PRINCIPIA **26**(1): 39–54 (2022) doi: 10.5007/1808-1711.2022.e86052
Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

MODELS AND MODELING IN SCIENCE: THE ROLE OF METAMATHEMATICS

DÉCIO KRAUSE
Graduate Program in Logic and Metaphysics – UFRJ, BRAZIL
deciokrause@gmail.com

Abstract. The use of models of scientific theories should not be done without qualifications about the mathematics being used to build the models. This looks obvious, at least for logicians, but generally, it is not to the philosopher of science. Thus, some details about this point seem useful for both. Since any quick revision in the literature shows that in most cases, mainly after the raising of the semantic approach (to scientific theories), the models are taken to be set-theoretical structures, in discussing the issue we shall be concerned more with set theories, the *locus* where the play is usually developed (yet sometimes unconsciously).

Keywords: Scientific theories • structures • models • set theories • categories • metaphysics • non-individuals • quasi-set theory

RECEIVED: 05/02/2022 ACCEPTED: 04/04/2022

> To present a theory is to specify a family of structures, its *models*
>
> Bas van Fraassen 1980, p.64

> [A] possible realization of a theory is a set-theoretical entity of the appropriate logical type.
>
> Patrick Suppes 2002, p.21

1. Introduction

It is presented here a quite particular way to understand the scientific enterprise (at least for those called 'physical sciences' and mathematics) and scientific theories in general (mainly in these fields). Agreeing that scientific activity is a conceptual activity, it is recalled that a simple collecting of concepts means nothing. For instance, in classical particle mechanics, important concepts (or notions) are those of 'particle', 'position', 'time', 'mass', 'inner forces', 'external forces', among others, such as 'momentum', 'acceleration', etc. But, what can we do with these concepts? Of course, we need to *organize* them, since we intend to say that 'every *particle* has a *mass*', that 'the sum of the *inner forces* with the *external forces* of a *particle* in any instant of *time* is equal to the second derivative of the *position* relative to *time*', and so on. In the same vein, if we admit that Euclidean geometry can be described from the concepts of *point*, *line*, and *plane*, *right angle*, *parallel lines* and etc., then we need to relate them by sentences such as 'two distinct *points* have just one *line* in common', 'parallel lines have no point in common', and so on. In other words, scientific knowledge is not only conceptual knowledge, but *structured* conceptual knowledge (Krause and Arenhart, 2017).[1] How do we organize the chosen notions that will base our theory? The answer is immediate: we do it by the formulation of the theory's *postulates* or *axioms*.

Together, the postulates define a *class of structures*, the *models* or the *realisations* of the theory. For most of the standard theories, such as those of mathematics, physics and much more, the defined structures are *sets* in a standard set theory, such as the ZFC system (Zermelo-Fraenkel with the Axiom of Choice). Firstly, the reader of course should notice that we are speaking of the axiomatic method. The class of models of a theory has its own 'model', which is an abstract mathematical structure that shows how the models are to be, defining an *species of structures* they belong to (more on this below). What we are trying to say is that the abstract notion of *structure* is essential in the scientific enterprise and also in the mathematical description of scientific theories, which in standard cases corresponds to presenting a set-theoretical predicate, as indicated in Suppes' quotation at the beginning.[2]

Although a scientific theory is usually born from a motivation,[3] after its development as an axiomatic theory it becomes *autonomous* from the original field that motivated its development, becoming an abstract mathematical entity that may be interpreted differently from the original motivation. Take for instance the natural numbers 0, 1, 2, ..., which we use daily in the supermarket and notice that this sequence can be characterized by a first-order language whose 'typical' structure is of the form $\mathscr{N} = \langle N, s, \hat{0} \rangle$, where N is a non-empty set, s is an injective function from N to N and $\hat{0} \in N$. These notions are linked (structured) by Peano's axioms. The standard interpretation considers the set $\mathbb{N}$ of the natural numbers, the operation $s(x) = x + 1$ and $\hat{0}$ being 0. But there exist also other interpretations, say taking the domain yet as $\mathbb{N}$ but now $s(x) = x + 2$ and $\hat{0}$ as 1; Peano's axioms are *true*[4] also in this structure. As we see, in mathematics the different 'realizations' can come from other abstract structures, while in science it is expected that they come from some outside 'reality'. But this raises a huge problem to be considered below (see Section 5).

Anyway, the former structure gives origin to a class of mathematical structures, the *models* of the theory (described by the axioms). We refer to a *class* of structures because, in general, such a collection is not a *set* in standard set theories; for instance, although any real vector space can be seen as a set of a theory such as the ZFC system (the Zermelo-Fraenkel set theory with the Axiom of Choice, see Fraenkel, Bar-Hillel and Levy 1973), the collection of all real finite-dimensional vector spaces is not a *set* of ZFC; we shall explain this important detail later (Section 5).

