

FORTHCOMING IN *Manuscripto*

**ISSUES IN THE FOUNDATIONS OF SCIENCE, I:  
LANGUAGES, STRUCTURES, AND MODELS**

NEWTON C. A. DA COSTA

*Department of Philosophy  
Federal University of Santa Catarina  
Florianópolis, SC 88040-900, Brazil  
ncacosta@terra.com.br*

DÉCIO KRAUSE

*Department of Philosophy  
Federal University of Santa Catarina  
Florianópolis, SC 88040-900, Brazil  
deciokrause@gmail.com*

OTÁVIO BUENO

*Department of Philosophy  
University of Miami  
Coral Gables, FL 33124-4670, USA  
otaviobueno@mac.com*

**Abstract:** In this first paper of a series of works on the foundations of science, we examine the significance of logical and mathematical frameworks used in foundational studies. In particular, we emphasize the distinction between the *order of a language* and the *order of a structure* to prevent confusing models of scientific theories (as set-theoretical structures) with first-order structures (called here *order-1* structures), and which are studied in standard (first-order) model theory. All of us are, of course, bound to make abuses of language even in putatively precise contexts. This is not a problem—in fact, it is part of scientific and philosophical practice. But it is important to be sensitive to the different uses that *structure*, *model*, and *language* have. In this paper, we examine these topics in the context of classical logic; only in the last section we touch upon briefly on non-classical ones.

**Keywords:** Languages. Structures. Models of scientific theories.

## QUESTÕES SOBRE OS FUNDAMENTOS DA CIÊNCIA, I: LINGUAGENS, ESTRUTURAS E MODELOS

**Resumo:** Neste primeiro artigo de uma série de trabalhos sobre os fundamentos da ciência, investigamos a importância dos arcabouços lógicos e matemáticos empregados nos estudos de fundamentos. Em particular, enfatizamos a distinção entre a *ordem de uma linguagem* e *ordem de uma estrutura*, com o objetivo de evitar confundir modelos de teorias científicas (como estruturas conjuntistas) com estruturas de primeira ordem (que denominamos de estruturas de *ordem-1*), e que são tratadas pela teoria usual de modelos de primeira ordem. Mesmo em contextos que presumivelmente deveriam ser precisos, acabamos por nos valer de abusos de linguagem. Não reputamos esse uso como um problema—com efeito, é parte tanto da prática científica como filosófica. Mas cumpre atentar aos diferentes usos que *estrutura*, *modelo* e *linguagem* possuem. Neste artigo, examinamos esses tópicos no contexto da lógica clássica; apenas na última seção consideramos brevemente lógicas não-clássicas.

**Palavras chave:** Linguagens. Estruturas. Modelos de teorias científicas.

## 1 Introduction

The most common formal framework for mathematical discussions in the philosophy of science (and in particular, the philosophy of physics) is provided by first-order Zermelo-Fraenkel set theory with the axiom of choice: ZFC (see, for instance, Muller & Saunders 2008, Muller & Seevinck 2009).<sup>1</sup> Why ZFC? As is well known, there are several non-equivalent set theories, some of them have theorems that contradict those of ZFC. For instance, Quine and Rosser's NF contains a universal set, while ZFC, if consistent, does not.<sup>2</sup> Foundational discussions can also be articulated in terms of higher-order logics or even category theory. So, why ZFC?

Suppose that it is claimed that that ZFC was chosen for pragmatic reasons, such as familiarity, scope, simplicity, and so on. But such reasons do not end

---

<sup>1</sup>Although the axiom of choice is usually not mentioned in this context, it is important for certain considerations. We will return to this point below.

<sup>2</sup>This remark must be read with some care, for *sets* in ZFC and in NF do not amount to the same thing.

the discussion. As usually presented, ZFC is a *pure* set theory, that is, it does not involve things that are not *sets*, namely, *Urelemente*. But in the empirical sciences—and sometimes even in mathematics—we may need to deal with collections of objects that are not sets, such as collections of ants or planets. We may say that, in cases such as these, we should use a set theory with *Urelemente*, for sets are abstract objects, while ants and planets are not. But we have now moved beyond purely pragmatic considerations to reasons dealing with the adequacy of the formal framework to perform certain tasks.

Our goal here is to identify a cluster of philosophical issues that emerge as soon as we pay attention to the mathematical frameworks that are used in foundational issues about science. We do not intend to resolve such issues, but to indicate their significance. In turn, the identification of these issues should help to illuminate various aspects of foundational studies themselves. We will consider them in what follows.

## 2 The relevance of mathematical frameworks

Although we could use a distinct framework, we shall follow here common practice and assume first-order ZFC.<sup>3</sup> This allows us to speak of sets, functions, relations, and all the usual mathematical concepts that are needed in ‘standard’ mathematics and science (particularly, in physics).<sup>4</sup> Assuming ZFC is, of course, no trivial matter, given the plurality of non-equivalent set theories available (some of them articulate quite distinct conceptions of a set-theoretic universe).

Even in the scope of ‘standard’ set theory, in 1958 Skolem called attention to a significant fact: the dependence of various semantic concepts (such as Tarski’s characterization of truth) on the set-theoretic resources that are invoked in their formulation. In a remarkable (yet little known) passage, Skolem notes:

It is self-evident that the dubious character of the notion of set renders other notions dubious as well. For example, the semantic definition of mathematical truth proposed by A. Tarski and other logicians presupposes the general notion of set. (Quoted in Moore 2009.)

In other words, as Skolem saw very clearly, Tarski’s semantic conception of truth depends on the set theory that is used to express it. If that set theory changes,

---

<sup>3</sup>As is well known, the acceptance of ZFC was not immediate. For a discussion of the issue in the context of the debates between Skolem, Zermelo, Bernays, Tarski and others, see Moore 2009.

<sup>4</sup>By ‘standard’ mathematics we mean the mathematics formulated in Bourbaki’s books.

the properties of the concept of truth, as characterized by the theory, change accordingly.

