THE REVIEW OF SYMBOLIC LOGIC
Volume 13, Number 3, September 2020

DISTANCES BETWEEN FORMAL THEORIES

MOHAMED KHALED
Faculty of Engineering and Natural Sciences, Bahçeşehir University
GERGELY SZÉKELY
Alfréd Rényi Institute of Mathematics, and Department of Natural Sciences, National University of Public Service
KOEN LEFEVER
Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel
and
MICHÈLE FRIEND
George Washington University

Abstract. In the literature, there have been several methods and definitions for working out whether two theories are "equivalent" (essentially the same) or not. In this article, we do something subtler. We provide a means to measure distances (and explore connections) between formal theories. We introduce two natural notions for such distances. The first one is that of *axiomatic distance*, but we argue that it might be of limited interest. The more interesting and widely applicable notion is that of *conceptual distance* which measures the minimum number of concepts that distinguish two theories. For instance, we use conceptual distance to show that relativistic and classical kinematics are distinguished by one concept only.

§1. Introduction. One very important topic in the philosophy of science is how different scientific theories can be compared to each other, especially in the case of competing theories. The first criterion for theory comparison is empirical adequacy, but this is not, at all times and in all circumstances, simple and straightforward (see, e.g., [12]). So, when we have competing theories, each empirically adequate, we look elsewhere to make sense of the present state of science. For this reason, investigating the relations between theories, independent of their relation to reality, becomes also very important.

So far this investigation has been made mainly in one direction: whether two given theories have the same essential content. There have been several attempts to capture the concept of equivalence between theories (henceforth: "theory-equivalence"), e.g., *logical equivalence*, *definitional equivalence*, and *categorical equivalence*. When we say that two theories are not equivalent, we mean that there is at least one difference between them and that this difference can be formulated from the point of view from which one decides to explore the equivalence between the theories in question.

Received: July 30, 2018.
2010 *Mathematics Subject Classification*: Primary 03B99, 03C07, 03A10, Secondary 03G99, 03B80.
Key words and phrases: network of theories, degrees of nonequivalence, conceptual distance, relativistic and classical kinematics.

doi:10.1017/S1755020319000558

In the present article, we propose a new direction in the analysis of the connections between formal theories. Our proposal aims to provide a *qualitative* and *quantitative* study of the differences between theories. In order to investigate how far two theories are from each other, we introduce some notions for the measure of distances between theories, such distances count the minimum number of differences distinguishing two given theories. We do this by measuring the degree of nonequivalence of two nonequivalent theories (according to any chosen definition of theory-equivalence).

We focus on formal theories that are formulated in any of the following logical systems: sentential logic, ordinary first-order logic (FOL), finite variables fragments of FOL and/or infinitary versions of FOL. We develop, discuss and compare some notions for distances between theories.

The idea is very simple: based on a symmetric relation capturing a notion of minimal change, we introduce a general way to define a distance on any class of objects (not just theories) equipped with an equivalence relation. The idea is a generalization of the distance between any two nodes in the same graph, in graph theory. Later, we give particular examples when the given class is a class of theories and the equivalence relation is a fixed notion of theory-equivalence.

The first particular example is that of logical equivalence. As a measure for the degree of logical nonequivalence, we introduce the concept of *axiomatic distance*. The idea is to count the minimum number of axioms that are needed to be added or "removed" to get from one theory to the other.[1] We prove that the axiomatic distance between theories formulated in the same language must be ≤ 3. See Proposition 3.11 on Page 642. Although, counting axioms separating theories seems to be a natural suggestion for defining the desired metric, this result gives the sense that the measure of this axiomatic distance is of limited use.

Then we turn to definitional equivalence. Two theories are definitionally equivalent if they cannot be distinguished by a concept (a formula defining some notion). As a measure for the degree of definitional nonequivalence, we define *conceptual distance*. This distance counts the minimum number of (nondefinable) *concepts* that separate two theories. We find that this distance is of special interest in the study of logic. We give examples and we count conceptual distance between some specific theories, see, e.g., Theorem 4.3. We also explore a connection between conceptual distance and *spectrum of theories* which is a central topic in model theory, cf., Theorem 4.9.

Such quantitative study might be useful, and it may provide new insights in comparing formal theories. Given a metric on a class of formal theories formulated in a fixed logical system, some classifications for the theories in this class can be achieved. For example, a classification can be achieved by measuring the axiomatic distance from the empty theory, see Theorem 3.13 herein. For a classification using conceptual distance, we refer the reader to Theorem 4.9 on page 646. The task now is to find the relationship between the properties of theories and the categories of these classifications, and also to set boundary lines between these categories. So, when a new theory is constructed, its properties can be estimated, given its category in a classification.

With definitions and metrics on distance developed here, we have maps of the network of logical theories. When we draw such maps of networks, the topology may suggest very interesting and fruitful questions. For instance, if there is a distance other than zero or

[1] By "removing an axiom" here we only mean the trivial converse of adding an axiom in the following sense: T is a theory resulting from "removing" one axiom from T' if T' can be reached from T by adding one axiom.

one, then is there already a known theory in between? Or if not, we can ask what are the limitative properties of that theory and what is its philosophical significance? By engaging such questions, we see the "edge" of the limitative results, and by examining this edge, we understand with precision the relationship between meta-logical limitative results and physical phenomena.

Furthermore, we investigate the possible application of conceptual distance in the logical foundation of physical theories in ordinary first-order logic. We prove that conceptual distance between classical and relativistic kinematics is equal to one. In other words, only one concept distinguishes classical and relativistic kinematics: the existence of a class of observers who are at absolute rest. This is indeed an interesting result in its own right, not only for logicians but also for physicists. Such a result opens several similar questions about how many concepts differ two theories in physics whose phenomena can be described in FOL and also questions about the precise nature of the differentiating concepts.

In the philosophy of physics, this might be important because, on the one hand, it is clear that we are not presently converging towards one unified theory of physics in the sense of converging to one set of laws from which all the phenomena of physics can be derived. On the other hand, we can give a logical foundation for several physical theories: Newtonian mechanics, relativity theories and some parts of quantum theory. Given these logical representations, we would like to know the exact logical and conceptual relationship between physical theories. If we have a complete overview of this, then we can form an impression of how far we are from such a philosophical dream—the dream of the unity of physics. Or, we can adjust our hopes and expectations, and rest content with a unity of science at a more general level: as a network of logical theories with precise relations between them. For some philosophers, this is a radical reconception as to what "the unity of physics" consists in. It is worth exploring this conception, since it more accurately mirrors the actual state of our various physical theories and their relations to each other.

1.1. The algebraic idea behind conceptual distance. By a concept, we understand a definable notion, no matter how many different ways one can define it. In other words, a concept is a maximal set of logically equivalent formulas. Our understanding here comes from the theory of cylindric algebras. These algebras were defined by A. Tarski around 1947 to capture the intrinsic algebraic side of FOL. Cylindric algebras are often introduced as algebras of different concepts of the corresponding theories, see, e.g., [19, sec. 4.3] and [6].

Now, we want to define a distance counting the minimum number of concepts that distinguish two theories T and T'. It is very natural to explain the idea within the framework of cylindric algebras, since these are concept algebras and we want to count concepts. Assume that $\mathfrak{A}$ and $\mathfrak{A}'$ are the cylindric algebras corresponding to theories T and T', respectively.

If $\mathfrak{A}$ is isomorphic to $\mathfrak{A}'$, then the two theories are definitionally equivalent [19, Theorem 4.3.43], and so the conceptual distance between them is zero. Now, assume that $\mathfrak{A}$ and $\mathfrak{A}'$ are not isomorphic. For simplicity, let us assume that $\mathfrak{A}$ is embeddable into $\mathfrak{A}'$. Thus, the minimum number of concepts distinguishing the two theories is equal to the minimum number of elements of $\mathfrak{A}'$ that we can add to $\mathfrak{A}$ (more precisely, to one of its copies inside $\mathfrak{A}'$) to generate the algebra $\mathfrak{A}'$, see Figure 1.

In the case of Figure 1, we can say that we need one step to move from T to T'. Now, the minimal distance between two theories can be defined as the minimum number of steps needed to move from one theory to the other. Our definition of conceptual distance

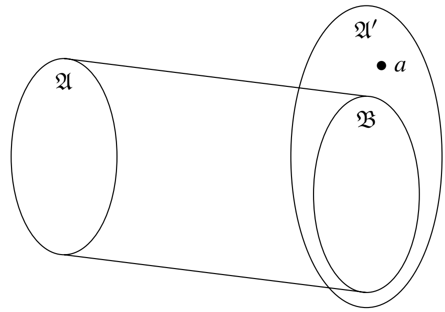

Fig. 1. The distance between $\mathfrak{A}$ and $\mathfrak{A}'$ is one if $\langle B, a\rangle \cong \mathfrak{A}'$ and $\mathfrak{A} \cong \mathfrak{B} \ncong \mathfrak{A}'$.

(Definitions 4.1 and 4.2) illustrates this idea, but in terms of logic instead of algebras. In a future algebra-oriented article, we plan to discuss in detail the correspondence between our logical definition herein and the above algebraic idea. We note that it happens quite often that one obtains interesting results in mathematical logic by using algebraic tools, e.g., [17], [7], [1], [9], and [22].

We start with a quick review for the notions of logic that we are going to use. We assume familiarity with the basic notions of set theory. For instance, what is a set, a class, a relation, etc. The only difference is that in several places in this article, we decided not to distinguish different kinds of infinities. Therefore, together with the standard notion of cardinality, we are going to speak about the *size of set X*, defined as follows:

$$||X|| \stackrel{\text{def}}{=} \begin{cases} k & X \text{ is finite and has exactly } k\text{-many elements,} \\ \infty & \text{if } X \text{ is an infinite set.} \end{cases}$$

We also make use of von Neumann ordinals. For example, ω is the smallest infinite ordinal, sometimes we denote ω by $\mathbb{N}$ to indicate that it is the set of natural numbers (nonnegative integers).

§2. Notions of logic. In the course of this article, let α and $\beta \leq \alpha + 1$ be two fixed ordinals. We consider a natural generalisation of ordinary first-order logic, we denote it by $\mathbf{L}_\alpha^\beta$, which is inspired from the definitions and the discussions in [19, sec. 4.3]. Roughly, the formulas of $\mathbf{L}_\alpha^\beta$ uses a fixed set of individual variables $\{v_i : i \in \alpha\}$ and relation symbols of rank strictly less than β. For simplicity, we assume that our languages do not contain any function symbols and/or constant symbols.

In particular, $\mathbf{L}_0^1$ is sentential (propositional) logic, while $\mathbf{L}_\omega^\omega$ is ordinary first-order logic. The so-called finite variable fragments of first-order logic are the logics $\mathbf{L}_n^{n+1}$, for finite ordinals n. When α and β are infinite, $\mathbf{L}_\alpha^\beta$ is called infinitary logic. Throughout, since α and β are fixed, languages, theories, etc., are understood to be languages for $\mathbf{L}_\alpha^\beta$, theories in $\mathbf{L}_\alpha^\beta$, etc.