Let us consider the following question: can we believe the two sentences put in the preamble? They were made by two authorities in the field, so the chances are great that they are to be agreed, independently of any fallacy by the appeal to authority. Really, in both quotations, there are some hidden notions that, for certain kinds of analysis (philosophical), should be made explicit and as we shall see, they cannot be taken without considering the contexts. Thus, this paper is written in order to consider the following questions at least:

(1) What is a structure? (as in van Fraassen's quotation, but in his (van Fraassen 1980) he does not say what a structure is supposed to be)
(2) What is a model? ("a structure which satisfies the axioms of a theory" van Fraassen 1980, p.43)
(3) Where do they live? That is, *where* the models are constructed? (no one of them considers this explicitly; Suppes refers to naïve set theory).
(4) What kind of things do models model? (wide discussion)
(5) Why do we usually consider set theory and not category theory or higher-order logic, or perhaps a still different framework? (no clear answer).

We hope to provide a way to understand these important questions from a particular perspective. Let us move.

2. Structures

The word 'structure' has many senses and is used in different contexts, from linguistic to architecture and chemistry. But here we shall be occupied with the mathematical sense of the term since we are pursuing the quotations that motivate this paper. In our basic logic courses, we usually learn something about logic systems starting from a certain formal language, to which it is later given an *interpretation*, which means to define a *mathematical structure*. For first-order systems, an interpretation consists of a non-empty domain (as we shall assume here) and an interpretation function, which attributes to the non-logical symbols of the language something related to the elements of the domain (see Mendelson, 1997). But in general, in science and even in mathematics we proceed from the other way around. The case of arithmetics given above is a good example; we know in advance what we intend to organize, or at least we think that we know; in this case, the sequence of the natural numbers.[5] We could also be interested in the set of the real numbers under its usual order and operations, or the system of classical mechanics, the theory of natural selection, and so on. That is, we start with the structures and *then* we try to find a language for speaking of *that* structure, at least in principle, and later realize that the axiomatics we have elaborate is open to 'other' interpretations.[6]

We need first to acknowledge that mathematics and scientific theories require more than what we term *order-1* structures.[7] These are composed of one or more domains and operations, relations and distinguished elements over these domains. There are no relations whose arguments are also relations, for instance. But, in several situations, we need to quantify over subsets of these domains or on relations that relate not only the elements of the domain(s) but other relations and operations over these elements. For example, take topological spaces. A topological space is a structure of the following species (see below for the term 'species of structures'): $\mathscr{T} = \langle D, \tau \rangle$ where D is non-empty and τ is a collection of subsets of D, the *topology*. Some axioms must be obeyed by the elements of τ. Thus, we are involved with things that do not relate the elements of the domain, but collections of sets of elements of D.

This is something the philosopher of science should take into account: most of the scientific (and even mathematical) structures are not order-1 structures, and so cannot be dealt with by standard Model Theory (Button and Walsh 2018). By the way, we notice that there is no general model theory for higher-order structures. The lesson is that the languages for speaking of most structures we can find in science are also not elementary (first-order), and this appears also in mathematics. Typical examples are well-orderings: a well-ordering over a set S is a binary relation R over this set so that (i) R is a partial ordering (irreflexive and transitive),[8] and (ii) every non-empty subset of S has a least element relative to R, that is, for every non-empty $X \subseteq S$, there exists $m \in X$ such that mRa for every $a \in X$.[9] As we see, we need to speak of *for every subset of* X, that is, we need to quantify over sets of elements of the domain, and such a sentence is not elementary (first-order). As for the above example of a topological space, a typical axiom is the following: *for any* two elements $A, B \in \tau$, their intersection $A \cap B$ does belong to τ. Of course, we are quantifying over subsets of the domain, and not over its elements.

But, how do we get structures? As we have said before, very roughly speaking a structure is an n-tuple formed by one or more sets and collections of distinguished elements, relations and operations over these sets, or over any element of the *scale* based on such sets, that is, it may comprise relations whose *relata* are also relations, sets of sets, and so on. Perhaps the best way to grasp the idea is to have a look at Bourbaki's notion of *species of structures*, but without all the tedious details (which can be found in Chapter 4 of his (Bourbaki, 2006)). We work in a set theory, such as the ZFC system. Then consider a collection of *principal* non-empty sets $E_1, E_2, \ldots, E_n$ and a collection of *auxiliary* sets $A_1, \ldots, A_m$. For the sake of simplicity, we shall consider just three principal sets, E, F, G and two auxiliary sets A and B. Using set-theoretical basic operations of cartesian product and power, we can get other sets, such as $E \times A$, $\mathscr{P}(F \times \mathscr{P}(G))$, $\mathscr{P}(E \times E \times E)$, etc. This way, we obtain a *scale* of sets based on the principal and auxiliary sets we have chosen and can form the structures we are interested in. We can introduce, as Bourbaki did, several related notions, such as those of *deduced* and *derived* structures, *homomorphism* and *isomorphism* between structures, *canonic extensions* of structures, and so on. But the development of these topics is not relevant for our argumentations here.