But the concept of truth is not the only one that depends on set-theoretic resources. An important concept in quantum mechanics is that of an unbounded operator, such as position and momentum linear operators. For instance, in the Hilbert space  $L^2(\mathbb{R})$  of the equivalence classes of square integrable functions, we are dealing with unbounded operators. Recall that if  $A$  is a linear operator, then  $A$  is bounded just in case for any  $M > 0$  there exists a vector  $\alpha$  such that  $\|A(\alpha)\| \leq M\|\alpha\|$ . Otherwise,  $A$  is unbounded. In this context, Robert Solovay proved a significant result. Suppose that ZF is consistent (here ‘ZF’ stands for the theory ZFC without the axiom of choice), and suppose that DC stands for a weakened form of the axiom of choice according to which a ‘countable’ version of that axiom can be obtained. It then follows that ZF plus DC has a model in which each subset of real numbers is Lebesgue measurable. (Assuming the full version of the axiom of choice, it can be shown that there are subsets of the set of real numbers that are not Lebesgue measurable.)<sup>5</sup> Let us call ‘Solovay’s axiom’ (SA) the statement to the effect that ‘Any subset of  $\mathbb{R}$  is Lebesgue measurable’. In the set theory ZF+DC+SA (Solovay’s set theory), it can be proved that any linear operator is bounded (see Maitland Wright 1973). As a result, if we use Solovay’s set theory instead of standard ZF to build quantum structures, it is unclear how unbounded operators can be accommodated. In other words, depending on the set theory we consider, different properties of the relevant mathematical objects emerge.

Here is an additional example of this situation (which also involves Robert Solovay). One of the fundamental theorems related to quantum mechanics is Gleason’s theorem, which shows the existence of certain probability measures in separable Hilbert spaces. (There is no need for us to formulate the theorem explicitly here.) Solovay obtained a generalization of the theorem also for non-separable spaces, but it was necessary to assume the existence of a *gigantic orthonormal basis* whose cardinal is a measurable cardinal (see Chernoff 2009). However, the existence of measurable cardinals cannot be proved in ZFC (assuming that the latter is consistent). Thus, in order to obtain the generalization, we need to go beyond ZFC. Examples such as these show that, for certain considerations, it is extremely important to take into account the mathematical framework under use. Once again, depending on the framework, different mathematical resources are available.

---

<sup>5</sup>We shall not define here what a Lebesgue measure is. Intuitively speaking, it generalizes the usual notion of measure, involving lengths, areas etc.

Why are situations such as these important for the philosophy of science? One of the tasks of foundational studies is to provide precise formulations of the concepts under consideration, as well as to offer critical discussions and careful expositions of relevant scientific theories. In particular, among the concepts that are often used in philosophical examinations of science is that of a *model* of a scientific theory. There are, however, different senses of ‘model’. On one formulation, models are the structures that satisfy the postulates of a given theory (whether in mathematics or in the empirical sciences), and these structures, in turn, are typically sets formulated in a particular set theory, such as ZFC. Consider, for example, group theory. The *groups*—that is, the models of group theory—are structures of the form  $\mathcal{G} = \langle G, \star \rangle$ , where  $G$  is a non-empty set and  $\star$  is a binary operation on  $G$ , obeying well-known postulates. The ordered pair  $\mathcal{G}$  is then a set (of ZFC, say). Similar characterizations of models is found in geometry, algebra, classical analysis, and in the mathematical counterpart of any physical theory.

But Solovay’s ‘model’ of ZF + DC mentioned above is not a ‘model’ in this sense. Neither are the various ‘models’ of ZFC. That is, although ZF and ZFC, if consistent, have ‘models’, the latter are not *sets* of ZF or ZFC. These ‘models’ must be formulated in stronger theories (see Brignole & da Costa 1971; Fraenkel et al. 1973, §6.3).

Now, when we speak of the *models* of a scientific theory, such as usual quantum mechanics (the formalism plus a certain interpretation), which mathematical framework should we use to build such models? Presumably, in the quantum mechanical case, it cannot be Solovay’s set theory in ZF+DC+SA, since we need unbounded operators. Furthermore, in the Hilbert space formalism, we need to speak of certain basis of the relevant spaces (e.g. those formed by eigenvectors of self-adjoint operators), and for that we need the axiom of choice.<sup>6</sup>

As these examples illustrate, the choice of a suitable mathematical framework is crucial. And the proper characterization of that framework depends, in part, on the types and orders of the relevant structures. This is the topic to which we now turn.

---

<sup>6</sup>Using the axiom of choice—in one of its versions, called Zorn’s lemma—it can be proved that any vector space has a basis (see Halmos 1987, Appendix). In fact, the formalism of quantum mechanics requires the existence of the orthogonal complement of a subspace. But the existence of such a complement is equivalent to the axiom of choice (see Jech 1973, theorem 10.12, p. 148).

### 3 Structures

Model theory, as usually presented (see, for instance, Shoenfield 1967, Chapter 5; Chang & Keisler 1990) deals with what we call *order–1 structures*,<sup>7</sup> namely, structures of the kind:

$$\mathfrak{A} = \langle D, \{a_i\}_i, \{R_j\}_j, \{f_k\}_k \rangle_{i \in I, j \in J, k \in K} \quad (1)$$

where  $D$  is a non-empty set,  $a_i$  are distinguished elements of  $D$ ,  $R_j$  are  $n$ -ary relations on  $D$ , and  $f_k$  are  $n$ -ary functions on  $D$ . Since  $n$ -ary functions can be formulated as  $(n + 1)$ -ary relations, and given that each  $a_i$  can be expressed as a 0-ary function, the above structure can be rewritten as follows:

$$\mathfrak{A} = \langle D, r_i \rangle, \quad (2)$$

where  $r_i$  stands for a sequence of  $n$ -ary relations defined on  $D$ .

Note that the relations here relate the elements of the domain  $D$ . They do not relate subsets of  $D$  or more complex sets constructed from  $D$ , such as subsets of  $D$  and relations involving subsets of  $D$  and elements of  $D$ . That is, the elements of  $r_i$  are what we call *order–1 relations* only (and the corresponding structures will be called *order–1 structures*). However, for more complex applications, mainly in the empirical sciences, more general structures are required, sometimes involving several domains and relations that are not of order–1. In some cases, we need to consider relations whose relata are subsets of the domain, subsets of subsets of the domain and other complex sets, such as functions, derivatives, differentiable manifolds, Hilbert spaces, etc. In order to be more specific, we need to introduce some basic definitions (based on da Costa & Rodrigues 2007). What we will consider now is developed in ZF set theory, perhaps with the axiom of choice.