2.1. The syntax of $\mathbf{L}_\alpha^\beta$**.** A *language* $\mathcal{L}$ *for* $\mathbf{L}_\alpha^\beta$ is a tuple $(\mathcal{R}, \mathsf{rank})$, where $\mathcal{R}$ is a set of relation symbols and $\mathsf{rank} : \mathcal{R} \to \beta$ is a function assigns for each relation symbol $R \in \mathcal{R}$ a rank $\mathsf{rank}(R)$. Relation symbols P with rank $\mathsf{rank}(P) = 0$ are called *sentential constants*.

From now on, and for simplicity, we will use $\mathcal{L}$ for a language and its set of relation symbols, and the rank of any relation symbol $R \in \mathcal{L}$ will be denoted by $\mathsf{rank}(R)$. A language for $\mathbf{L}_\alpha^\beta$ can be also considered to be a language for $\mathbf{L}_\alpha^{\beta'}$, for any ordinal β' with $\beta \leq \beta' \leq \alpha + 1$.

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

To construct the formulas of a language $\mathcal{L}$, we also need some other symbols: equality "="[2], brackets "(" and ")," conjunction "$\wedge$," negation "$\neg$," and the existential quantifier "$\exists$". We also use the necessary symbols to write sequences of variables $(v_{i_m} : m \in I)$, for any indexing set $I \subseteq \alpha$. We assume that all of these symbols are part of the logic $\mathbf{L}_\alpha^\beta$ itself.

The set of formulas Fm *of* $\mathcal{L}$ is the smallest set that satisfies:

(a) Fm contains each *basic formula of* $\mathcal{L}$, where the basic formulas are the following two types of formulas:

  (i) The equalities $v_i = v_j$, for any $i, j \in \alpha$.
  (ii) $R(v_{i_m} : m < \mathsf{rank}(R))$, for any relation symbol R.

(b) Fm contains $(\varphi \wedge \psi)$, $(\neg\varphi)$ and $(\exists v_i \varphi)$, for each $\varphi, \psi \in \mathsf{Fm}$.

We use the usual conventions for dropping brackets in FOL, e.g., we may drop the outside brackets of a formula. We also use the following abbreviations:

- If P is a sentential constant, then we just write P instead of $P()$.
- If R is a relation symbol of finite positive rank, say k, then we write $R(v_{i_0}, \dots, v_{i_{k-1}})$ instead of $R(v_{i_m} : m < k)$.
- We use disjunction, implication, equivalence and universal quantifier as:

$$\varphi \vee \psi \stackrel{\text{def}}{=} \neg(\neg\varphi \wedge \neg\psi) \qquad \varphi \to \psi \stackrel{\text{def}}{=} \neg(\varphi \wedge \neg\psi)$$
$$\varphi \leftrightarrow \psi \stackrel{\text{def}}{=} (\varphi \to \psi) \wedge (\psi \to \varphi) \qquad \forall v_i\, \varphi \stackrel{\text{def}}{=} \neg(\exists v_i\, \neg\varphi).$$

- We also use grouped conjunction and disjunction: Empty disjunction is defined to be $\varphi \wedge \neg\varphi$ and empty conjunction is defined to be $\varphi \vee \neg\varphi$ (for any arbitrary but fixed formula $\varphi \in \mathsf{Fm}$).[3] Let $\varphi_0, \dots, \varphi_m \in \mathsf{Fm}$, then

$$\bigvee_{0 \le i \le m} \varphi_i \stackrel{\text{def}}{=} \big(\dots(\varphi_0 \vee \varphi_1) \vee \cdots \varphi_m\big) \quad \text{and} \quad \bigwedge_{0 \le i \le m} \varphi_i \stackrel{\text{def}}{=} \big(\dots(\varphi_0 \wedge \varphi_1) \cdots \wedge \varphi_m\big).$$

Let us note that our choice of restricted vocabulary excludes some logics, such as intuitionistic logic, where $\vee$ is not definable from $\wedge$ and $\neg$. Moreover, our design herein does not allow us to recover the logic $\mathbf{L}_\alpha^\beta$ from a given language $\mathcal{L}$, but this is not important for us, since the logic $\mathbf{L}_\alpha^\beta$ is fixed throughout the article.

2.2. The semantics of $\mathbf{L}_\alpha^\beta$***.*** A *model* $\mathfrak{M}$ *for language* $\mathcal{L}$ is a nonempty set M enriched with operations $R^\mathfrak{M} \subseteq M^{\mathsf{rank}(R)}$, for each $R \in \mathcal{L}$ (for a sentential constant P, $P^\mathfrak{M} \subseteq M^0 = \{\emptyset\}$).[4] An *assignment in* $\mathfrak{M}$ is a function τ that assigns for each variable an element of the set M. Let $\varphi \in \mathsf{Fm}$ be any formula. The satisfiability relation $\mathfrak{M}, \tau \models \varphi$ is defined recursively as follows:

---

[2] It is also important to note that the equality symbol is always assumed, even if $\beta \le 2$. Therefore, the set of formulas is not empty unless $\mathcal{L} = \emptyset$ and $\alpha = 0$.

[3] These are nondeterministic definitions; $\exists v_0(v_0 \neq v_0)$ for the empty disjunction and $\exists v_0(v_0 = v_0)$ for the empty conjunction could be better ones, but these deterministic definitions require the assumption $\alpha \ge 1$.

[4] So the meaning $P^\mathfrak{M}$ of a sentential constant P can be either true ($T = \{\emptyset\}$) or false ($F = \emptyset$).

$$\begin{aligned}
\mathfrak{M}, \tau \models R(v_{i_m} : m < \mathsf{rank}(R)) \quad &\text{iff} \quad (\tau(v_{i_m}) : m < \mathsf{rank}(R)) \in R^{\mathfrak{M}},^{5}\\
\mathfrak{M}, \tau \models v_i = v_j \quad &\text{iff} \quad \tau(v_i) = \tau(v_j),\\
\mathfrak{M}, \tau \models \varphi \wedge \psi \quad &\text{iff} \quad \mathfrak{M}, \tau \models \varphi \text{ and } \mathfrak{M}, \tau \models \psi,\\
\mathfrak{M}, \tau \models \neg\varphi \quad &\text{iff} \quad \mathfrak{M}, \tau \not\models \varphi,\\
\mathfrak{M}, \tau \models \exists v_i \varphi \quad &\text{iff} \quad \text{there is } a \in M \text{ such that } \mathfrak{M}, \tau[v_i \mapsto a] \models \varphi,
\end{aligned}$$

where $\tau[v_i \mapsto a]$ is the assignment which agrees with τ on every variable except $\tau[v_i \mapsto a](v_i) = a$. The cardinality of $\mathfrak{M}$ is defined to be the cardinality of M. A formula φ is said to be *true in* $\mathfrak{M}$, in symbols $\mathfrak{M} \models \varphi$, iff $\mathfrak{M}, \tau \models \varphi$, for every assignment τ in $\mathfrak{M}$. A formula φ is said to be a *tautology* iff it is true in every model for $\mathcal{L}$. The *theory of* $\mathfrak{M}$ is defined as

$$\mathsf{Th}(\mathfrak{M}) \stackrel{\text{def}}{=} \{\varphi \in \mathsf{Fm} : \mathfrak{M} \models \varphi\}.$$

We say that *two models* $\mathfrak{M}$ *and* $\mathfrak{N}$ *for language* $\mathcal{L}$ *are isomorphic* iff there is a bijection $f : M \to N$ between their underlying sets that respects the meaning of the relation symbols, i.e., for each $R \in \mathcal{L}$,

$$(a_i : i < \mathsf{rank}(R)) \in R^{\mathfrak{M}} \iff (f(a_i) : i < \mathsf{rank}(R)) \in R^{\mathfrak{N}}.$$

2.3. Theories in the logic $\mathbf{L}_\alpha^\beta$.

DEFINITION 2.1. *A* theory T *is a pair* $(\mathcal{L}, A)$*, where* $\mathcal{L}$ *is a language and* $A \subseteq \mathsf{Fm}$ *is a subset of its set of formulas.*

We use the same superscripts and subscripts for theories and their corresponding languages. For example, if we write T' is a theory, then we understand that T' is a theory of language $\mathcal{L}'$ whose set of formulas is Fm'. For simplicity, we will loosely assume that a theory T is a set of formulas, but we know that a language is given in the background. In this sense, we may assume that the same theory can be given in different languages, or in different logics.

A *model for theory* T is a model for $\mathcal{L}$ in which every $\psi \in T$ is true. We say that theory T is *consistent* iff there is at least one model for T.

DEFINITION 2.2. *Let T be a theory and let κ be any cardinal. The* spectrum of T*, in symbols* $\mathsf{I}(T, \kappa)$*, is the number of its different models (up to isomorphism) of cardinality κ. This number is defined to be ∞ if T has infinitely many nonisomorphic models of cardinality κ.*

We say that *a formula φ is a consequence of theory T*, in symbols $T \models \varphi$, iff φ is true in every model for T. With this definition, one may have: if $T \models \varphi$, then $T \models \forall v_i \varphi$, but one does not have: if $T \cup \{\varphi\} \models \psi$ then $T \models \varphi \to \psi$. Kit Fine [14, p. 65] calls such an approach "truth to truth" rather than "case to case." The *set of consequences of theory T* is defined as follows:

$$\mathsf{Cn}(T) \stackrel{\text{def}}{=} \{\varphi \in \mathsf{Fm} : T \models \varphi\}.$$

DEFINITION 2.3. *Two theories T_1 and T_2 are called* logically equivalent*, in symbols $T_1 \equiv T_2$, iff they have the same consequences, i.e.,* $\mathsf{Cn}(T_1) = \mathsf{Cn}(T_2)$.

[5] If $\mathsf{rank}(P)$ is 0, then $(\tau(v_{i_m}) : m < \mathsf{rank}(P))$ is the empty sequence $\emptyset$. Hence $\mathfrak{M}, \tau \models P$ iff $P^{\mathfrak{M}}$ is true.

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

2.4. More notions for theory-equivalence. A *translation of language $\mathcal{L}_1$ into language $\mathcal{L}_2$* is a map $\mathsf{tr} : \mathsf{Fm}_1 \to \mathsf{Fm}_2$ such that the following are true for every $\varphi, \psi \in \mathsf{Fm}$ and every v_i, v_j.