Let us give some simple examples, starting with mathematics. More detailed cases can be seen in da Costa and Doria (1992). Take groups, one of the most important structures of all science (mathematics, physics, crystallography, learning, and so on). Roughly speaking a group is a structure formed by just one principal set G and a binary operation '$*$' over G which obeys well-known postulates: (i) $*$ is associative, (ii) the operation admits a neutral element, and (iii) every element of G has an 'inverse' relative to $*$ which also belongs to G. Note that there are no auxiliary sets. Thus, in order to grasp such a structure, we start with the set G, and then form the scale $G \times G$, then $G \times G \times G$ and, finally (for our necessities), $\mathscr{P}(G \times G \times G)$. Now we choose an element $* \in \mathscr{P}(G \times G \times G)$ satisfying conditions which reflect the axioms for groups,[10] getting a structure of the form $\mathscr{G} = \langle G, * \rangle$. This development gives us a class of structures, anyone of them modeling the group postulates. Each structure of this species is a group, and the above abstract structure $\mathscr{G} = \langle G, * \rangle$ characterizes a *species of structures* of groups (more details below).

Now let us consider vector spaces. Here we need a principal set of 'vectors' $\mathscr{V}$ and one auxiliary set, the set F which stands for the domain of a *field* $\mathscr{F} = \langle F, +, \cdot, 0, 1, \rangle$.[11] The other elements we need are the vectors' 'addition', $+ \in \mathscr{P}(\mathscr{V} \times \mathscr{V} \times \mathscr{V})$ and the multiplication of vectors by scalars, $\cdot \in \mathscr{P}(F \times \mathscr{V} \times \mathscr{V})$, obeying the known postulates we shall not repeat here. This qualifies the species of structures of vector spaces, that is, structures of the form $\mathscr{E} = \langle \mathscr{V}, \mathscr{F}, +, \cdot \rangle$.

More precisely, a *species of structure* Σ is defined by Bourbaki this way. We take a collection of principal base sets $x_1, \ldots, x_n$, a collection of auxiliary bases sets $A_1, \ldots, A_m$ and a specific *echelon construction schema* he writes

$$S(x_1, \ldots, x_n, A_1, \ldots, A_m).$$

This is the notation for $\mathscr{P}(G \times G \times G)$ in the case of groups and $\mathscr{P}(\mathscr{V} \times \mathscr{V} \times \mathscr{V}) \times \mathscr{P}(F \times \mathscr{V} \times \mathscr{V})$ in the case of vector spaces.

An element $\mathfrak{s} \in S(x_1, \ldots, x_n, A_1, \ldots, A_m)$ is the *typification* of Σ. The typification is written by Bourbaki as a formula $T(x_1, \ldots, x_n, \mathfrak{s})$. In our samples, this corresponds to the selections we made: $* \in \mathscr{P}(G \times G \times G)$ and $\langle +, \cdot \rangle \in \mathscr{P}(\mathscr{V} \times \mathscr{V} \times \mathscr{V}) \times \mathscr{P}(F \times \mathscr{V} \times \mathscr{V})$.

Now let $R(x_1, \ldots, x_n, \mathfrak{s})$ be a transportable formula with respect to the given typification, with the x_i as the principal sets and the A_j as the auxiliary sets. This formula corresponds to Suppes' set-theoretical predicate, although there are differences in the two approaches, as pointed out in Krause and Arenhart (2017) and in da Costa and Krause (2020).[12] This formula will be the *axiom* of the species of structures with typification T. If we select some particular sets $E_1, \ldots, E_n, U$ so that both $T(E_1, \ldots, E_n, U)$ and $R(E_1, \ldots, E_n, U)$ hold, then U is said to be *a structure of species* Σ (as a particular group or a real vector space). Bourbaki's first example is that of the species of structures of ordered sets, where from a set A, we get (by a suitable echelon construction schema S) the set $\mathscr{P}(A \times A)$ and the typification $\mathfrak{s} \in \mathscr{P}(A \times A)$ (a binary relation on A), with the axiom $\mathfrak{s} \circ \mathfrak{s} = \mathfrak{s}$ (reflexivity) and $\mathfrak{s} \cap \mathfrak{s}^{-1} = \Delta_A$ (transitivity), being Δ_A the diagonal of A (informally, the set $\Delta_A = \{(x, x) : x \in A\}$). Other examples can be found in Bourbaki (2006), pp.263ff.

Thus, the restrictions imposed to $\mathfrak{s}$ constitute the axioms of the species of structure (in the case of groups, the restriction is that $\mathfrak{s}$ (which we have called '$*$', must be associative, admits a neutral element and that every element of the domain has an inverse also in the domain).

All these examples of structures (and models) given above lie in a set theory, that is, they are *sets*, and so do not cause any troubles for the practising mathematician or for the philosopher. Next, we shall present a sample case study that will enable us to emphasize the importance of considering metamathematics in presenting the models of a physical scientific theory.