We denote by  $\mathbb{T}$  the set of *types*, which is the smallest set such that (i)  $i \in \mathbb{T}$  ( $i$  is the type of the *individuals*), and (ii) if  $t_1, \dots, t_n \in \mathbb{T}$ , then  $\langle t_1, \dots, t_n \rangle \in \mathbb{T}$ . Thus,  $i, \langle i \rangle, \langle i, i \rangle, \langle \langle i \rangle, i \rangle, \langle \langle i \rangle \rangle$  are types. Intuitively speaking, in this list we have: types for individuals, for sets (or properties) of individuals, for binary relations on individuals, binary relations whose relata are properties of individuals and individuals, and properties of properties of individuals.

We define the order of a type  $t$ , which is denoted by  $\text{Ord}(t)$ , as follows: (i)  $\text{Ord}(i) = 0$ , and (ii)  $\text{Ord}(\langle t_1, \dots, t_n \rangle) = \max\{\text{Ord}(t_1), \dots, \text{Ord}(t_n)\} + 1$ . Thus,  $\text{Ord}(\langle i \rangle) = \text{Ord}(\langle i, i \rangle) = 1$ , while  $\text{Ord}(\langle \langle i \rangle, i \rangle) = 2$ . Now let  $D$  be a set. Similarly to da Costa and Rodrigues 2007, we also introduce the function  $t_D$  with domain

<sup>7</sup>It will become clear below why this terminology is used.

$\mathbb{T}$ : (i)  $t_D(i) = D$ ; (ii) if  $t_1, \dots, t_n \in \mathbb{T}$ , then  $t_D(\langle t_1, \dots, t_n \rangle) = \mathcal{P}(t_D(t_1) \times \dots \times t_D(t_n))$ . For any  $t \in \mathbb{T}$ , the objects belonging to  $t_D(t)$  are said to be of type  $t$ , and the objects of  $t_D(i)$  are the *individuals*. If  $\varepsilon(D) := \bigcup \text{range}(T_D)$ , we call  $\varepsilon(D)$  a *scale based on D*. The elements of  $\varepsilon(D)$  that have types whose order is  $n > 0$  are called *relations* (of finite rank). In particular, unary relations are called *sets* or *properties*. By definition, the *order of a relation* is the order of its type.<sup>8</sup>

Relations of types  $\langle i \rangle$  (properties of individuals),  $\langle i, i \rangle$  (binary relations on individuals),  $\langle i, i, i \rangle$  (ternary relations on individuals), and so on, are *order-1 relations*. We shall not use the expression *first-order relation* to prevent confusion with the order of a language, which will be defined below. Similarly, relations of types  $\langle \langle i \rangle \rangle$ ,  $\langle \langle i \rangle, i \rangle$ , and so on, are *order-2 relations*. Individuals are identified with order-0 relations.

A structure  $\mathfrak{A}$  based on the set  $D$  is an ordered pair  $\mathfrak{A} = \langle D, r_i \rangle$ ; here,  $D \neq \emptyset$  and  $r_i$  is a sequence of  $n$ -ary relations belonging to  $\varepsilon(D)$ . These relations are called the *primitive elements* of the structure. Furthermore, we call  $k_D$  the cardinal associated with  $\mathfrak{A}$ , which is defined as  $k_D = \sup\{|D|, |\mathcal{P}(D)|, |\mathcal{P}^2(D)|, \dots\}$ ; here  $|X|$  is the cardinal of the set  $X$ .

Let  $\mathfrak{A} = \langle D, r_i \rangle$  be a structure. Its order,  $\text{Ord}(\mathfrak{A})$ , is defined as follows: if there is the greatest order of the relations in  $r_i$ , the order of the structure is that order; otherwise, it is  $\omega$ . If a structure has order  $\kappa$ , we say that it is an *order- $\kappa$  structure*. Groups are examples of order-1 structures, for they contain only order-1 relations (actually, they contain just one order-1 relation or of type  $\langle i, i, i \rangle$ , for the binary operation on the domain can be seen as a ternary relation).<sup>9</sup> More involved structures can contain several sets, divided up into *principal* and *secondary* sets (see Bourbaki 1968, Chapter 4). For instance, vector spaces are structures of the form  $\mathfrak{E} = \langle \mathcal{V}, F, +, \cdot \rangle$ ; here  $\mathcal{V}$  is the set of vectors,  $F$  is a field,  $+$  is a binary operation on  $\mathcal{V}$  (addition of vectors) and  $\cdot$  is the multiplication of a vector by an element of  $F$ , all of them obeying well-known axioms. Such structures can still be represented via the general form (3), for the domain can be formulated as the union of all principal and secondary sets, with suitable adaptations in the definition of the relations in question.

However, the structures that are used in the formulation of significant scientific theories, such as classical particle mechanics (Suppes 2002), non-relativistic quantum mechanics, quantum field theory, special and general relativity, and so

<sup>8</sup>Note that, as an element of the scale  $\varepsilon(D)$ , a relation has a type. For instance, a binary relation on  $D$  is a subset of  $D \times D$ , hence it has type  $t = \langle i, i \rangle$ .

<sup>9</sup>If we denote the binary operation on a set  $G$  as above by  $\star$ , so that the image of the pair  $\langle a, b \rangle$  is written  $a \star b$ , then this binary operation can be treated as the relation  $R = \{\langle a, b, a \star b \rangle : a, b, c \in G\}$ .

on, are not order-1 structures. After all, the formulation of such structures involves various domains and relations that are not of order-1, e.g., as we noted above, subsets of the domain, subsets of subsets of the domain, and more intricate sets (derivatives, differentiable manifolds, Hilbert spaces, and so on). Thus, the mathematical theory in which such structures are studied is not the usual first-order model theory. Although we still lack a thorough theory of higher-order structures, generalized Galois theory provides a framework for the study of such structures (see da Costa & Rodrigues 2007). But there is much yet to be done.