- The free variables of $\mathsf{tr}(\varphi)$ are among the free variables of φ.
- $\mathsf{tr}(v_i = v_j)$ is $v_i = v_j$.
- tr commutes with the Boolean connectives:

$$\mathsf{tr}(\neg\varphi) = \neg\mathsf{tr}(\varphi) \ \text{ and } \ \mathsf{tr}(\varphi \wedge \psi) = \mathsf{tr}(\varphi) \wedge \mathsf{tr}(\psi).$$

- Finally, $\mathsf{tr}(\exists v_i\, \varphi) = \exists v_i\, \mathsf{tr}(\varphi)$.[6]

DEFINITION 2.4. *Suppose that T_1 and T_2 are theories in languages $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively, and* tr *is a translation of $\mathcal{L}_1$ into $\mathcal{L}_2$. The translation* tr *is said to be an* interpretation of T_1 into T_2 *iff it maps consequences of T_1 into consequences of T_2, i.e., for each formula $\varphi \in \mathsf{Fm}_1$,*

$$T_1 \models \varphi \implies T_2 \models \mathsf{tr}(\varphi).$$

(a) *An interpretation* tr *of T_1 into T_2 is called* a faithful interpretation of T_1 into T_2 *iff for each formula $\varphi \in \mathsf{Fm}_1$,*

$$T_1 \models \varphi \iff T_2 \models \mathsf{tr}(\varphi).$$

(b) *An interpretation* tr_{12} *of T_1 into T_2 is called* a definitional equivalence between T_1 and T_2 *iff there is an interpretation* tr_{21} *of T_2 into T_1 such that*

- $T_1 \models \mathsf{tr}_{21}\big(\mathsf{tr}_{12}(\varphi)\big) \leftrightarrow \varphi$,
- $T_2 \models \mathsf{tr}_{12}\big(\mathsf{tr}_{21}(\psi)\big) \leftrightarrow \psi$.

for every $\varphi \in \mathsf{Fm}_1$ and $\psi \in \mathsf{Fm}_2$. In this case, tr_{21} *is also a definitional equivalence.*

DEFINITION 2.5. *Two theories T_1 and T_2 are said to be* definitionally equivalent, *in symbols $T_1 \rightleftarrows T_2$, iff there is a definitional equivalence between them.*

In the literature, there are several ways to define definitional equivalence, e.g., [3], [10], [21], and [29]. Here, we use a variant of the definition in [19, Definition 4.3.42 and Theorem 4.3.43]. Our notion here for definitional equivalence was shown to be an equivalence relation [26], this will play a role in the following sections. For a discussion on the different definitions of definitional equivalence, see [26], and we refer to [31] for a category theory based discussion.

REMARK 2.6. *Let T_1 and T_2 be two theories and suppose that* $\mathsf{tr}_{12} : \mathsf{Fm}_1 \to \mathsf{Fm}_2$ *is a definitional equivalence between T_1 and T_2, then* tr_{12} *is also a faithful interpretation.*

DEFINITION 2.7. *Let T_1 and T_2 be two theories. We say that T_2* is a conservative extension of T_1, *in symbols $T_1 \sqsubseteq T_2$, iff $\mathsf{Fm}_1 \subseteq \mathsf{Fm}_2$ and, for all $\varphi \in \mathsf{Fm}_1$, $T_2 \models \varphi \iff T_1 \models \varphi$.*

---

[6] In the case of ordinary first-order logic (when $\alpha = \beta = \omega$), to define a translation $\mathsf{tr} : \mathsf{Fm}_1 \to \mathsf{Fm}_2$, it suffices to define tr on the basic formulas in Fm_1 of the form $R(v_0, \ldots, v_{m-1})$. Then, using Tarski's substitution observation, we can define

$$\mathsf{tr}(R(v_{i_0}, \ldots, v_{i_{m-1}})) = \exists y_0(y_0 = v_{i_0} \wedge \cdots \wedge \exists y_{m-1}(y_{m-1} = v_{i_{m-1}} \wedge \\ \exists v_0(v_0 = y_0 \wedge \cdots \wedge \exists v_{m-1}(v_{m-1} = y_{m-1} \wedge \mathsf{tr}(R(v_0, \ldots, v_{m-1})))))),$$

where $y_i = v_{l+i}$ and l is the maximum of $0, \ldots, m-1, i_0, \ldots, i_{m-1}$. This can be extended in a unique way to a translation that covers the whole Fm_1.

We note that $T_1 \sqsubseteq T_2$ iff the identity translation $\mathsf{id} : \mathsf{Fm}_1 \to \mathsf{Fm}_2$ is a faithful interpretation. It is also worth mentioning that $T_1 \sqsubseteq T_2 \iff T_1 \equiv \mathsf{Cn}(T_2) \cap \mathsf{Fm}_1$.

§3. Cluster networks & step distance. Now, we introduce a general way of defining a distance on any given class X. We note that our target is to define distances on the class of all theories, thus we need to work with classes which are not necessarily sets.[7]

DEFINITION 3.1. *By a* cluster (X, E) *we mean a nonempty class X equipped with an equivalence relation E.*

We are interested in distances according to which some different objects are indistinguishable. Indeed, it is natural to treat equivalent theories as if they were of distance 0 from each other. As we mentioned in the introduction, there are several notions of equivalence between theories. Such equivalence thus can be represented in the cluster of theories by the relation E.

DEFINITION 3.2. *A* cluster network *is a triple (X, E, S), where (X, E) is a cluster and S is a symmetric relation on X.*

Given a cluster network (X, E, S). A *path leading from $x \in X$ to $x' \in X$ in (X, E, S)* is a finite sequence $b_1, \dots, b_m$ of 0's and 1's such that there is a sequence $x_0, \dots, x_m$ of members of X with $x_0 = x$, $x_m = x'$ and, for each $1 \leq i \leq m$,

$$b_i = 0 \implies x_{i-1}\, E\, x_i \quad \text{and} \quad b_i = 1 \implies x_{i-1}\, S\, x_i.$$

The *length of this path* is defined to be $\sum_{i=1}^{m} b_i$. Two objects $x, x' \in X$ *are connected in* (X, E, S) iff there is a path leading from one of them to the other in (X, E, S).

DEFINITION 3.3. *Let $\mathcal{X} = (X, E, S)$ be a cluster network. The* step distance *on $\mathcal{X}$ is the function $d_{\mathcal{X}} : X \times X \to \mathbb{N} \cup \{\infty\}$ defined as follows. For each $x, x' \in X$:*

- *If x and x' are not connected in (X, E, S), then $d_{\mathcal{X}}(x, x') \stackrel{\text{def}}{=} \infty$.*
- *If x and x' are connected in (X, E, S), then*

$$d_{\mathcal{X}}(x, x') \stackrel{\text{def}}{=} \min\{k \in \mathbb{N} : \exists \text{ a path leading from } x \text{ to } x' \text{ whose length is } k\}.$$

The equivalence relation E represents pairs that cannot be distinguished by the step distance, whereas the symmetric relation S represents the pairs of objects that are (at most) one step away from each other. The step distance then counts the minimum number of steps needed to reach an object starting from another one.

EXAMPLE 3.4. *Let X be any class, let E be the identity relation on X, and let $S = X \times X$. Then, $\mathcal{X} = (X, E, S)$ is a cluster network and its step distance is the following discrete distance:*

$$d_{\mathcal{X}}(x, x') = \begin{cases} 0 & \text{if } x = x', \\ 1 & \text{if } x \neq x'. \end{cases}$$

THEOREM 3.5. *Let $\mathcal{X} = (X, E, S)$ be a cluster network and let $d_{\mathcal{X}} : X \times X \to \mathbb{N} \cup \{\infty\}$ be the step distance on $\mathcal{X}$. The following are true for each $x_1, x_2, x_3 \in X$:*

(a) $d_{\mathcal{X}}(x_1, x_2) \geq 0$, *and* $d_{\mathcal{X}}(x_1, x_2) = 0 \iff x_1\, E\, x_2$.

[7] All definitions in this section can be formulated within von Neumann–Bernays–Gödel set theory.

(b) $d_{\mathcal{X}}(x_1, x_2) = d_{\mathcal{X}}(x_2, x_1)$.

(c) $d_{\mathcal{X}}(x_1, x_2) \leq d_{\mathcal{X}}(x_1, x_3) + d_{\mathcal{X}}(x_3, x_2)$, *(where addition with ∞ is defined in the natural way).*

(d) *If $d(x_1, x_2) = n$ and $m + k = n$, then there is a $y \in X$ such that $d(x_1, y) = m$ and $d(y, x_2) = k$.*

Proof. Straightforward. □

REMARK 3.6. *Let $\mathcal{X} = (X, E, S)$ and $\mathcal{X}' = (X', E', S')$ be cluster networks such that $X \subseteq X'$, $E \subseteq E'$, and $S \subseteq S'$. Since every path in $\mathcal{X}$ is contained in $\mathcal{X}'$, it is easy to see that $d_{\mathcal{X}}(x_1, x_2) \geq d_{\mathcal{X}'}(x_1, x_2)$ for each $x_1, x_2 \in X$.*

Now, we use the above general settings to define distances between theories. Before we start, we need the following convention: suppose that we are given two theories T and T' of the same language. We write $T \leftharpoondown T'$ iff there is $\varphi \in \mathsf{Fm}$ such that $T \cup \{\varphi\} \equiv T'$. We also write $T \rightleftharpoons T'$ iff either $T \leftharpoondown T'$ or $T' \leftharpoondown T$. Conventionally, we call the relation $\leftharpoondown$ *axiom adding*, whereas the converse relation $\rightharpoonup$ is called *axiom removal*. It is easy to see that the following are true for any theories T_1, T_2 and T_3:

$$\begin{aligned} T_1 \leftharpoondown T_2 \;\&\; T_2 \leftharpoondown T_3 &\implies T_1 \leftharpoondown T_3, \\ T_1 \equiv T_2 \;\&\; T_2 \leftharpoondown T_3 &\implies T_1 \leftharpoondown T_3, \\ T_1 \leftharpoondown T_2 \;\&\; T_2 \equiv T_3 &\implies T_1 \leftharpoondown T_3. \end{aligned}$$

DEFINITION 3.7. *Let $\mathcal{T}$ be a class of some theories (in the logic $\mathbf{L}_\alpha^\beta$) and consider the cluster network $(\mathcal{T}, \equiv, \rightleftharpoons)$. We call the step distance on this cluster network* axiomatic distance on $\mathcal{T}$. *This step distance will be denoted by $\mathsf{Ad}_{\mathcal{T}}$.*

Let $\mathcal{T}$ be a class of theories. We note the following: If there is a path between $T, T' \in \mathcal{T}$ in the cluster network $(\mathcal{T}, \equiv, \rightleftharpoons)$, then both T and T' must be formulated in the same language. In other words, if T, T' are formulated on different languages, then $\mathsf{Ad}_{\mathcal{T}}(T, T') = \infty$. This is because, if $T \equiv T'$ or $T \rightleftharpoons T'$, then $\mathcal{L} = \mathcal{L}'$.