3. Orthodox quantum mechanics

Here we shall present a more well-developed example of an important physical theory in order to enlighten the notions delineated above. We shall leave implicit the tedious details of making explicit the species of structures and their constructions, proceeding as the physicist does. By the way, this is what Bourbaki himself does; after detailed developments in his book (Bourbaki, 2006), the rest of his oeuvre proceeds as the standard mathematician does. For comments on his way of working, see Mashaal (2006).

In the standard formulation of orthodox (non-relativistic) quantum mechanics (henceforth, QM) via Hilbert spaces, *states* of quantum systems and *observables* over them are considered. We notice in advance that the quantum systems themselves are usually referred to, but do not play any relevant role in the axioms, to be summarized below. This will constitute an important fact for our argumentation about the need of considering the metamathematics where the models of the theory are being built. Let us follow Krause and Arenhart (2017), chap.5 for a detailed presentation.

The primitive notions are *system* (quantum system), *observable*, and *state*. The following postulates are them posed:

Postulate 1 Let S be the set of physical (quantum) systems. To each physical system $s \in S$ we associate a Hilbert space $\mathscr{H}$. Composite quantum systems are associated with complex Hilbert spaces that are the tensor product of the Hilbert spaces for each system, as usual.

Postulate 2 The one-dimensional subspaces of $\mathscr{H}$ denote the *states* the system may be in. These spaces are called *rays* by the physicists. To simplify the notation, usually, they are represented by unitary vectors ψ (or by $|\psi\rangle$ in Dirac's notation) that generate these spaces. Hence, $c\psi$, for a non-zero $c \in \mathbb{C}$, represents the same state as ψ, so as does $\psi.c$ (this is a typical physicists' abuse of notation).[13] These vectors are said to represent *pure states* of the system. It is also postulated that linear combinations of pure states, that is, vectors of the form $\psi = \sum_n a_n\psi_n$, for $a_n \in \mathbb{C}$ (the linear combination may comprise any finite number of vectors), called *superpositions* by the physicist, also denote pure states. This assumption is called *the Superposition Postulate*.

Postulate 3 To each observable (the physical quantity that can be measured) A we associate a self-adjoint operator $\hat{A}$ over $\mathscr{H}$

Postulate 4 The possible values of the measurement of observable A for the system s in the state ψ lie in the spectrum (the set of eigenvalues) of the associated operator $\hat{A}$. This was called *the Quantization Algorithm* by M. Redhead (1987), p.5.

Postulate 5 Here we have the *Born Rule*. Given a system s, which is associated to a 4-tuple σ as we describe below (Definition 3.1), let A be an observable to be measured on the system in state ψ. First we take the Hilbert space of the states of the system, $\mathscr{H}$. Now, let $\{\alpha_n\}$ be an orthonormal basis for $\mathscr{H}$ formed by eigenvectors of $\hat{A}$ (something that is possible to assume, since $\hat{A}$ is diagonalizable[14]), so that there are complex numbers c_n such that $\psi = \sum_n c_n\alpha_n$, with $\sum |c_n|^2 = 1$. The c_n are the *Fourier coefficients* $c_n = \langle \alpha_n|\psi\rangle$, where $\langle \cdot | \cdot \rangle$ is the inner product. Let us denote the eigenvalues associated to the vectors α_n by a_n, that is, $\hat{A}\alpha_n = a_n\alpha_n$. Then we have *the Statistical Algorithm* (Redhead 1987, p.8): the probability that the measurement of observable A gives the value a_n when the system is in the state ψ is

$$\mathsf{prob}_A^{\psi}(a_n) = |c_n|^2 = |\langle \alpha_n|\psi\rangle|^2$$

for the non-degenerate state (that is, all eigenvalues of $\hat{A}$ are distinct); when the operator is degenerate, the probability is obtained by summing the $|c_j|^2$ for all α's associated to the same eigenvalue.[15]

Other possible states that are not pure are called *mixtures*. They can be briefly described by means of *statistical operators* (or statistical matrices) (Redhead 1987, pp.15–6). We assign probabilities w_k to a set of pure states $\{\beta_k\}$ in which the system may be found, so that we have a *statistical ensemble* of several quantum (possible) states. Let P_{β_p} denote the projection operator whose range is the unitary sub-space generated by β_p. Then the statistical operator for the system becomes

$$\rho = \sum_k w_k P_{\beta_k},$$

and the *expectation value* of an observable A is given in terms of its Hermitean associated operator by

$$\langle A \rangle := \mathsf{Tr}(\rho \cdot \hat{A}),$$

where Tr is the trace function.

Postulate 6 Let us call this *the Dynamic Postulate*. It says that if the system is in the instant t_0 in state $\psi(t_0)$ — here the notation is adapted in order to consider the state as depending on time — then in a distinct time t the system evolves to the state $\psi(t)$ according to the Schrödinger equation (Penrose 2005, p.536; Redhead 1987, p.12)

$$\psi(t) = \hat{U}(t)\psi(t_0),$$

where $\hat{U}$ is a unitary operator.