## 4 Languages

The explicit separation between first- and higher-order languages appeared for the first time in Hilbert and Ackermann's 1928 book (see Hilbert & Ackermann 1950), although it had been considered before, for instance, by Löwenheim in 1915 and by Skolem in the early 1920s. The distinction is now familiar: first-order languages allow for the quantification over only individuals of the domain of interpretation, not over properties, functions, or sets of such individuals. To quantify over all of the latter, we need higher-order languages.

As is also well known, ZFC can be axiomatized as a first-order theory. Although today the first-order formulation of ZFC is widely accepted, this was not always the case. In fact, Zermelo himself was never comfortable with Skolem's axiomatization of ZFC as a first-order theory.<sup>10</sup>

The first-order language of ZFC will be denoted by  $\mathcal{L}_\in$ , and we assume that it has only one non-logical symbol — namely, the binary predicate  $\in$  — and no individual constants (although the latter can be considered in an alternative formulation).

Suppose we have a structure  $\mathfrak{A} = \langle D, r_i \rangle$ , defined in  $\mathcal{L}_\in$ , whose domain  $D$  comprises several sets, and the relations in  $r_i$  are order- $n$  relations ( $n \geq 1$ ). Since  $\mathcal{L}_\in$  does not contain individual constants, we need to enlarge  $\mathcal{L}_\in$  with new objects in order to represent, for instance,  $\mathfrak{A}$ ,  $D$ , and the primitive relations of the structure. In this way, for instance, by enlarging the language  $\mathcal{L}_\in$  with the symbols  $G$  and  $\star$ , a group structure can be built, and by adding  $V$ ,  $K$ ,  $+$ ,  $\cdot$  (among other symbols), vector spaces can be formulated as well. These are the primitive symbols of the (corresponding) structures.

---

<sup>10</sup>Gregory Moore provides an insightful discussion of the issue in the context of the debate regarding the underlying logic of set theory (see Moore 1980, and Moore 2009).



However, not only do we need to speak of these structures, but also of all the objects of the scale built on  $D$ ,  $\varepsilon(D)$ . And depending on our needs, infinitary languages may be used. In order to show the power of  $\mathcal{L}_\varepsilon$ , we shall exemplify below how an infinitary language (and higher-order languages in general) can be built using the first-order extended language  $\mathcal{L}_\varepsilon$  plus specific symbols. A typical case is that of a language  $\mathcal{L}_{\mu\kappa}^\eta$ ; here  $\mu, \kappa$ , and  $\eta$  are cardinals such that  $\kappa \leq \mu$  and  $1 \leq \eta \leq \omega$ , and  $\omega$  is the first infinite cardinal. In this language, we can form conjunctions and disjunctions of sets of formulas of cardinality less than  $\mu$ , and blocks of quantifiers of length less than  $\kappa$ . We then have *first-order languages* when  $\eta = 1$ , *second-order languages* when  $\eta = 2$ , and so on, until  $\eta = \omega$ , which is an *order- $\omega$  language* suitable for type theory.

As an example, consider a language such as  $\mathcal{L}_{\omega_1\omega}^1$ . In this language, we can write formulas with denumerably many conjunctions and disjunctions, such as the formula we abbreviate by:  $x = 0 \vee x = 1 \vee x = 2 \vee \dots$ . This formula states that  $x$  is a natural number. (Note, however, that this expression is not a formula of our language, for the dots ‘ $\dots$ ’ are not symbols of  $\mathcal{L}_\varepsilon$ .) Usual first-order languages are of the kind  $\mathcal{L}_{\omega\omega}^1$ .

To give an idea of how significantly strong languages can be constructed in first-order ZFC, let us sketch the language  $\mathcal{L}_{\omega_1\omega}^\omega(R)$ . Here ‘ $R$ ’ stands for the collection of the primitive symbols (relations) of a certain structure we are interested in. For instance, we could consider all those languages whose cardinals can be defined in ZFC.<sup>11</sup> The language  $\mathcal{L}_{\omega_1\omega}^\omega(R)$  can be described in the following terms. First, we suppose that set  $\mathbb{T}$  of types has been defined as above. The primitive symbols are as follows:

- (i) Sentential connectives:  $\neg$ ,  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\bigwedge$ , and  $\bigvee$ .
- (ii) Quantifiers:  $\forall$  and  $\exists$ .
- (iii) For each type  $t$ , a family of variables of type  $t$  whose cardinal is  $\omega$ .
- (iv) For each type  $t$ , a family of constants of type  $t$  (some of them may be empty). All constants for the set  $R$ .
- (v) Primitive relations: for any type  $t$ , a collection of constants of that type (some of them may be empty).
- (vi) Parentheses: left and right parentheses (‘(’ and ‘)’), and comma (‘,’).

<sup>11</sup>This excludes, of course, measurable cardinals, since their existence cannot be proved in ZFC (assuming the latter’s consistency).

(vii) Equality: For each type  $t$ , an equality symbol  $=_t$  of type  $t = \langle t_1, t_2 \rangle$ , with  $t_1$  and  $t_2$  of the same type.