EXAMPLE 3.8. *Suppose that $\alpha \geq 1$ or $\beta \geq 1$. Let $\mathcal{T}$ be a class of theories. Let $T, T_\perp \in \mathcal{T}$ be two theories formulated in the same language. Suppose that T is consistent while $T_\perp$ is inconsistent. Then, adding a contradiction to T ensures that $\mathsf{Ad}_{\mathcal{T}}(T, T_\perp) = 1$. Consequently, if $T, T' \in \mathcal{T}$ are formulated on the same language and an inconsistent theory $\mathcal{T}_\perp$ of that language is in $\mathcal{T}$, then $\mathsf{Ad}_{\mathcal{T}}(T, T') \leq 2$ since we have $T \rightharpoonup \mathcal{T}_\perp \leftharpoondown T'$.*

EXAMPLE 3.9. *Let $\mathcal{T}$ be a class of theories. Let $T, \emptyset_{\mathcal{L}} \in \mathcal{T}$ be two theories formulated in the same language $\mathcal{L}$ such that $\emptyset_{\mathcal{L}}$ is the empty theory of $\mathcal{L}$ (i.e., empty set of formulas). Suppose that T is finitely axiomatizable, then we have either*

$$\mathsf{Ad}_{\mathcal{T}}(T, \emptyset_{\mathcal{L}}) = 1 \;\; or \;\; T \equiv \emptyset_{\mathcal{L}}.$$

Thus, in the class of all theories, the axiomatic distance between any two finitely axiomatizable theories of the same language is ≤ 2.

EXAMPLE 3.10. *Suppose that $\alpha \geq 3$ and $\beta \geq 3$. Let $\mathcal{T}$ be the set of all consistent theories of binary relations, let T_P be the theory of strict partial orders, and let T_E be the theory of equivalence relations. Then $\mathsf{Ad}_{\mathcal{T}}(T_P, T_E) = 2$. Clearly, $\mathsf{Ad}_{\mathcal{T}}(T_P, T_E) \geq 2$ because none of T_P or T_E implies the other, and, by Example 3.9 and Theorem 3.5 (c), $\mathsf{Ad}_{\mathcal{T}}(T_P, T_E) \leq \mathsf{Ad}_{\mathcal{T}}(T_P, \emptyset) + \mathsf{Ad}_{\mathcal{T}}(\emptyset, T_E) = 2$.*

All these examples suggest that the axiomatic distance, in most of the cases, has restricted measures. Let $\mathsf{CON}_\alpha^\beta$ be the class of all consistent theories in $\mathbf{L}_\alpha^\beta$. For simplicity, we will denote the axiomatic distance in the class $\mathsf{CON}_\alpha^\beta$ by Ad instead of $\mathsf{Ad}_{\mathsf{CON}_\alpha^\beta}$.

PROPOSITION 3.11. *Let $T, T' \in \mathsf{CON}_\alpha^\beta$ be two theories with the same language $\mathcal{L}$. Then the axiomatic distance $\mathsf{Ad}(T, T') \leq 3$. Moreover, if $T \not\equiv T'$ and they are complete, then $\mathsf{Ad}(T, T') = 2$.*

Proof. Let $T, T' \in \mathsf{CON}_\alpha^\beta$ be two theories with the same language $\mathcal{L}$. If $T \equiv T'$, then $\mathsf{Ad}(T, T') = 0$. Otherwise, and without loss of generality, we may assume that there is a formula φ such that $T \not\models \varphi$ and $T' \models \varphi$. We note that, since $T \not\models \varphi$, the theory $T \cup \{\neg\varphi\}$ is consistent. We define a theory W on the language $\mathcal{L}$ as follows:

$$W \stackrel{\text{def}}{=} \{\neg\varphi \to \psi : \psi \in T\} \cup \{\varphi \to \chi : \chi \in T'\}.$$

Theory $W \cup \{\neg\varphi\}$ is consistent since $T \cup \{\neg\varphi\}$ is so. Since $T' \models \varphi$, we find that $W \cup \{\varphi\} \equiv T'$. We now have:

$$T \rightleftharpoons T \cup \{\neg\varphi\} \equiv W \cup \{\neg\varphi\} \rightleftharpoons W \rightleftharpoons W \cup \{\varphi\} \equiv T'.$$

Therefore, the axiomatic distance between T and T' in $\mathsf{CON}_\alpha^\beta$ is at most 3. Now, suppose that theories T and T' are complete and $T \not\equiv T'$, then their axiomatic distance cannot be ≤ 1. Let φ be as in our previous argument. By the completeness of T, it follows that $T \models \neg\varphi$. Consequently,

$$T \equiv W \cup \{\neg\varphi\} \rightleftharpoons W \rightleftharpoons W \cup \{\varphi\} \equiv T',$$

which implies that the axiomatic distance between T and T' in $\mathsf{CON}_\alpha^\beta$ is precisely 2. □

COROLLARY 3.12. *Let $T, T' \in \mathsf{CON}_\alpha^\beta$. Then,*

$$\mathsf{Ad}(T, T') = \infty \iff T \text{ and } T' \text{ are formulated in different langauges}.$$

THEOREM 3.13. *Fix a language $\mathcal{L}$ for the logic $\mathbf{L}_\alpha^\beta$. Let $T \in \mathsf{CON}_\alpha^\beta$ be any theory on language $\mathcal{L}$ and let $\emptyset_\mathcal{L}$ be the empty theory on $\mathcal{L}$. The following are true*:

1. $\mathsf{Ad}(\emptyset_\mathcal{L}, T) = 0$ *iff T is trivial ($T \equiv \emptyset_\mathcal{L}$).*
2. $\mathsf{Ad}(\emptyset_\mathcal{L}, T) = 1$ *iff T is finitely axiomatizable and nontrivial.*
3. $\mathsf{Ad}(\emptyset_\mathcal{L}, T) = 2$ *iff T is not finitely axiomatizable, but has a finitely axiomatizable consistent extension.*
4. $\mathsf{Ad}(\emptyset_\mathcal{L}, T) = 3$ *iff T has no finitely axiomatizable consistent extension.*

Proof. The proof of this is easy (given that the last item is the remaining case of the other items by the result in Proposition 3.11). □

EXAMPLE 3.14. *Let PA stand for Peano Arithmetic with its usual axioms in the relation language for arithmetic $\mathcal{L}$. Then $\mathsf{Ad}(\emptyset_\mathcal{L}, \mathsf{PA}) = 3$, because PA has no finitely axiomatizable consistent extension, see the Ryll-Nardzewski Theorem* [*31*].

The above results may point to the fact that the measure of distance Ad is of limited use; however, Theorem 3.13 suggests that Ad does capture a reasonably natural notion. It might be true that the distance defined in this way does not give us much information on the nature of the axioms separating two theories, adding any axiom is considered to be one step. One can overcome this problem by giving weights to the axiom adding steps, e.g., considering the addition of certain kind of axioms as two or more steps.

§4. Conceptual distance.

Now, we introduce the notion of conceptual distance by a careful translation of the algebraic idea, described in the introduction, in terms of logic.

DEFINITION 4.1. *We say that* theory T' is a one-concept-extension of theory T *and we write* $T \rightsquigarrow T'$ *iff* $\mathcal{L}' = \mathcal{L} \cup \{R\}$*, for some relation symbol* R*, and* $T \sqsubseteq T'$ *(i.e.,* T' *is a conservative extension of* T*). We also write* $T \leftrightsquigarrow T'$ *iff* $T \rightsquigarrow T'$ *or* $T' \rightsquigarrow T$*, and in this case we say that* T and T' are separated by at most one concept.[8]

DEFINITION 4.2. *Recall that* $\rightleftarrows$ *denotes definitional equivalence, see Definition 2.5. Let* $\mathcal{T}$ *be a class of theories. The step distance induced by the cluster network* $(\mathcal{T}, \rightleftarrows, \leftrightsquigarrow)$ *is called* conceptual distance on $\mathcal{T}$ *and is denoted by* $\mathsf{Cd}_{\mathcal{T}}$*. In the case when* $\mathcal{T}$ *is the class of all theories in* $\mathbf{L}^{\beta}_{\alpha}$*, we denote the conceptual distance on* $\mathcal{T}$ *by* $\mathsf{Cd}^{\beta}_{\alpha}$*.*

Let $\mathcal{T}$ be a class of theories and let T_1, T_2, T_1', and T_2' be members of $\mathcal{T}$. An immediate observation can be formulated as follows:

$$\mathsf{Cd}_{\mathcal{T}}(T_1, T_2) = 0 \text{ if } T_1 \rightleftarrows T_2, \text{ and } \mathsf{Cd}_{\mathcal{T}}(T_1, T_2) = \mathsf{Cd}_{\mathcal{T}}(T_1', T_2') \text{ if } T_1 \rightleftarrows T_1' \text{ and } T_2 \rightleftarrows T_2'.$$

Moreover, since logically equivalent theories are also definitionally equivalent by translating formulas to themselves, conceptual distance has also the following two desirable properties:

$$\mathsf{Cd}_{\mathcal{T}}(T_1, T_2) = 0 \text{ if } T_1 \equiv T_2, \text{ and } \mathsf{Cd}_{\mathcal{T}}(T_1, T_2) = \mathsf{Cd}_{\mathcal{T}}(T_1', T_2') \text{ if } T_1 \equiv T_1' \text{ and } T_2 \equiv T_2'.$$

By Remark 3.6, it is clear that $\mathsf{Cd}^{\beta}_{\alpha}(T, T') \geq \mathsf{Cd}^{\gamma}_{\alpha}(T, T')$ for any ordinal $\beta \leq \gamma \leq \alpha + 1$ and any theories T and T' in $\mathbf{L}^{\beta}_{\alpha}$. It is also apparent that an inconsistent theory is of an infinite conceptual distance from any consistent theory, because relations $\rightleftarrows$ and $\rightsquigarrow$ cannot make a consistent theory inconsistent and also cannot make an inconsistent theory consistent. Now, we give more examples.