Postulate 7 This is *the Collapse Postulate*. It says that if immediately after the measurement of observable A for the system in state $\psi = \sum_n c_n \alpha_n$, gives the value $|c_n|^2 = |\langle \alpha_n | \psi \rangle|^2$, then the system enters in the state described by the corresponding eigenvector α_n.

The above axiomatics gives rise to a mathematical abstract structure of the following kind:

$$\mathcal{QM} = \langle S, \{H_i\}, \{\hat{A}_{ij}\}, \{U_{ik}\}, \mathscr{B}(\mathbb{R}) \rangle,\ i \in I, j \in J, k \in K \tag{1}$$

being I, J, K sets of indices, and where

(i) S is a collection[16] whose elements are called *physical objects*, or *physical systems*.

(ii) $\{H_i\}$ is a collection of mathematical structures, namely, complex separable Hilbert spaces whose dimension is defined in the particular application of the theory.

(iii) $\{\hat{A}_{ij}\}$ is a collection of self-adjunct (or Hermitian) operators over a particular Hilbert space H_i.

(iv) $\{U_{ik}\}$ is a collection of unitary operators over a particular Hilbert space H_i

- (v) $\mathscr{B}(\mathbb{R})$ is the collection of Borel sets over the set of real numbers.

Notice that postulates do not speak of space (Hilbert spaces are not where the 'real' things live; they are just the place for the state-vectors), something essential when we intend to apply it to the 'real world'. This notion is introduced as follows.

Definition 3.1 *To each quantum system*[17] $s \in S$ *we associate a 4-tuple of the form*

$$\sigma = \langle \mathbb{E}^4, \psi(\mathbf{x}, t), \Delta, \mathsf{prob} \rangle.$$

Here, $\mathbb{E}^4$ is the Galilean spacetime;[18] each point is denoted by a 4-tuple $\langle x, y, z, t\rangle$ where $\mathbf{x} = \langle x, y, z\rangle$ denote the coordinates of the system and t is a parameter representing time, $\psi(\mathbf{x}, t)$ is a function over $\mathbb{E}^4$ called the *wave function* of the system, $\Delta \in \mathscr{B}(\mathbb{R})$ is a Borelian,[19] and prob is a function defined, for some i (determined by the physical system s), in $\mathscr{H}_i \times \{\hat{A}_{ij}\} \times \mathscr{B}(\mathbb{R})$ and assuming values in $[0, 1]$, so that the value $\mathsf{prob}(\psi, \hat{A}, \Delta) \in [0, 1]$ is the probability that the measurement of the observable A (represented by the self-adjoint operator $\hat{A}$) for the system in the state $\psi(\mathbf{x}, t)$ lies in the Borelian set Δ. We can see the relationship between the state vector and the wave function as follows. Let $(\mathbf{x}, t)$ denote the location operation at time t. Then we put $\psi(\mathbf{x}, t) = \langle(\mathbf{x}, t)|\psi\rangle$, that is, the wave-function is described by the coefficients of the expansion of the state vector in the orthonormal basis of the position operator.

It seems clear that the particular structures of this species are not order-1 structures, and doubtful can be transformed in some. We notice also that, contrary to the standard approaches, we are here explicitly introducing the set S of quantum systems, and this brings us the following question: would S be really a *set*, a collection of *distinct* objects, as Cantor has said? The answer is that this is disputable; the most acknowledged interpretation (some variant of the Copenhagen interpretation) accept that quantum objects of the same kind (electrons, protons, photons, etc.) are *indistinguishable*, or *indiscernible*, and in some situations, cannot be discerned from each other in any way. So, if the quantum system s is formed by entities of this kind, say a cluster of bosons in a bosonic condensate, nothing we can suppose can discern them. This poses a challenge to the semantics of our axiomatics, that is, the background where the particular structures of the species (1) can be constructed.

Thus, we have proposed that S should be not a *set* of standard set theories such as ZFC,[20] but a *quasi-set*, an entity described by the theory of quasi-sets (Krause 1990, 1992; French and Krause 2006). In such a theory, collections of indiscernible objects can be considered. Once more, the details do not concern us here; the interested reader can consult French and Krause (2006), and Krause, Arenhart, and Bueno (2022) for the corresponding philosophy.

4. The role of the metamathematics

We have seen that the mathematical structures that serve as models of theories are built in a certain metamathematical framework. We have also claimed that there is no precise sense in attributing *directly* some physical object (that is, an object living in our physical world) to the terms of the language (of the theory), for we cannot use the rules of logic for dealing with such an association; what we have said is that we need to consider some other mathematical structure which 'represents' the relevant elements of the domain and then we use this structure as an interpretation of the language of our theory, so we are apt to use the standard semantic rules.

But all of this is done within certain metamathematics. Here we schematize some situations where the use of an informal set theory or some 'standard' set theory such as the ZFC system may be put into parenthesis and questioned. Let us see.