Variables and constants of type  $t$  are *terms* of that type. If  $T$  is a term of type  $\langle t_1, \dots, t_n \rangle$  and  $T_1, \dots, T_n$  are terms of types  $t_1, \dots, t_n$  respectively, then  $T(T_1, \dots, T_n)$  is an *atomic formula*. If  $T_1$  and  $T_2$  are terms of the same type  $t$ , then  $T_1 =_{\langle t_1, t_2 \rangle} T_2$  is an *atomic formula*. We shall write  $T_1 = T_2$  for this last formula, leaving the types implicit. If  $\alpha, \beta, \alpha_i$  are formulas ( $i = 1, \dots$ ), then  $\neg\alpha, \alpha \wedge \beta, \alpha \vee \beta, \alpha \rightarrow \beta, \bigwedge \alpha_i$ , and  $\bigvee \alpha_i$  are formulas. It is then possible to construct formulas with denumerably many conjunctions and disjunctions, which we write informally, for instance, as:  $\alpha_1 \wedge \alpha_2 \wedge \dots$ . Furthermore, if  $X$  is a variable of type  $t$ , then  $\forall X\alpha$  and  $\exists X\alpha$  are also formulas (only finite blocks of quantifiers are allowed). These are the only formulas of the language.<sup>12</sup>

What does it mean to *construct* a language such as  $\mathcal{L}_{\omega_1\omega}^\omega(R)$  in ZFC? In the intended interpretation, the individual variables of  $\mathcal{L}_\epsilon$  range over sets. However, if we consider ZFC as a *formal* theory, there is no interpretation associated with its language. Thus, all the primitive symbols of  $\mathcal{L}_{\omega_1\omega}^\omega(R)$  are terms of  $\mathcal{L}_\epsilon$ . If we assign to them the intended interpretation, these symbols stand for sets. So, ‘(’, the left parenthesis, is a name of a set, and the same point goes for all the sequences of symbols in the language. In this way, by assigning suitable sets to symbols in  $\mathcal{L}_{\omega_1\omega}^\omega(R)$ , the latter is eventually constructed in ZFC.<sup>13</sup>

Let again  $\mathfrak{A} = \langle D, r_i \rangle$  be a structure. And let us suppose that this structure represents some domain of knowledge, perhaps in the empirical sciences.<sup>14</sup> In some cases, the relations in  $\mathfrak{A}$  are partial, thereby reflecting the partiality of the information available about the domain in question (see da Costa & French 2003). The relations in  $r_i$  are the primitive relations of the structure. For instance, in Suppes’s humans paternity theory, besides a domain  $D$  of human beings, he admits two subsets of  $D$ ,  $M$  and  $L$ , which stand for the set of males and the set of living human beings, respectively. We can consider these subsets as representing *properties* of human beings: those humans who are male belong to  $M$ ; those who are alive belong to  $L$ . Moreover, the structure also includes a binary relation  $P$ , defined on  $D$ , such that for every  $x$  and  $y$  in  $D$ ,  $Pxy$  is satisfied as long as  $x$  is the father of  $y$ . In this case,  $M$ ,  $L$ , and  $P$  are the primitive relations (which here also

<sup>12</sup>The concepts of free and bound variables and other syntactic concepts are introduced as usual.

<sup>13</sup>As noted above, in order to accommodate the objects of the empirical sciences, which clearly are not sets, it would be better to employ ZFU rather than ZFC.

<sup>14</sup>We find here yet another sense of ‘model’, which stands for the ‘mathematization’ of a given domain, as exemplified by the work of engineers and applied scientists.

include primitive properties) of the structure, satisfying suitable postulates (see Suppes 1969).

Let us now suppose that  $\mathfrak{A} = \langle D, r_i \rangle$  is a structure whose relations are of order- $n$ , with  $n \geq 1$ , and  $\varepsilon(D)$  is the scale based on  $D$ . In order to speak of this structure and the objects of its scale, we need a higher-order language. According to the definitions above,  $r_i$  is a sequence of relations of the scale  $\varepsilon(D)$ , that is, it is a mapping from a finite ordinal into a collection of relations in the scale. In this case,  $\text{rng}(r_i)$  stands for the set of these relations. And  $\mathcal{L}_{\omega\omega}^\omega(\text{rng}(r_i))$  (or  $\mathcal{L}^\omega(\text{rng}(r_i))$  for short) is the language whose only constants are the primitive relations of the structure. This language will be called *the basic language of the structure*. (Of course, this is not the only possible language, for stronger languages incorporating it could be used as well.)

In this way, we can interpret a sentence of  $\text{rng}(r_i)$  in  $\mathfrak{A} = \langle D, r_i \rangle$ , and define the concept of truth for sentences of this language (according to the structure in question) in the Tarskian sense. As usual, this is denoted by:

$$\mathfrak{A} \models S. \quad (3)$$

We can similarly define the concept of *validity*. A sentence  $S$  is valid (which we denote by  $\models S$ ) if  $\mathfrak{U} \models S$ , for every structure  $\mathfrak{U}$ . The concept of truth can be extended to the whole scale (since the latter is also a set), and in this way we can also speak of a sentence being *true in a scale*, as we shall do below. If we consider partial relations as primitive relations in the relevant structure, the concept of truth must be that of *partial truth*, which generalizes Tarski's approach and is more adequate for the empirical sciences (for details, see da Costa & French 2003).

Finally, note that we have been describing a higher-order language using the resources of first-order ZFC. Even the structure  $\mathfrak{A} = \langle D, r_i \rangle$ , which need not be an order-1 structure, is being constructed in that set theory.

## 4.1 Definability

Our goal now is to formulate some conditions under which an object of a scale  $\varepsilon(D)$  is definable in a structure  $\mathfrak{A} = \langle D, r_i \rangle$  by a formula of  $\mathcal{L}^\omega(\text{rng}(r_i))$  in such a way that an element of the scale is expressible in the structure with respect to a sequence of objects of the scale. We will not discuss all of the details here, but only provide the main definitions in order to extend the standard definitions for first-order structures (see Shoenfield 1967, p. 135). The issue is important, since in order to represent a domain by a certain structure in a given language,

we need to know whether the relevant objects can in fact be so represented with the resources available.

Let  $R$  be a relation of type  $t = \langle t_1, \dots, t_n \rangle$  and  $\mathfrak{A} = \langle D, r_i \rangle$  be a structure. We say that  $R$  is *definable* in  $\mathfrak{A}$  if there is a formula  $F(x_1, x_2, \dots, x_n)$  of  $\mathcal{L}^\omega(\text{rng}(r_i))$  whose only free variables are  $x_1, \dots, x_n$  of types  $t_1, \dots, t_n$ , respectively, such that in  $\mathcal{L}^\omega(\text{rng}(r_i)) \cup \{R\}$ , the formula

$$\forall x_1 \dots \forall x_n (R(x_1, \dots, x_n) \leftrightarrow F(x_1, x_2, \dots, x_n))$$

is true in  $\varepsilon(D)$ .