THEOREM 4.3. *Suppose that* $\beta \geq 1$*. For every* $n \in \mathbb{N} \cup \{\infty\}$*, there are theories* T *and* T' *in* $\mathbf{L}^{\beta}_{\alpha}$ *such that* $\mathsf{Cd}^{\beta}_{\alpha}(T, T') = n$*.*

Proof. Let $\mathcal{L}_\infty = \{R_0, R_1, \ldots\}$ be a language for $\mathbf{L}^{\beta}_{\alpha}$ that consists of infinitely many relation symbols of arbitrary ranks less than β (such a language describes infinitely many different concepts). For each $k \in \mathbb{N}$, let $\mathcal{L}_k$ be the language that consists of the first k-many relation symbols of $\mathcal{L}_\infty$, i.e.,

$$\mathcal{L}_0 = \emptyset \text{ and } \mathcal{L}_k = \{R_0, \ldots, R_{k-1}\} \text{ if } k \geq 1,$$

and let $T^{\star}_k = \emptyset$ be the empty theory on language $\mathcal{L}_k$. It is clear that $\mathsf{Cd}^{\beta}_{\alpha}(T^{\star}_0, T^{\star}_n) \leq n$ for each $n \in \mathbb{N} \cup \{\infty\}$. To prove the other direction, we count models of cardinality 1. It is

[8] One may think that there is a mismatch between this definition and the algebraic motivation discussed in the introduction. In the case of FOL, the one-concept extension corresponds to the one-generator extension as described on page 636. The reason is that $R(v_{i_0}, \ldots, v_{i_{n-1}})$ is generated from $R(v_0, \ldots, v_{n-1})$ (see Footnote 6). However, if R has, say, arity ω, then, e.g., $R(v_0, v_1, \cdots)$ and $R(v_1, v_2, \ldots)$ are separate generators (meaning that none of them can be generated from the other). In fact, there is no mismatch here, because cylindric algebras correspond to logics with restricted formulas, i.e., formulas where the variables appear in the atomic subformulas only in their natural order. So formulas of the form $R(v_1, v_2, \ldots)$ are not included.

easy to see that $\mathsf{I}(T_n^\star, 1) = 2^n$. For any two theories T_1 and T_2,

$$T_1 \rightsquigarrow T_2 \implies \mathsf{I}(T_2, 1) \leq 2 \cdot \mathsf{I}(T_1, 1) \tag{1}$$

because in a model of cardinality 1 there are at most two relations (of any rank). Therefore, we need at least n-many steps to increase $\mathsf{I}(T_0^\star, 1) = 2^0 = 1$ to $\mathsf{I}(T_n^\star, 1) = 2^n$. Therefore, $\mathsf{Cd}_\alpha^\beta(T_0^\star, T_n^\star) = n$ as desired. □

The theories we use in the above proof all have models of cardinality 1. In (5) of Theorem 4.9 below, we show what happens if the theories in question do not have models of size 1. First, we need the following lemma.

LEMMA 4.4. *Suppose that $\alpha = \beta = \omega$. Let T_1, T_2 and T_3 be theories such that*

$$\mathsf{I}(T_1, 1) = \mathsf{I}(T_2, 1) = \mathsf{I}(T_3, 1) = 0.$$

Then, if $T_1 \rightsquigarrow T_2 \rightsquigarrow T_3$, then there is a theory T such that $T_1 \rightsquigarrow T \rightleftarrows T_3$.

Proof. Suppose T_1, T_2 and T_3 are as required in the statement of the lemma above, and assume that $T_1 \rightsquigarrow T_2 \rightsquigarrow T_3$. Then $\mathcal{L}_3 = \mathcal{L}_1 \cup \{R, S\}$ for some relation symbols R and S. Suppose that $\mathsf{rank}(R) = n$ and $\mathsf{rank}(S) = m$. Let $l = \max\{n, m\} + 2$ (choosing $l = \max\{n, m\} + 1$ is enough if $n, m \geq 1$). Let $\mathcal{L}^+ \stackrel{\text{def}}{=} \mathcal{L}_3 \cup \{B\} = \mathcal{L}_1 \cup \{R, S, B\}$, for some new relation symbol B of rank l. Every model $\mathfrak{M}$ for $\mathcal{L}_3$ can be extended to a model $\mathfrak{M}^+$ for $\mathcal{L}^+$ by defining $B^{\mathfrak{M}^+}$ as follows:

$$B^{\mathfrak{M}^+} \stackrel{\text{def}}{=} \big\{(a_0, \dots, a_{l-1}) \in M^l : \text{ there is an assignment } \tau \text{ for which } \\ \tau(v_0) = a_0, \dots, \tau(v_{l-1}) = a_{l-1} \text{ and } \mathfrak{M}, \tau \models \beta\big\},$$

where

$$\beta(v_0, \dots, v_{l-1}) \stackrel{\text{def}}{=} \big(R(v_0, \dots, v_{n-1}) \wedge v_{l-2} = v_{l-1}\big) \vee \big(S(v_0, \dots, v_{m-1}) \wedge v_{l-2} \neq v_{l-1}\big).$$

Let $\mathcal{L} \stackrel{\text{def}}{=} \mathcal{L}_1 \cup \{B\}$ and let

$$T \stackrel{\text{def}}{=} \big\{\varphi \in \mathsf{Fm} : \mathfrak{M}^+ \models \varphi, \text{ for every model } \mathfrak{M} \text{ for } T_3\big\}.$$

We will prove that $T_1 \rightsquigarrow T$ and $T \rightleftarrows T_3$. To prove that $T_1 \rightsquigarrow T$, it is enough to show that $T_1 \sqsubseteq T$ (because $\mathcal{L} = \mathcal{L}_1 \cup \{B\}$). Let $\varphi \in \mathsf{Fm}_1$. We have

$$T_1 \models \varphi \iff T_2 \models \varphi \iff T_3 \models \varphi \iff T \models \varphi,$$

where the first two equivalences follow by the assumption $T_1 \sqsubseteq T_2 \sqsubseteq T_3$, and the last equivalence follows by the definition of T. To show that $T \rightleftarrows T_3$, we define translations $\mathsf{tr} : \mathsf{Fm} \to \mathsf{Fm}_3$ and $\mathsf{tr}' : \mathsf{Fm}_3 \to \mathsf{Fm}$ as follows:

$$\mathsf{tr} : B(v_0, \dots, v_{l-2}, v_{l-1}) \mapsto \beta(v_0, \dots, v_{l-1}) \quad \text{and}$$

$$\mathsf{tr}' : R(v_0, \dots, v_{n-1}) \mapsto \exists v_n \dots \exists v_{l-1} \big(B(v_0, \dots, v_{l-2}, v_{l-1}) \wedge (v_{l-2} = v_{l-1})\big)$$
$$\mathsf{tr}' : S(v_0, \dots, v_{m-1}) \mapsto \exists v_m \dots \exists v_{l-1} \big(B(v_0, \dots, v_{l-2}, v_{l-1}) \wedge (v_{l-2} \neq v_{l-1})\big)$$

We have defined tr and tr' on specific basic formulas and these can be extended in a unique way to their domains, see Footnote 6. Let $\mathfrak{M}$ be a model for T_3. By definition of $\mathfrak{M}^+$,

$$\mathfrak{M}^+ \models B(v_0, \dots, v_{l-1}) \leftrightarrow \beta(v_0, \dots, v_{l-1}). \tag{2}$$

Thus, by (2), we have

$$\begin{aligned}\mathfrak{M}^+ \models \mathsf{tr}'(S(v_0,\ldots,v_{m-1})) &\leftrightarrow \exists v_m \ldots \exists v_{l-1}\left(B(v_0,\ldots,v_{l-2},v_{l-1}) \wedge v_{l-2} \neq v_{l-1}\right)\\ &\leftrightarrow \exists v_m \ldots \exists v_{l-1}\Big(\Big(\big(R(v_0,\ldots,v_{n-1}) \wedge v_{l-2} = v_{l-1}\big) \vee \\ &\qquad \big(S(v_0,\ldots,v_{m-1}) \wedge v_{l-2} \neq v_{l-1}\big)\Big) \wedge v_{l-2} \neq v_{l-1}\Big)\\ &\leftrightarrow \exists v_m \ldots \exists v_{l-1}\left(S(v_0,\ldots,v_{m-1}) \wedge v_{l-2} \neq v_{l-1})\right)\\ &\leftrightarrow S(v_0,\ldots,v_{m-1}).\end{aligned}$$

The last $\leftrightarrow$ follows by the assumption that the cardinality of $\mathfrak{M}$ is at least 2 (and hence the same is true for $\mathfrak{M}^+$). Similarly, $\mathfrak{M}^+ \models \mathsf{tr}'(R(v_0,\ldots,v_{n-1})) \leftrightarrow R(v_0,\ldots,v_{n-1})$. Therefore, by the fact that tr and tr' are translations, it follows that

$$\mathfrak{M}^+ \models \varphi \leftrightarrow \mathsf{tr}(\varphi) \text{ and } \mathfrak{M}^+ \models \psi \leftrightarrow \mathsf{tr}'(\psi) \tag{3}$$

for all $\varphi \in \mathsf{Fm}$ and $\psi \in \mathsf{Fm}_3$. By (3), it is not hard to see that tr and tr' are definitional equivalences, and the desired follows. □

The above lemma is a direct consequence of the following elementary fact. In $\mathbf{L}^{\omega}_{\omega}$ (under some conditions), for any two relations R and S, there is a relation M such that M is definable in terms of R and S and, conversely, both R and S are definable in terms of M, see [16]. The idea of the above proof is distilled from [18, Theorem 2.3.22].

COROLLARY 4.5. *Suppose that* $\alpha = \beta = \omega$. *Let* $T_1, T_2, \ldots, T_n$, *for some* $n \geq 2$, *be theories such that* $\mathsf{l}(T_i, 1) = 0$, *for each* $1 \leq i \leq n$. *Then,*

$$T_1 \rightsquigarrow T_2 \rightsquigarrow \cdots \rightsquigarrow T_n \implies T_1 \rightsquigarrow T \rightleftarrows T_n \textit{ for some theory } T.$$

Proof. This can be proved by a simple induction on n. If $n = 2$, then we are obviously done. Suppose that $n \geq 3$ and $T_1 \rightsquigarrow T_2 \rightsquigarrow \cdots T_{n-1} \rightsquigarrow T_n$. If by induction hypothesis we can assume that there is T' such that $T_2 \rightsquigarrow T' \rightleftarrows T_n$, then $T_1 \rightsquigarrow T_2 \rightsquigarrow T' \rightleftarrows T_n$. Therefore, by the Lemma 4.4, there is theory T such that $T_1 \rightsquigarrow T \rightleftarrows T_n$. □

PROPOSITION 4.6.

$$T_1 \rightsquigarrow T \rightleftarrows T_n \textit{ for some theory } T \iff T_n \textit{ faithfully interprets } T_1,$$

if $\alpha = \beta = \omega$ *and* T_n *is formulated on a finite language.*

Proof. From left to right, the statement is clear because T faithfully interprets T_1 since it is a conservative extension of T_1, and hence T_n also faithfully interprets T_1 since $T \rightleftarrows T_n$. From right-to-left, we can find a theory T_* whose language contains just one relation symbol and $T_* \rightleftarrows T_n$ by the trick used in the proof of Lemma 4.4. We have a faithful interpretation $\mathsf{tr}_{1*} : T_1 \to T_*$. Let $\widetilde{T}_*$ be T_* expanded with the language of T_1 using the faithful interpretation tr_{1*}. Then we have $T_1 \rightsquigarrow \widetilde{T}_* \rightleftarrows T_* \rightleftarrows T_n$. □

Now we are going to investigate the connection between the spectrum of theories and conceptual distance in $\mathbf{L}^{\omega}_{\omega}$. To do so, let us introduce some notations. Let $\mathcal{L}_{\emptyset}$ denote *the empty language*, i.e., the language of pure identity $=$. Let T' be a theory and $\mathcal{L}$ an arbitrary language. We define *the restriction of* T' *to* $\mathcal{L}$ as

$$T'|_{\mathcal{L}} \overset{\text{def}}{=} \{\varphi \in \mathsf{Fm}' \cap \mathsf{Fm} : T' \models \varphi\}.$$