The quantum case shown above tells us that in certain situations we need to have some care with the (meta)mathematics where the relevant structures are to be built. Really, in the Hilbert space formalism, one usually makes use of unbounded operators over the relevant Hilbert spaces, such as those that stand for position, momentum or energy.[21] So, the metamathematics needs to be able to accept their existence. But what happens if instead of a standard set theory (such as the ZFC system) or even quasi-set theory, we use the so-called Solovay's set theory (or Solovay's 'model'), which is ZF (ZFC without the axiom of choice) plus DC, the Axiom of Dependent Choice (that is, $S = ZF + DC$)? In such a theory, every linear operator over a Hilbert space is bounded (Maitland-Wright, 1973); we would be in trouble for using the above formalism. A well-documented case that not all models of ZFC are adequate to support the development of quantum mechanics is given by the two papers by Paul Benioff from 1976 (Benioff 1976a; 1976b).

The same would happen if instead of a standard set theory such as the ZFC system we would make use of ZFA, the Zermelo-Fraenkel system with atoms, entities that are not sets but which can be elements of sets (Suppes, 1960). The problem is that we can construct 'permutation models' such as those of Läuchli, which enable the construction of Hilbert spaces with no basis or then with bases of different cardinalities (Jech, 1977). Since the existence of bases is fundamental for the H-space formalism, we would be again in trouble.

The third example is that one mentioned above, namely, the situation where S (see again the structure (1)) is not to be counted as a *set*, an entity whose elements can always (even if only in principle) be discerned from one another. As we have said, in certain situations, collections of quantum objects cannot be discerned in any way or, then, even if they have some distinctive value of some property, such as spin in a given direction, in most cases we cannot know which is which, that is, which quantum has a certain value for the property, and which one has a different value for the same property. They remain indiscernible.

Finally, an example involving mathematics. Think of a set theory such as the ZFC system, axiomatized as a first-order theory. If consistent, it has models but is not categorical, that is, its models are not isomorphic. The worst situation is this: where these models are built? Notice that while the structures presented in the previous section are *sets* of, say, ZFC or of quasi-set theory, the models of ZFC cannot be sets of ZFC. This is impeded by the Second Incompleteness Theorem (look at Smith 2021, chap.17, for a clear exposition you can surpass "without too many tears"). That is, a model of a theory satisfying certain conditions of recursivity, expressiveness, and consistency, cannot enable the constructions of models for itself. The models of a theory like ZFC need to be considered in strong theories such as those involving *universes* or (equivalently) assuming the existence of inaccessible cardinals. So, since any theory of sets is also a 'scientific theory', we need to be careful in considering its models.

5. Debts paid

In the introduction, we have promised to explain two points that are relevant to the present discussion; here we pay the debt.

Physical theories and 'reality' A physical theory is not just a mathematical abstract theory; physics is not mathematics. So, although a well developed physical theory is elaborated to make reference to a certain domain of knowledge or a field of application, it has a *mathematical counterpart* which is a purely abstract mathematical structure to which interpretations can be done under the canons of standard logic, which (once the axioms of the theory are true in the corresponding structure) turn to be the models of the theory.

But, as discussed earlier, this mathematical structure (which in general can be constructed in different ways) may express also 'other models'. We choose one of them to be the *intended* one, which *by hypothesis* stands for the mathematization of the field of knowledge we had in mind when we started working. Here is where the problem lies. In my opinion, in precise terms, we cannot ascribe to a certain term of the language of the theory an element in our lab or in our surroundings; we do it only informally. Logical precision, in my account, is impossible to get. How should we do it? The only way could be to stick a tag in the object in which the name of the term was written, but we know that this is impossible to do in most cases (for instance, when the term is a derived one, such as the specific weight of a material, which is the quotient of the weight by the unity of volume).

Just as a parenthesis, I agree with M. L. Dalla Chiara and G. Toraldo di Francia in that there is no distinction between theoretical terms and observable terms as in the old logical empiricism (Suppe, 1977). According to them, all terms are 'theoretical', even those that correspond to a supposed 'direct observation' with our eyes. Our eyes are like lenses, and the thing we observe depends on several conditions; let us read them:

> Let us suppose that we are observing a fish in a basin. Do we observe it directly, yet the light rays perform part of their trajectories in water? Very probably we shall say that yes, we do it. But what if the basin is an aquarium, so that part of the trajectory is done in the glass? Furthermore, what if part of the glass is a little convex in such a way that the fish is seen a little bit great? What is the glass is substituted by a lens? And what is instead of a lens we have two lenses as microscopy? Dalla Chiara and Toraldo di Francia (1981), p.39.

They conclude that "it is absurd to think that the distinction between direct observation and observation throughout an instrument has a precise sense and that should have importance in principle" and that in reality there is no distinction between observable and theoretical terms: all of them are theoretical.

