theta$ , we may also conclude  $(\tau_\Phi \tau_\Theta = 1)_{T_1}$ ,  $(\tau_\Theta \tau_\Phi = 1)_{T_2}$ ,  $T_2 \equiv D_\Phi T_1$  and  $T_1 \equiv D_\Theta T_2$ . For example, if  $T_2 + d_\Theta \vdash d_\Phi$ , from Lemma 10(8), by relabelling, we have  $T_2 \vdash \beta \leftrightarrow \tau_\Theta(\tau_\Phi(\beta))$ , for  $\beta \in L_Q$ , and thus  $(\tau_\Theta \tau_\Phi = 1)_{T_2}$ .

## Acknowledgements

The author acknowledges support from a grant from the National Science Centre in Kraków (NCN), Grant Number 2018/29/B/HS1/01832.

**Open Access.** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from

the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- [1] Andr eka, H., Madar asz, J.X., N emeti, I.: Mutual definability does not imply definitional equivalence, a simple example. *Math. Log. Q.* **51**(6), 591–597 (2005)
- [2] Andr eka, H., N emeti, I.: Definability Theory Course Notes (2014). <https://old.renyi.hu/pub/algebraic-logic/DefThNotes0828.pdf>
- [3] Boolos, G., Jeffrey, R.C.: *Computability and Logic*, 3rd edn. Cambridge University Press, Cambridge (1989)
- [4] de Bouv ere, K.L.: Logical synonymy. *Indagat. Math.* **27**, 622–9 (1965)
- [5] Enderton, H.: *A Mathematical Introduction to Logic*, 2nd edn. Harcourt/Academic Press, San Diego (2001)
- [6] Friedman, H., Visser, A.: When Bi-Interpretability Implies Synonymy. Preprint 2014
- [7] Ganea, M.: Arithmetic on semigroups. *J. Symb. Log.* **74**(1), 265–278 (2009)
- [8] Grzegorzczak, A.: Undecidability without arithmetization. *Stud. Log.* **79**(2), 163–230 (2005)
- [9] Grzegorzczak, A., Zdanowski, K.: Undecidability and concatenation. In: Ehrenfeucht, A., Marek, V.W., Srebrny, M. (eds.) *Andrzej Mostowski and Foundational Studies*, pp. 72–91. IOS Press, Amsterdam (2008)
- [10] Hodges, W.: *A Shorter Model Theory*. Cambridge University Press, Cambridge (1997)
- [11] Kanger, S.: Equivalent theories. *Theoria* **34**(1), 1–6 (1968)
- [12] Ketland, J., Schindler, T.: Arithmetic with fusions. *Log. Anal.* **234**, 207–226 (2016)
- [13] Kaye, R., Wong, T.L.: On interpretations of arithmetic and set theory. *Notre Dame J. Formal Log.* **48**(4), 497–510 (2007)
- [14] Machover, M.: *Set Theory, Logic & Their Limitations*. CUP, Cambridge (1996)
- [15] Quine, W.V.: Concatenation as a basis of arithmetic. *J. Symb. Log.* **11**(4), 105–114 (1946)
- [16] Švejdar, V.: On interpretability in the theory of concatenation. *Notre Dame J. Formal Log.* **50**(1), 87–95 (2009)
- [17] Tarski, A., Mostowski, A., Robinson, R.M.: *Undecidable Theories*. North-Holland, Amsterdam (1953)
- [18] Visser, A.: Growing commas: a study of sequentiality and concatenation. *Notre Dame J. Formal Log.* **50**(1), 61–85 (2009)
- [19] Visser, A.: The Interpretability of Inconsistency: Feferman's Theorem and Related Results. Preprint 2014



Bases for Structures and Theories I

Jeffrey Ketland  
Institute of Philosophy  
University of Warsaw  
Krakowskie Przedmieście 3  
00-927 Warsaw  
Poland  
e-mail: [jeffreyketland@gmail.com](mailto:jeffreyketland@gmail.com)

Received: May 26, 2020.

Accepted: July 7, 2020.