For instance, in  $\mathcal{L}^\omega(\text{rng}(r_i))$ , for each type  $t$  we can define an identity relation  $=_t$  of type  $\langle i, i \rangle$  as follows. Let  $Z$  be a variable of type  $\langle t \rangle$ . Then we can easily see that for suitable structures and scales, the following is true:

$$\exists! I_t \forall x \forall y (I_t(x, y) \leftrightarrow \forall Z (Z(x) \leftrightarrow Z(y))).$$

$I_t$  can be called the identity of type  $t$ , and we can write  $x =_t y$  for  $I_t(x, y)$ . Intuitively, this means that identity, as usual, is defined by Leibniz Law. In fact, the definition just given can be rewritten in the following way:

$$x =_t y := \forall Z (Z(x) \leftrightarrow Z(y)).$$

Usually, we suppress the index  $t$  and write  $x = y$ , leaving the type implicit. (The relation  $=$  is of type  $\langle t, t \rangle$ , while  $x$  and  $y$  are both of type  $t$ .)

This kind of definability, which involves structures and scales, is called *semantic definability*, and goes back to Tarski's work. Another important case, also involving semantic definability, is the following. Let  $\mathfrak{A} = \langle D, r_i \rangle$ ,  $\varepsilon(D)$  and  $\mathcal{L}^\omega(\text{rng}(r_i))$  be a structure, a scale, and a language, respectively, as introduced above. Given an object  $a \in \varepsilon(D)$  of type  $t$ , we say that it is  $\mathcal{L}^\omega(\text{rng}(r_i))$ -definable, or definable in the strict sense, in  $\mathfrak{A} = \langle D, r_i \rangle$  if there is a formula  $F(x)$ , whose only free variable  $x$  is of type  $t$ , such that:

$$\mathfrak{A} \models \forall x (x =_t a \leftrightarrow F(x)). \quad (4)$$

As an example, let  $\mathfrak{N} = \langle \omega, +, \cdot, s, 0 \rangle$  be an order-1 structure for first-order arithmetic.<sup>15</sup> In order to define any particular natural number, we only need a finitary language, such as  $\mathcal{L}_{\omega\omega}^\omega(R)$  with  $R = \{0, s\}$ . It is then easy to see that:

$$\mathfrak{N} \models \forall x (x = n \leftrightarrow x = s s \dots s(0)),$$

<sup>15</sup>Strictly speaking, we should have written  $\{0\}$  rather than  $0$ , but we will proceed in this way.

here  $ss \dots s(0)$  abbreviates a formula of the language. If we use a suitable infinitary language, such as  $\mathcal{L}_{\omega_1\omega}$  (in which denumerably many conjunctions and disjunctions are allowed), we can write inside the parentheses of the previous formula:

$$x \in \omega \leftrightarrow x = 0 \vee x = 1 \vee \dots,$$

which allows us to define all natural numbers (assuming, of course, that we have devised a way of referring to each one of them).

Can everything be defined in this way? Not really, since definability is a relative concept, which depends on the resources of the languages we use. Here are a couple of examples to motivate this point. First, as is well known, in ZFC (assuming its consistency), the set  $\mathbb{R}$  of real numbers is not denumerable. Thus, if we use a standard denumerable language, we do not have enough *names* to name each real number, and cannot define them by the condition just given above. However, if we adopt a suitable infinitary language  $\mathcal{L}_{\nu\kappa}$ , with appropriate  $\nu$  and  $\kappa$ , it is possible to provide a name for each real number (after all, each real number can be used to name itself). In this infinitary language, real numbers can then be defined by the condition above. Second, it can be proved in ZFC that every set is well-ordered. In particular,  $\mathbb{R}$  has a well-ordering. However, as is also well known, we cannot define this well-ordering by a formula of  $\mathcal{L}_\epsilon$ . In fact, the least element of the subset  $(0, 1)$ , for instance, which exists, is also not definable in that language. As these examples indicate, definability and expressibility (of elements in a given language) depend on the language that is used. In the end, these are language dependent, relativized concepts — rather than absolute ones.

Far more could be said about these issues. But we hope we have said enough to motivate the distinction between the order of a language and the order of a structure, and to illustrate that even with the resources of first-order languages (such as  $\mathcal{L}_\epsilon$ , the language of ZFC), we can construct higher-order structures and languages (for further details, see da Costa & Rodrigues 2007). We will now consider the significance of these issues to the empirical sciences.

## 5 Leaving standard frameworks behind

In our discussion so far we have assumed a ‘classical’ framework, in the sense that we have invoked classical logic and the standard mathematics built in ZFC. But we also noted some cases in which we need to go beyond ZFC. Solovay’s generalization of Gleason’s theorem is such a case, as well as the use of category theory to formulate certain physical theories, as recently discussed by John

Baez.<sup>16</sup>

But we should also consider the use of non-classical logics, and not just ‘extensions’ of classical mathematics (which is the case of category theory and ZFC supplemented by strong axioms, such as those that postulate the existence of universes). For instance, if we intend to axiomatize Bohr’s theory of the atom, we may need some kind of paraconsistent logic, since Bohr’s theory seems to be inconsistent (da Costa & French 2003). Bressan claimed that modal logic is needed in physics (Bressan 1974), while Reichenbach (1944), Février (1951), da Costa & Krause (2005) and others have suggested the use of many-valued logics in quantum mechanics.<sup>17</sup>

Needless to say, the strength of languages such as  $\mathcal{L}_\epsilon$  is so great that virtually everything that is needed for scientific practice can be expressed in it. In fact, when scientists develop their informal mathematical theories (that is, when they formulate ‘informal mathematical models’, since they typically do not work in an axiomatic setting), there is usually no need for them to go beyond the mathematics that can be developed in a fragment of ZFC. However, this fact may conceal a significant philosophical issue. For when we consider foundational problems, the use of non-classical frameworks may be needed to articulate certain assumptions scientists implicitly make when they develop their own theories. Moreover, a non-classical framework may also be required to express and maintain certain philosophical views about science. In either case, the need for going beyond the classical framework emerges. To illustrate this point, let us consider an example.