Going back to our discussion, in ascribing meaning to a (theoretical) term of our scientific language, we need to make another previous construction, one that *represents* the field of knowledge we are dealing with. In other words, there exists a middle term between theory and 'reality', namely, the *structuring* this 'reality' in a mathematical structure so that we can apply the rules of standard semantics to associate meaning to the terms of the language. In short, we never attribute (meaningfully) to the term 'alpha particle' *that* thing we supposedly observe in our laboratory, say in a bubble chamber, since we do not 'observe' it in any way, but just indications of its existence. What we do, to be precise is to construct a mathematical representation (of course in terms of structures) of an alpha particle and associate it to the term of our language as being *this* entity, which stands for an alpha particle, according to the semantic rules. This is what we have said before that between theory and 'reality' there is an intermediary usually hidden entity, a mathematical structure that represents that reality and which can interact, via standard semantics, with the theory. Theory and reality, strictly speaking, never interact.

There is a cluster of questions to be discussed here, but we shall leave them for another work. The resume is that the association of physical theories and 'reality' is never direct, but pass by an intermediary mathematical representation of such a 'reality', for if not the semantic rules cannot be applied.

Classes of models Suppose we have a theory T and let $\mathscr{C}$ denote its class of models. What is $\mathscr{C}$? Suppose T is the theory of groups. Any model of T is a group, and a group, as we have seen, is a structure of the form $\mathscr{G} = \langle G, * \rangle$. It is easy to realize,[22] as we have done before, $\mathscr{G}$ is a *set*, say in ZFC. But the collection of all groups (all models of the theory) does not constitute a set of ZFC. It is what mathematicians call a *proper class* and exist not in ZFC, but in some strong theory such as the system NBG (von Neumann, Bernays and Gödel, see Mendelson 1997, chap.4). In this system, all entities are *classes*, and those classed that are members of other classes are called *sets* and coincide with the sets of ZFC. The others are *proper classes*. So, in order to accept a phrase as van Fraassen's in our beginnings, we need to consider *all* the models of T and this of course requires us to be aware of the possibility of their existence. As we see, metamathematics matters, and for certain philosophical questions, should be not taken arbitrarily.

6. Note added in proof

I would like to say something more about the structure introduced in the definition (3.1). There, we said that to each quantum system $s \in S$ we associate a structure of that type, but this is not strictly correct. Well, it is partial. In order to justify the reasons, we follow Leslie Ballentine (1998), p.99. The wave function $\psi(\mathbf{x}, t)$ can be said to represent something propagating in what Ballentine calls "the ordinary space", namely, $\mathbb{R}^3$ (recall that the support of $\mathbb{E}^4$ is $\mathbb{R}^3 \times \mathbb{R}$) only for one-particle states, when the configuration space is isomorphic to the ordinary space. But, when we have multi-particle systems (suppose N of them), the same reasoning would conduce to N wave-functions interacting in ordinary space, something that cannot be accepted (Schrödinger's equation shows that this is not the case, as recalled by Ballentine). Really, the wave-function of the whole system is a function of space and time but in a large dimensional space (of dimension $3N$), the *configuration space*. So, we need to take some care in discussing whatever 'reality' we associate with the wave-function (by the way, this theme is still being debated in the philosophy of physics literature).

References

Ballentine, L. E. 1998. *Quantum Mechanics: A Modern Development*. Singapore: World Scientific Publishing Company.

Benioff, P. A. 1976a. Models of Zermelo Fraenkel set theory as carriers for the mathematics of physics, i. *Journal of Mathematical Physics* **17**: 618–28.

Benioff, P. A. 1976b. Models of Zermelo Fraenkel set theory as carriers for the mathematics of physics, ii. *Journal of Mathematical Physics* **17**: 629–40.

Bourbaki, N. 2006[1970]. *Théorie des Ensembles*. Berlin and Heidelberg: Springer-Verlag.

Button, T.; Walsh, S. 2018. *Philosophy and Model Theory*. Oxford: Oxford University Press.

da Costa, N. C. A.; Doria, F. A. 1992. Suppes predicates for classical physics. In J. Echeverria; A. Ibarra; T. Mormann (ed.) *The Space of Mathematics: Philosophical, Epistemological, and Historical Explorations*, pp.168–91. Berlin and New York: Walter de Gruyter.

da Costa, N. C. A.; Krause, D. 2020. Suppes predicates for classes of structures and the notion of transportability. In: A. Costa-Leite (ed.) *Abstract Consequence and Logics: Essays in Honor of Edelcio G. de Souza*, pp.79–98. Tributes 42. London: College Publications.

Dalla Chiara, M. L.; di Francia, G. T. 1981. *Le Teorie Fisiche: Un'Analisi Formale*. Torino: Boringhieri.

Dalla Chiara, M. L.; di Francia, G. T. 1993. Individuals, kinds and names in physics. In: G. Corsi; M. L. Dalla Chiara; G. C. Ghirardi (ed.) *Bridging the Gap: Philosophy, Mathematics, and Physics*, pp.261–84. Boston Studies in the Philosophy of Science 140. Dordrecht: Kluwer.