REMARK 4.7. *It is straightforward to check that if* $\alpha \geq 1$ *and* $\mathsf{Fm} \subseteq \mathsf{Fm}'$*, then* T' *is a conservative extension of* $T'|_{\mathcal{L}}$ *and up to logical equivalence* $T'|_{\mathcal{L}}$ *is the only theory in language* $\mathcal{L}$ *whose conservative extension is* T'*.*

LEMMA 4.8. *Suppose that* $\alpha = \beta = \omega$*. Let* T_1 *and* T_2 *be two arbitrary theories of countable languages. Then*

$$T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset} \iff (\forall \textit{ cardinal } \kappa)\big[\mathsf{I}(T_1,\kappa) \neq 0 \iff \mathsf{I}(T_2,\kappa) \neq 0\big].$$

Proof. For every $n \in \mathbb{N}$, let us introduce formula $\Psi^{(n)}$ saying that there are exactly n-many objects:

$$\Psi^{(n)} \stackrel{\text{def}}{=} \exists v_0\, \exists v_1 \cdots \exists v_{n-1} \left(\left(\bigwedge_{0 \leq i \neq j \leq n-1} v_i \neq v_j \right) \land \forall v_n \left(\bigvee_{0 \leq i \leq n-1} v_n = v_i \right) \right).$$

Let us first assume that $T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset}$ and let κ be any cardinal. Suppose first that κ is finite, then T_1 does not have a model of size κ iff $T_1 \models \neg\Psi^{(\kappa)}$. But,

$$T_1 \models \neg\Psi^{(\kappa)} \iff T_2 \models \neg\Psi^{(\kappa)}$$

as $\neg\Psi^{(\kappa)} \in \mathsf{Fm}_\emptyset$. Hence, T_2 has a model of cardinality κ iff T_2 has a model of cardinality κ. If κ is infinite, then by Lövenheim–Skolem Theorem, T_1 and T_2 have models of cardinality κ iff they have infinite models. Let us also introduce formulas $\Psi^{(\leq n)}$ saying that there are at most n-many objects:

$$\Psi^{(\leq n)} \stackrel{\text{def}}{=} \forall v_0\, \forall v_1 \cdots \forall v_n \left(\bigvee_{0 \leq i \neq j \leq n} v_i = v_j \right).$$

Theory T_1 has an infinite model iff $T_1 \not\models \Psi^{(\leq n)}$ for all $n \in \mathbb{N}$. Since $\Psi^{(\leq n)} \in \mathsf{Fm}_\emptyset$,

$$T_1 \models \Psi^{(\leq n)} \iff T_2 \models \Psi^{(\leq n)}.$$

Hence, T_1 has an infinite model and thus a model of cardinality κ iff T_2 has such a model. Consequently, $\mathsf{I}(T_1,\kappa) \neq 0$ iff $\mathsf{I}(T_2,\kappa) \neq 0$ for all cardinal κ.

The converse direction follows from the simple fact that the validity of a formula $\varphi \in \mathsf{Fm}_\emptyset$ in a model depends only on the cardinality of the model. □

THEOREM 4.9. *Suppose that* $\alpha = \beta = \omega$*. Let* T_1 *and* T_2 *be two theories formulated in countable languages. Then,*

$$\mathsf{Cd}^{\omega}_{\omega}(T_1,T_2) < \infty \implies (\forall \textit{ cardinal } \kappa)\big[\mathsf{I}(T_1,\kappa) \neq 0 \iff \mathsf{I}(T_2,\kappa) \neq 0\big]. \tag{4}$$

If T_1 *and* T_2 *are formulated in finite languages, then the converse of* (4) *is also true and*

$$\big[\mathsf{Cd}^{\omega}_{\omega}(T_1,T_2) < \infty \textit{ and } \mathsf{I}(T_1,1) = \mathsf{I}(T_2,1) = 0\big] \implies \mathsf{Cd}^{\omega}_{\omega}(T_1,T_2) \leq 2. \tag{5}$$

Proof. Since the validity of formulas of the empty language $\mathcal{L}_\emptyset$ is preserved under conservative extensions and definitional equivalences, we have

$$\mathsf{Cd}^{\omega}_{\omega}(T_1,T_2) < \infty \implies T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset}. \tag{6}$$

Hence, by Lemma 4.8, T_1 has a model of cardinality κ iff T_2 has such a model because T_1 and T_2 are formulated on countable languages.

To prove the converse of (4), let us assume that T_1 and T_2 are formulated in finite languages and that, for every cardinal κ, T has a model of cardinality κ iff T' has such

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

a model. By Lemma 4.8, we have $T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset}$. Let $T_\emptyset := T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset}$. Thus, by Remark 4.7, $T_\emptyset \sqsubseteq T_1$ and $T_\emptyset \sqsubseteq T_2$. Now, we can add the whole $\mathcal{L}$ to $\mathcal{L}_\emptyset$ in finitely many steps, because there only finitely many relation symbols in $\mathcal{L}$, thus $\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) < \infty$. Similarly, $\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_2) < \infty$. Therefore,

$$\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) \leq \mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) + \mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_2) < \infty.$$

Now suppose that $\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) < \infty$ and $\mathsf{I}(T_1, 1) = \mathsf{I}(T_2, 1) = 0$, and that T_1 and T_2 are formulated in finite languages. Then, by (6), we can introduce $T_\emptyset := T_1|_{\mathcal{L}_\emptyset} = T_2|_{\mathcal{L}_\emptyset}$ as before. We claim that

$$\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) \leq 1 \text{ and } \mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_2) \leq 1. \tag{7}$$

To show (7), let us assume that $\mathcal{L}_1 = \{R_i : i < m\}$, for some finite m. If $m = 0$, then $T_\emptyset \equiv T_1$ and thus $\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) = 0$. Assume that $m \neq 0$. Let $\mathcal{L}^\star_0 = \{R_0\}, \ldots, \mathcal{L}^\star_{m-1} = \{R_0, \ldots, R_{m-1}\}$ and, for each $0 \leq i \leq m-1$, $T^\star_i = T_1|_{\mathsf{Fm}^\star_i}$. Hence, $T_\emptyset \rightsquigarrow T^\star_0 \rightsquigarrow \cdots \rightsquigarrow T^\star_{m-1} \equiv T_1$. Clearly, for each $0 \leq i \leq m-1$, $\mathsf{I}(T^\star_i, 1) = 0$ because $\neg\Psi^{(1)} = \neg\Big(\exists v_0\, \forall v_1\, (v_0 = v_1)\Big)$ is a consequence of T_1, and hence is a consequence of $T^\star_i$. Thus, by Corollary 4.5, it follows that $\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) \leq 1$. Similarly, one can show that $\mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_2) \leq 1$. Therefore, $\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) \leq \mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_1) + \mathsf{Cd}^{\omega}_{\omega}(T_\emptyset, T_2) \leq 2$. □

COROLLARY 4.10. *The conceptual distance between the theories of any two finite models of different cardinalities is infinite. More precisely, if $\mathfrak{A}$ and $\mathfrak{B}$ are two finite models of different cardinality, then* $\mathsf{Cd}^{\omega}_{\omega}\big(\mathsf{Th}(\mathfrak{A}), \mathsf{Th}(\mathfrak{B})\big) = \infty$.

For instance, given two cyclic groups $\langle k_1\rangle$ and $\langle k_2\rangle$ of orders 5 and 7, respectively, the conceptual distance between the theories of these groups is ∞. This might seem strange; these theories are about similar structures. But if we look carefully at the statement of the above corollary, we will find that it talks about theories of structures, not structures themselves. In other words, the conceptual distance between the theories of $\langle k_1\rangle$ and $\langle k_2\rangle$ cannot be granted as a distance between these two groups as algebraic structures. Instead, this conceptual distance can be considered to be a distance between the Lindenbaum–Tarski algebras of the theories of these groups, which are of course of different nature than the groups themselves.

COROLLARY 4.11. *There are infinitely many theories that are, in terms of conceptual distance, infinitely far from each other in* $\mathbf{L}^{\omega}_{\omega}$.

THEOREM 4.12. *Let* $\mathsf{CC}^{\omega}_{\omega}$ *be the class of all complete and consistent theories in* $\mathbf{L}^{\omega}_{\omega}$. *Then*

$$\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) = \mathsf{Cd}_{\mathsf{CC}^{\omega}_{\omega}}(T_1, T_2)$$

for all $T_1, T_2 \in \mathsf{CC}^{\omega}_{\omega}$.

Proof. Let $T_1, T_2 \in \mathsf{CC}^{\omega}_{\omega}$. Then, clearly, $\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) \leq \mathsf{Cd}_{\mathsf{CC}^{\omega}_{\omega}}(T_1, T_2)$. In this proof, let us denote that theory $\widetilde{T}$ is a complete and consistent extension of theory T in the same language as $T \subseteq_{\mathsf{cc}} \widetilde{T}$. We are going to show the converse inequality using the following three simple facts:

a.) If $T \rightleftarrows T'$ and $T \subseteq_{\mathsf{cc}} \widetilde{T}$, then there is $\widetilde{T}' \in \mathsf{CC}^{\omega}_{\omega}$ such that $T' \subseteq_{\mathsf{cc}} \widetilde{T}'$ and $\widetilde{T} \rightleftarrows \widetilde{T}'$.

We can take $\widetilde{T}' := \big\{\mathsf{tr}(\varphi) : \varphi \in \widetilde{T}\big\}$, where tr is the interpretation of T to T' showing $T \rightleftarrows T'$.

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

b.) If $T \rightsquigarrow T'$ and $T \subseteq_{\mathsf{cc}} \widetilde{T}$, then there is $\widetilde{T}' \in \mathsf{CC}^{\omega}_{\omega}$ such that $T' \subseteq_{\mathsf{cc}} \widetilde{T}'$ and $\widetilde{T} \rightsquigarrow \widetilde{T}'$.

We can take $\widetilde{T}'$ to be any complete and consistent extension of $\widetilde{T} \cup T'$. We have that $\widetilde{T} \cup T'$ is also consistent because otherwise there was a formula φ in the language of T such that $T' \models \varphi$ and $\widetilde{T} \models \neg\varphi$, but then $T \not\models \varphi$ contradicting that T' is a conservative extension of T.

c.) If $T \leftsquigarrow T'$ and $T \subseteq_{\mathsf{cc}} \widetilde{T}$, then there is $\widetilde{T}' \in \mathsf{CC}^{\omega}_{\omega}$ such that $T' \subseteq_{\mathsf{cc}} \widetilde{T}'$ and $\widetilde{T} \leftsquigarrow \widetilde{T}'$.

We can take $\widetilde{T}'$ to be the restriction of $\widetilde{T}$ to the language to T'.