Steven French (forthcoming) has argued that a metaphysical underdetermination arises from quantum statistics. The (non-relativistic) quantum mechanical formalism is compatible with two ‘metaphysical packages’: according to one of them, elementary particles are individuals; according to the other, they are non-individuals. As he notes:

The (now)<sup>18</sup> standard example of metaphysical underdetermination arises from quantum statistics. Philosophical reflection on the ‘new’ quantum mechanics was entwined with the development of the physics itself, with Born and Heisenberg, for example, suggesting that quantum statistics—both the Bose-Einstein and Fermi-Dirac varieties—implied that particles could no longer be regarded as individuals (see

<sup>16</sup>See his homepage, <http://math.ucr.edu/home/baez/categories.html>.

<sup>17</sup>Additional works along these lines, mainly dealing with quantum theory, are discussed in Jammer (1974), Chapter 8.

<sup>18</sup>French has also discussed other instances of underdetermination.

French and Krause 2006, pp. 94-115). For many years this was effectively the ‘received’ view of the matter, until it was argued that such particles could be regarded as individuals, subject to certain constraints (French 1989; van Fraassen 1989; French and Krause 2006). With the development of ‘non-standard’ logico-mathematical frameworks suitable for accommodating the ‘Received’ view’s ‘non-individuals’ and a detailed understanding of the aforementioned constraints, two distinct metaphysical packages can be elaborated, consistent with the physics: particles-as-non-individuals (described via quasi-set theory) [see French and Krause 2006] and particles-as-individuals (subject to certain state accessibility constraints). (French, forthcoming.)

This is an important point, and we can echo it from a different point of view. Consider the case of bosons. As is well known, in certain situations, these quantum entities share the same quantum state and are *absolutely indistinguishable*, that is, it is not possible to distinguish them. As an example, consider a system of two bosons,  $a$  and  $b$ , and two possible states,  $A$  and  $B$ . According to quantum mechanics, there are three possibilities of distributing such bosons over these states. These possibilities are known as ‘Bose-Einstein statistics’, and they are expressed by suitable vectors in a given Hilbert space:

(i)  $|\psi_{ab}\rangle = |\psi_a^A\rangle|\psi_b^A\rangle$  (that is, the joint system is composed by both bosons in the same state  $A$ );

(ii)  $|\psi_{ab}\rangle = |\psi_a^B\rangle|\psi_b^B\rangle$  (i.e., both bosons are in state  $B$ );

(iii)  $|\psi_{ab}\rangle = (1/\sqrt{2})(|\psi_a^A\rangle|\psi_b^B\rangle + |\psi_b^A\rangle|\psi_a^B\rangle)$  (that is, *one* boson is in state  $A$ , while the *other* is in  $B$ ; here ‘ $1/\sqrt{2}$ ’ is just a mathematical trick).

A fundamental postulate then states that, in (iii), it is not possible to determine which boson is in state  $A$  and which is in  $B$ . The labels,  $a$  and  $b$ , are thus just convenient devices: they do not presuppose the individuation of each boson.

Consider now the mathematics that is used to formulate quantum theory. It is typically based on a standard set theory, such as ZFC. The simple fact that we have two bosons — that is, two objects whose set has cardinal 2 — is enough to make them different. A set, Cantor noted a long time ago, is a collection of *distinct* objects. But as a result, there does not seem to be room for indiscernible objects in standard set theories, that is, objects that cannot be distinguished in

any way, except if we confine ourselves to a non-rigid structure (built in ZFC).<sup>19</sup>  
<sup>20</sup>

The ‘solution’ to this problem sketched above in the framework of standard mathematics — namely, to work with symmetric vectors — only conceals the real difficulty. When we consider the symmetric vector (iii), we are simply putting aside the distinction between  $a$  and  $b$  that we had initially made. In fact, we had *supposed* that the vectors  $|\psi_a\rangle$  and  $|\psi_b\rangle$  describe the bosons, we had labeled or ‘named’ the latter, and thus we had counted them as two. The maneuver is fine for certain purposes, and the practice of physics works quite well in this way. However, from a foundational point of view, there is a philosophical difficulty here. The simple fact that two elementary particles, e.g. two electrons, are characterized by different wave functions in classical logic makes them distinct. It is then unclear how they can possibly be indiscernible (or ‘identical’ in the physicist’s jargon).

An answer we favor, and that we can only sketch here, consists in distinguishing three levels of the language in use. Let us suppose again that  $\mathfrak{A} = \langle D, r_i \rangle$  is a structure, whose language (in the sense discussed above) is  $\mathcal{L}_\epsilon(R)$ . We then have the following ‘working levels’ of language: (i) In the *syntactic level*, we formulate and articulate scientific concepts, define the relevant elements, and so on. (ii) The *semantic level* is divided into two additional levels: (a) In the *formal semantic level*, we work *inside* the structure  $\mathfrak{A}$  and obey its limits, e.g. we call ‘indiscernible’ (relative to that structure) those elements (of the domain) which are invariant under the automorphisms of the structure. (b) In the *informal semantic level*, we find implicit ‘interpretations’, according to which, e.g., ‘electron’ designates a certain entity etc. (iii) Finally, the third level is *pragmatic*, and it involves assumptions that strictly speaking are not part of the relevant theories, but which are required for the latter to function properly; for instance, the be-

---

<sup>19</sup>But even in this case the indistinguishability is only relative to the relations of the structure, for any structure in ZFC can be extended to a rigid one (see da Costa & Rodrigues 2007). A rigid structure only has one automorphism, namely, the identity function. In general, two objects of the domain of a structure  $\mathfrak{A}$  are said to be indiscernible in  $\mathfrak{A}$  if there is an automorphism of  $\mathfrak{A}$  (i.e. a bijective function on its domain that ‘preserves’ all the relations of the structure) which leads one of them into the other. In a rigid structure, the elements are indiscernible only to themselves.