Using a.), b.) and c.), any chain of theories from T_1 to T_2 realizing $\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2)$ can be replaced step-by-step with one that contains only complete theories and represents the same distance. Hence $\mathsf{Cd}^{\omega}_{\omega}(T_1, T_2) \geq \mathsf{Cd}_{\mathsf{CC}^{\omega}_{\omega}}(T_1, T_2)$. □

§5. Conceptual distance in physics. Each physical theory is established based on some preliminary decisions. These decisions are suggested by the accumulation and the assimilation of new knowledge. The methods used to improve physical theories are intuitively conceived and applied in a fruitful way, but many obvious ambiguities have appeared. To eliminate these ambiguities, it was critical to introduce the *logical foundation of the physical theories*.

Even today the logic based axiomatic foundation of physical theories is intensively investigated by several research groups. For example, the Andréka–Németi school axiomatizes and investigates special and general relativity theories within ordinary first-order logic, see, e.g., [5], [2], and [4]. For similar approaches related to other physical theories, see, e.g., [8] and [23].

Following the tradition of Andréka–Németi school, two theories ClassicalKin and SpecRel are formulated in ordinary first-order logic $\mathbf{L}^{\omega}_{\omega}$ to capture the intrinsic structures of classical and relativistic kinematics. For the precise definitions of these theories, one can see [25, p.67 and p. 69]. In this section, we will investigate the conceptual distance between these two theories.

In [24] and [25], it was shown that these two theories can be turned definitionally equivalent by the following two concept manipulating steps:

(1) adding the concept of an observer “being stationary” to the theory of relativistic kinematics SpecRel, and

(2) removing the concept of observers “not moving slower than light” from the theory of classical kinematics ClassicalKin.

Then, it was shown that even if observers “not moving slower than light” are removed from ClassicalKin the resulting theory remains definitionally equivalent to ClassicalKin and hence adding only the concept of “being stationary” to SpecRel is enough to make the two theories equivalent. Thus, it follows that the conceptual distance between relativistic and classical kinematics is 1.

THEOREM 5.1. *Classical and relativistic kinematics are distinguished from each other by only one concept, namely the existence of some distinguished observers captured by formula* (8) *below, i.e.,* $\mathsf{Cd}^{\omega}_{\omega}(\mathsf{ClassicalKin}, \mathsf{SpecRel}) = 1$.[9]

[9] It is worth noting that in the proof of Theorem 5.1, we add only a unary concept E to SpecRel to get a theory definitionally equivalent to ClassicalKin.

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

Proof. The key to this result is the surprising theorem stating that the only concept which needs to be added to $\mathsf{SpecRel}$ to make it definitionally equivalent to $\mathsf{ClassicalKin}$ is a concept distinguishing a set of observers that are "being at absolute rest" as shown in [24, p.72] and [25, p.110]. Let E be a unary relation symbol corresponding to this basic concept. Axiom $\mathsf{AxPrimitiveEther}$, see [24, p. 46] and [25, p. 87], defines E as follows:

$$\exists v_0\big[\mathsf{IOb}(v_0) \wedge \forall v_1\big(E(v_1) \leftrightarrow [\mathsf{IOb}(v_1) \wedge \mathsf{st}(v_0, v_1)]\big)\big], \tag{8}$$

where IOb is a unary relation symbol that represents inertial observers and $\mathsf{st}(v_0, v_1)$ is a formula in the language of $\mathsf{SpecRel}$ capturing the idea that observers v_0 and v_1 are stationary with respect to each other. In this proof, we only need that $\mathsf{st}(v_0, v_1)$ is a formula with two free variables in the language of $\mathsf{SpecRel}$; its concrete definition plays no role here. Let

$$\mathsf{SpecRel}^{\mathsf{E}} = \mathsf{SpecRel} \cup \{\mathsf{AxPrimitiveEther}\}.$$

First, we need to prove that $\mathsf{SpecRel} \rightsquigarrow \mathsf{SpecRel}^{\mathsf{E}}$. To do so, it is enough to show that $\mathsf{SpecRel}^{\mathsf{E}}$ is a conservative extension of $\mathsf{SpecRel}$, i.e., $\mathsf{SpecRel} \sqsubseteq \mathsf{SpecRel}^{\mathsf{E}}$, because the languages of these theories differ only in the unary relation symbol E. So, we need to show that for any formula ρ of the language of $\mathsf{SpecRel}$,

$$\mathsf{SpecRel} \models \rho \iff \mathsf{SpecRel}^{\mathsf{E}} \models \rho.$$

Let ρ be an arbitrary formula of the language of $\mathsf{SpecRel}$. Since $\mathsf{SpecRel} \subseteq \mathsf{SpecRel}^{\mathsf{E}}$, $\mathsf{SpecRel} \models \rho$ implies $\mathsf{SpecRel}^{\mathsf{E}} \models \rho$. We prove the other direction by proving that, if $\mathsf{SpecRel} \not\models \rho$, then $\mathsf{SpecRel}^{E} \not\models \rho$. Let $\mathfrak{M}$ be a model of $\mathsf{SpecRel}$. Since $\mathsf{SpecRel} \models \exists v_0\, \mathsf{IOb}(v_0)$, there exists an $a \in \mathsf{IOb}^{\mathfrak{M}}$. Let us fix such element a of $\mathsf{IOb}^{\mathfrak{M}}$ and let an extension $\mathfrak{M}'$ of $\mathfrak{M}$ be defined by adding the following relation to $\mathfrak{M}$:

$$E^{\mathfrak{M}'} = \Big\{b \in \mathsf{IOb}^{\mathfrak{M}} : \exists \text{ an assignment } \tau\big[\tau(v_0) = a, \tau(v_1) = b \text{ and } \mathfrak{M}, \tau \models \mathsf{st}\big]\Big\},$$

where $\mathsf{st}^{\mathfrak{M}}$ is the binary relation defined by formula $\mathsf{st}(v_0, v_1)$ in model $\mathfrak{M}$. By construction, $\mathfrak{M}'$ is a model of $\mathsf{SpecRel}^{\mathsf{E}}$. Therefore, if $\mathfrak{M} \models \neg\rho$, then $\mathfrak{M}' \models \neg\rho$, because $\mathfrak{M}'$ is an extension of $\mathfrak{M}$ means that $\mathsf{Th}(\mathfrak{M}) \sqsubseteq \mathsf{Th}(\mathfrak{M}')$. Consequently, $\mathsf{SpecRel} \not\models \rho$ implies $\mathsf{SpecRel}^{\mathsf{E}} \not\models \rho$, which is what we wanted to prove. This completes the proof of $\mathsf{SpecRel} \rightsquigarrow \mathsf{SpecRel}^{\mathsf{E}}$, and hence

$$\mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}, \mathsf{SpecRel}^{\mathsf{E}}) \leq 1.$$

By Corollary 9 in [24, p. 72] and [25, p. 110], $\mathsf{SpecRel}^{\mathsf{E}}$ is definitionally equivalent to $\mathsf{ClassicalKin}$. Hence,

$$\mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}^{\mathsf{E}}, \mathsf{ClassicalKin}) = 0.$$

Therefore,

$$\mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}, \mathsf{ClassicalKin}) \leq \mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}, \mathsf{SpecRel}^{\mathsf{E}}) + \mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}^{\mathsf{E}}, \mathsf{ClassicalKin}) = 1.$$

Moreover, $\mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}, \mathsf{ClassicalKin})$ cannot be 0 since $\mathsf{SpecRel}$ and $\mathsf{ClassicalKin}$ are not definitionally equivalent, see Theorem 5 in [24] or [25]. Consequently,

$$\mathsf{Cd}^{\omega}_{\omega}(\mathsf{SpecRel}, \mathsf{ClassicalKin}) = 1$$

and thus the desired result is reached. □

There are several ways of capturing the structures of relativistic and classical kinematics in first-order logic. Let us now introduce another way to capture these theories. Let $\mathbb{R}$ be

https://doi.org/10.1017/S1755020319000558 Published online by Cambridge University Press

the set of all real numbers. Let $\mathsf{Ph} \subseteq \mathbb{R}^4 \times \mathbb{R}^4$ be such that $(\bar{x}, \bar{y}) \in \mathsf{Ph}$ iff coordinate points $\bar{x}$ and $\bar{y}$ can be connected by a light signal, i.e., if $(x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 - (x_4 - y_4)^2 = 0$. Let $\mathsf{S} \subseteq \mathbb{R}^4 \times \mathbb{R}^4$ be the simultaneity relation, i.e., $(\bar{x}, \bar{y}) \in S$ iff $x_1 = y_1$. Consider the models $\mathfrak{R} = \langle \mathbb{R}^4, \mathsf{Ph} \rangle$ and $\mathfrak{N} = \langle \mathbb{R}^4, \mathsf{S}, \mathsf{Ph} \rangle$, these models capture the structure of special relativity and classical kinematics, respectively.

Let $T_{\mathfrak{N}} = \mathsf{Th}(\mathfrak{N})$ and $T_{\mathfrak{R}} = \mathsf{Th}(\mathfrak{R})$. Note that $T_{\mathfrak{N}}$ is in fact a conservative extension of $T_{\mathfrak{R}}$ and the conceptual distance between them is 1, i.e., $\mathsf{Cd}^{\omega}_{\omega}(T_{\mathfrak{N}}, T_{\mathfrak{R}}) = 1$.

PROBLEM 5.2 (Hajnal Andréka). *Let $\mathcal{T}$ be the class of all theories T such that $T_{\mathfrak{R}}$ is faithfully interpreted into T and T is faithfully interpreted into $T_{\mathfrak{N}}$. Is the following true: For all $T \in \mathcal{T}$,*

$$\mathsf{Cd}^{\omega}_{\omega}(T_{\mathfrak{R}}, T) + \mathsf{Cd}^{\omega}_{\omega}(T, T_{\mathfrak{N}}) = 1?$$

If the answer to the question to the above problem is yes, then no matter which classical (i.e., $T_{\mathfrak{N}}$-definable, but not $T_{\mathfrak{R}}$-definable) concept we add to special relativity ($T_{\mathfrak{R}}$) we will get classical kinematics ($T_{\mathfrak{N}}$). That would be an interesting insight for better understanding the connection between classical and relativistic concepts.

The investigation in this section opens so many questions: For any two concrete theories of physics, what is the conceptual distance between them? By Theorem 5.1, relativistic and classical kinematics are of conceptual distance one. However, the question "what is the distance between relativistic and classical dynamics?" remains open. Another natural related open problem is the following.

PROBLEM 5.3 (Jean Paul Van Bendegem). *What is the conceptual distance between classical and statistical thermodynamics*?

Of course, any answer to the above problems depends on the chosen axiomatizable theories capturing the physical theories in question. For an axiomatic approach of these thermodynamics theories, one can see, e.g., [11], [13], and [27].

§6. Concluding remarks. We have introduced a general framework to investigate measures of distances between formal theories, and we investigated some basic properties of two natural examples for concrete distances. One is based on counting axioms separating logically nonequivalent theories. The other is based on counting concepts separating definitionally nonequivalent theories.