<sup>20</sup>As an additional example, recall the fact that, given the axiom of choice, the set  $\mathbb{R}$  of real numbers is well ordered. This entails that any subset of  $\mathbb{R}$  has a least element. Thus, if we consider two disjoint subsets, e.g. the intervals  $(0, 1)$  and  $(2, 3)$ , their least elements are of course distinct, although we cannot define them in accordance with the definitions given above. (We cannot even name them using the standard language  $\mathcal{L}_\epsilon$  properly extended to accommodate real analysis.) In other words, even if we cannot represent the least element of a subset of  $\mathbb{R}$  relative to a well ordering of  $\mathbb{R}$ , this element is different from any *other* real number.



havior of measuring apparatuses, the way we read their results and advance the phenomenological counterpart of a given theory.

These levels can be, and often are, incompatible with each other. This fact need not trouble us. Indeed, scientists apparently are not that worried about contradictions, for they seem to avoid the latter by taking refuge in suitable (apparently) consistent sub-theories. Even from a foundational point of view, we can accommodate this apparent inconsistency by invoking a paraconsistent logic (da Costa et al. 2007). But this is, of course, a topic for another occasion.

## References

- [1] BRESSAN, A. 1974, 'On the usefulness of modal logic in axiomatizations of physics', in Kenneth F. Schaffner and Robert S. Cohen (eds.), *PSA 1972*, 285-303.
- [2] CARNAP, R. 1958, *Introduction to Symbolic Logic and its Applications*. New York: Dover.
- [3] CHANG, C.C. and KEISLER, H. J. 1990, *Model Theory*. (3rd. edition.) Amsterdam: North-Holland.
- [4] CHERNOFF, P. R. 2009, 'Andy Gleason and quantum mechanics', *Notices of the American Mathematical Society*, November, 1253-1259.
- [5] DA COSTA, N. C. A. and KRAUSE, D. 2005, 'Remarks on the applications of paraconsistent logics to physics', in M. Pietrocola and Olival Freire, Jr. (eds.), *Filosofia, Ciência e História: Michel Paty e o Brasil, uma homenagem aos 40 anos de colaboração*. São Paulo: FAPESP, Discurso Editorial, 337-362.
- [6] DA COSTA, N. C. A. and RODRIGUES, A. M. N. 2007, 'Permutation and invariance', *Studia Logica* 82, 1-30.
- [7] DA COSTA, N. C. A., KRAUSE, D. and BUENO, O. 2007, 'Paraconsistent logics and paraconsistency', in D. Jacquette (ed.), *Philosophy of Logic*. Amsterdam: North-Holland, 791-911.
- [8] DOMENECH, G., HOLIK, F. and KRAUSE, D. 2008, 'Q-spaces and the Foundations of Quantum Mechanics', *Foundations of Physics* 38, 969-994.

- [9] DOMENECH, G., HOLIK, F., KNIZNICK, D. and KRAUSE, D. 2009, 'No Labeling Quantum Mechanics of Indiscernible Particles', forthcoming in *International Journal of Theoretical Physics*.
- [10] FRAENKEL, A.A., BAR-HILLEL, Y. and LEVY, A. 1973, *Foundations of Set Theory*. (2nd edition.) Amsterdam: North-Holland.
- [11] FÉVRIER, P. D. 1951, *La Structure des Théories Physiques*. Paris: Presses Universitaires de France.
- [12] FRENCH S. 1989, 'Identity and individuality in classical and quantum mechanics', *Australasian Journal of Philosophy* 67, 432-446.
- [13] FRENCH, S. (forthcoming), 'Metaphysical underdetermination: why worry?', to appear in *Synthese*.
- [14] FRENCH, S. and KRAUSE, D. 2006, *Identity in Physics: A Historical, Philosophical, and Formal Analysis*. Oxford: Oxford University Press.
- [15] HEISENBERG, W. 1949, *The Physical Principles of Quantum Theory*. New York: Dover.
- [16] HILBERT, D. and ACKERMANN, W. 1950, *Principles of Mathematical Logic*. New York: Chelsea.
- [17] JAMMER, M. 1974, *The Philosophy of Quantum Mechanics*. New York: John-Wiley and Sons.
- [18] JECH, T. 1973, *The Axiom of Choice*. New York: Dover.
- [19] JECH, T. 1977, 'About the axiom of choice', in Barwise, J. (ed.), *Handbook of Mathematical Logic*. Amsterdam: North-Holland, pp. 345-370.
- [20] MAITLAND WRIGHT, J. D. 1973, 'All operators on a Hilbert space are bounded', *Bulletin of the American Mathematical Society* 79 (6), 1247-1250.
- [21] MOORE, G. H. 1980, 'Beyond first-order logic: the historical interplay between mathematical logic and axiomatic set theory', *History and Philosophy of Logic* 1, 95-137.
- [22] MOORE, G. H. 2009, 'Review of *Ernst Zermelo: An Approach to His Life and Work*, by Heinz-Dieter Ebbinghaus in cooperation with Volker Peckhaus', *Notices of the American Mathematical Society*, August, 823-827.

- [23] MULLER, F. A. and SAUNDERS, S. 2008, 'Discerning fermions', *British Journal for the Philosophy of Science* 59, 499-548.
- [24] MULLER, F. A. and SEEVINCK, M. P. 2009, 'Discerning elementary particles', arXiv: 0905.3273v1.
- [25] REICHENBACH, H. 1944, *Philosophic Foundations of Quantum Mechanics*. Berkeley and Los Angeles: University of California Press.
- [26] SCHRÖDINGER, E. 1957, *Science, Theory, and Man*. London: Routledge.
- [27] SHOENFIELD, J. R. 1967, *Mathematical Logic*. New York: Addison-Wesley.
- [28] SUPPES, P. 1969, *Studies in the Methodology and Foundations of Science*. Dordrecht: Reidel.
- [29] SUPPES, P. 2002, *Representation and Invariance of Scientific Structures*. Stanford: CSLI.
- [30] VAN FRAASSEN, B. C. 1980, *The Scientific Image*. Oxford: Clarendon Press.
- [31] VAN FRAASSEN, B. C. 1989, *Quantum Mechanics: An Empiricist View*. Oxford: Clarendon Press.