We have found that, even though it is probably the most natural idea, counting axioms does not give us much information about the distance between theories. More precisely, we have shown that axiomatic distance between two theories formulated in the same language is at most 3, and it is exactly 2 if the theories are complete but not logically equivalent.

Counting concepts is much more subtle and informative. We have shown that conceptual distance has a full range, i.e., any possible distance is realized. In ordinary first-order logic, we found that theories formulated in finite languages are of finite conceptual distance from each other if they have models over exactly the same cardinals. Consequently, there are infinitely many theories which are conceptually infinitely far from one another.

Are the notions introduced here the right ones? Do we get useful notions of distance in this way? Does this approach really contribute to our understanding of theories in physics (or of other sciences)? We do not know. These ideas are new and we need to work with them and develop them further before we can really assess their worth. What we do know is that it is worth developing a working framework in which one can tell how far certain nonequivalent theories are from each other. In this article, we tried to lay down the foundations and to initiate this research direction. Working in this direction may provide a "logical network" conception of the unity of physics.

To test or refine these notions, to learn which notion of theory-distance will be the most useful, and to understand how useful investigating distances between theories is, we need further investigations. Of course, there may be other reasonable ideas that may be worth considering. For example, it may be natural to check whether using (certain kinds of) interpretations as steps could lead to a useful notion of distance or not.

Although the problems introduced at the end of §5 and our result that relativistic and classical kinematics are of conceptual distance 1 from each other already indicates that the notion of conceptual distance might be useful to understand connections between scientific theories. It is worth noting that probably not just their distance but also the shortest paths connecting two theories in the corresponding cluster network is interesting.

Even though our framework is quite general, one may wish to generalize it even further. For example, in several cases, it might be natural not to assume the symmetry of distances between theories. For example, any inconsistent theory is understood to be of axiomatic distance 1 from any consistent theory; we just need to add a contradiction as an axiom. But starting from an inconsistent theory, we can never reach a consistent one by adding axioms; so considering this distance to be ∞ seems more natural. By dropping the symmetry requirement from our Definition 3.2, one can easily generalize our framework to investigate these kinds of *"directed" distances*.

Another reasonable way to generalize our framework is to define *bidirected distances* or *multidirected distances* where the minimal steps can be determined by two or more relations. For example, it might be natural to try introducing a "bidirected conceptual distance" that measures the minimum number of concepts needed to be added to or removed from a theory to reach another one up to definitional equivalence. We already have a notion for concept adding. So only a notion of concept removal is needed. The way the concept of faster-than-light observers was removed from the theory capturing classical kinematics in [24] and [25] might be a good starting point to find such a notion of concept removal. Similarly, one may desire to introduce and investigate a *bidirected axiomatic distance* where there are distinct steps for removing and/or adding an axiom.

One area of study that has a close relationship with the notion of conceptual distance is that of complexity. As we know, complexity, also can be measured in several ways: Turing complexity, in terms of the analytic hierarchy, and so on. If one theory is more complex than another in one of these measures, then it is natural to investigate the relationship between that and the distances we look at here. Some of the significance of the present work might be in its relationship to complexity theory. This is a subject of future investigation.

The idea of having a notion of distance between theories (of the same nature) seems applicable in any science. In computer science, programming languages and other systems can be seen as axiomatized theories. For more details about this, see, e.g., [15], [20], and [28]. Hence, it seems also natural to search for the best fit notion of equivalence between these theories. Developing this may give us insight to determine what can be one step difference between two such theories. Having these in mind, a distance can then be defined in the same way as §3 herein. The novelty here would be in choosing such equivalence and one step relation in a way that guarantees that the corresponding step distance is applicable.

Our work and approach leaves it open what we chose to do next in our investigations concerning the equivalence of theories and the distance between theories. We can change the languages of the theories; moving from FOL to another formal language. We can change the concept used to measure distance (axioms, concepts and so on). We can change the area of research we investigate (physics, mathematical theories, computer science languages and so on). It is in trying out the ideas presented here, in these new directions that we gain new insights, and refine the techniques and definitions started here.

§7. Acknowledgments. We shall thank Hajnal Andréka, Jean Paul Van Bendegem, Zalán Gyenis, Péter Juhász, István Németi and others for their fruitful comments and questions. We are especially grateful to Judit X. Madarász because of her insightful comments leading us to significantly improve some of the definitions herein. We are deeply indebted to the anonymous referees for their careful reading and for their contributions to some of the work presented in this article. We shall especially thank one of the referees for providing simple proofs for a number of theorems, e.g., Proposition 3.11, Theorems 3.13 and 4.12.

BIBLIOGRAPHY

[1] Andréka, H., Comer, S. D., Madarász, J. X., Németi, I., & Ahmed, T. S. (2009). Epimorphisms in cylindric algebras and definability in finite variable logic. *Algebra Universalis*, **61**(3), 261–282.

[2] Andréka, H., Madarász, J. X., & Németi, I. (2004). Logical analysis of relativity theories. In Hendricks, V., Neuhaus, F., Pedersen, S. A., Scheffler, U., and Wansing, H., editors. *First-order Logic Revisited*. Berlin: Logos Verlag, pp. 7–36.

[3] Andréka, H., Madarász, J. X., & Németi, I. (2005). Mutual definability does not imply definitional equivalence, a simple example. *Mathematical Logic Quarterly*, **51**(6), 591–597.

[4] Andréka, H., Madarász, J. X., Németi, I., & Székely, G. (2012). A logic road from special relativity to general relativity. *Synthese*, **186**(3), 633–649.

[5] Andréka, H., Madarász, J. X., Németi, I., with contributions from: Andai, A., Sági, G., Sain, I., & Tőke, C. (2002). *On the Logical Structure of Relativity Theories*. Research report, Alfréd Rényi Institute of Mathematics, Budapest: Hungarian Academy of Sciences. Available at: https://old.renyi.hu//pub/ algebraic-logic/Contents.html.

[6] Andréka, H. & Németi, I. (2017). How many varieties of cylindric algebras are there. *Transactions of the American Mathematical Society*, **369**, 8903–8937.

[7] Andréka, H., van Benthem, J., & Németi, I. (1998). Modal languages and bounded fragments of predicate logic. *Journal of philosophical logic*, **27**(3), 217–274.

[8] Baltag, A. & Smets, S. (2005). Complete axiomatizations for quantum actions. *International Journal of Theoretical Physics*, **44**(12), 2267–2282.

[9] Banerjee, A. & Khaled, M. (2018). First order logic without equality on relativized semantics. *Annals of Pure and Applied Logic*, **169**(11), 1227–1242.

[10] Barrett, T. W. & Halvorson, H. (2016). Glymour and quine on theoretical equivalence. *Journal of Philosophical Logic*, **45**(5), 467–483.

[11] Carathéodory, C. (1909). Untersuchungen über die grundlagen der thermodynamik. *Mathematische Annalen*, **64**, 355–386.

[12] Chang, H. (2012). *Is Water H_2O? Evidence, Realism and Pluralism*. Dordrecht: Springer.

[13] Cooper, J. L. B. (1967). The foundations of thermodynamics. *Journal of Mathematical Analysis and Applications*, **17**, 172–193.

[14] Fine, K. (1985). Natural deduction and arbitrary objects. *Journal of Philosophical Logic*, **14**(1), 57–107.

[15] Floyd, R. W. (1967). Assingining meanings to programs. *Proceedings of Symposia in Applied Mathematics*, **19**, 19–32.

[16] Goodman, N. (1943). On the simplicity of ideas. *The Journal of Symbolic Logic*, **8**(4), 107–121.

[17] Henkin, L. (1950). The Completeness of Formal Systems. Ph.D. Thesis, Princeton University.

[18] Henkin, L., Monk, J., & Tarski, A. (1971). *Cylindric Algebras Part I*. Amsterdam: North-Holland.

[19] Henkin, L., Monk, J., & Tarski, A. (1985). *Cylindric Algebras Part II*. Amsterdam: North-Holland.

[20] Hoare, C. A. R. (1969). An axiomatic basis for computer programming. *Communications of the ACM*, **12**(10), 576–583.

[21] Japaridze, G. & Jongh, D. (1998). The logic of provability. In Buss, S. R., editor. *Handbook of Proof Theory*, Vol. 137. Amsterdam: Elsevier, pp. 475–550.

[22] Khaled, M. (2019). The finitely axiomatizable complete theories of nonassociative arrow frames. *Advances in Mathematics*, **346**, 194–218.

[23] Krause, D. & Arenhart, J. R. B. (2017). *The Logical Foundations of Scientific Theories: Languages, Structures, and Models*. New York: Routledge.

[24] Lefever, K. (2017). Using Logical Interpretation and Definitional Equivalence to Compare Classical Kinematics and Special Relativity Theory. Ph.D. Thesis, Vrije Universiteit Brussel.

[25] Lefever, K. & Székely, G. (2018a). Comparing classical and relativistic kinematics in first-order-logic. *Logique et Analyse*, **61**(241), 57–117.

[26] Lefever, K. & Székely, G. (2018b). On generalization of definitional equivalence to nondisjoint languages. *Journal of Philosophical Logic*, **48**(4), 709–729.

[27] Lieb, E. H. & Yngvason, J. (2000). A fresh look at entropy and the second law of thermodynamics. *Physics Today*, **53**(4), 32–37.

[28] Meyer, A. R. & Halpern, J. Y. (1982). Axiomatic definitions of programming languages: A theoretical assessment. *Journal of the Association for Computing Machinery*, **29**(2), 555–576.

[29] Pinter, C. C. (1978). Properties preserved under definitional equivalence and interpretations. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **24**, 481–488.

[30] Ryll-Nardzewski, C. (1952). The role of the axiom of induction in elementary arithmetic. *Fundamenta Mathematicae*, **39**, 239–263.

[31] Visser, A. (2006). Categories of theories and interpretations. In Enayat, A., Kalantari, I., and Moniri, M., editors. *Logic in Tehran. Proceedings of the Workshop and Conference on Logic, Algebra and Arithmetic, held October 18–22, 2003*, Lecture Notes in Logic, Vol. 26. Wellesley, MA: ASL, A.K. Peters, Ltd., pp. 284–341.

FACULTY OF ENGINEERING AND NATURAL SCIENCES
BAHÇEŞEHIR UNIVERSITY
ISTANBUL, TURKEY
E-mail: mohamed.khalifa@eng.bau.edu.tr

SET THEORY, LOGIC AND TOPOLOGY
ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS
BUDAPEST, HUNGARY
E-mail: szekely.gergely@renyi.hu

CENTRE FOR LOGIC AND PHILOSOPHY OF SCIENCE
VRIJE UNIVERSITEIT BRUSSEL
BRUSSELS, BELGIUM
E-mail: koen.lefever@vub.be

DEPARTMENT OF PHILOSOPHY
GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC, USA
and
UNIVERSITÉ DE LILLE
LILLE, FRANCE
E-mail: michele@gwu.edu