Enderton, H. B. 1977. *Elements of Set Theory*. New York, San Francisco, London: Academic Press.

Fraenkel, A. A.; Bar-Hillel, Y.; Levy, A. 1973. *Foundations of Set Theory*. 2nd Edition. Studies in Logic and the Foundations of Mathematics 67. Amsterdam: Elsevier.

Oliveira, A. J. F. 2010. *Lógica e Aritmética: Uma Introdução à Lógica, Matemática e Computacional*. 3rd Edition (revised and enlarged). Lisboa: Gradiva, Lisboa.

French, S.; Krause, D. 2006. *Identity in Physics: A Historical, Philosophical, and Formal Analysis*. Oxford: Oxford University Press.

Jech, T. J. 1977. About the axiom of choice. In: J. Barwise (ed.) *Handbook of Mathematical Logic. Studies in Logic and the Foundations of Mathematics 90*, pp.345–70. Amsterdam: North Holland.

Krause, D. 1990. Non-Reflexivity, Indistinguishability, and Weyl's Aggregates (in Portuguese). *PhD thesis, Faculty of Philosophy, University of São Paulo, August*.

Krause, D. 1992. On a quasi-set theory. *Notre Dame Journal of Formal Logic* **33**(3): 402–11.

Krause, D.; Arenhart, J. R. B. 2017. *The Logical Foundations of Scientific Structures: Languages, Structures, and Models*. Routledge Studies in Philosophy of Mathematics and Physics. New York and London: Routledge.

Krause, D.; Arenhart, J. R. B.; Bueno, O. 2022. The non-individuals interpretation of quantum mechanics. In: O. Freire Jr. (ed.) *The Oxford Handbook of the History of Quantum Interpretation*, Chapter 46, pp.1135–54. Oxford: Oxford University Press.

MacLane, S.; Birkhoff, G. 1968. *Algebra*. 3rd printing edition. New York and London: The Macmillan Co. and Collier-Macmillan.

Maitland-Wright, J. D. 1973. All operators on a Hilbert space are bounded. *Bulletin of the American Mathematical Society* **79**(6): 1247–51.

Mashaal, M. 2006. *Bourbaki: A Secret Society of Mathematicians*. Providence: American Mathematical Society.

Mendelson, E. 1997. *Introduction to Mathematical Logic*. 4th edition. London: Chapman and Hall.

Penrose, R. 2005. *The Road to Reality: A Complete Guide to the Laws of the Universe*. New York: Vintage Books.

Pogorelov, A. V. 1987. *Geometry*. Moscow: Mir Publishers.

Redhead, M. 1987. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford: Clarendon Press.

Smith, P. 2021. *Gödel Without (Too Many) Tears*. Cambridge: Logic Matters.

Suppe, F. (ed.). 1977. *The Structure of Scientific Theories*. 2nd edition. Urbana and Chicago: University of Illinois Press.

Suppes, P. 1972. *Axiomatic Set Theory*. New York: Dover.

Suppes, P. 2002. *Representation and Invariance of Scientific Structures*. Stanford: Center for the Study of Language and Information, CSLI.

van Fraassen, B. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Notes

[1]Beyond 'point', 'line' and 'plan', Hilbert used three more relations as primitive, namely, *betweenness*, *lies on*, and *congruence* (between points and lines and between points and planes); see Pogorelov (1987).

[2]A set-theoretical predicate is a formula in the language of set theory that congregates the specific postulates of the theory; more on this below.

[3]Of course nothing impedes that someone wakes up on a certain day and decides to 'create' a theory from nothing, a purely formal 'theory', although in general, we have a motivation for doing that. Anyway, to consider this possibility, we would be in need of discussing the meaning of 'theory'.

[4]In Tarskian sense; see Mendelson (1997), chap.2.

[5]This is one of the most interesting and important results of axiomatization, namely, the noticing that in general there is more in the heavens than we noticed at the start. The typical example is first-order arithmetics, which has presented us with non-standard models; in these models, there are 'natural numbers' which are different from those of the above sequence (see Oliveira 2010).

[6]For more details, see Krause and Arenhart (2017).

[7]We speak of 'order-n' structures to avoid confusion with the order of languages since in first-order languages (such as that of ZFC) we can define higher-order structures ($n > 1$).

[8]See Enderton (1977), p.168.

[9]mRa is an abbreviation for $\langle m, a\rangle \in R$. See Enderton (1977), p.171.

[10]Notice that an element of $\mathscr{P}(G \times G \times G)$ is a collection (a set) of triples of elements of G. Intuitively, we can reason this way: from all triples $\langle a, b, c\rangle$ of elements of G, we select those where $c = a * b$ to compose $*$, that is, it is the result of the operation $*$ between a and b (in this order). For instance, for the addition of real numbers, we select from $\mathscr{P}(\mathbb{R}\times\mathbb{R}\times\mathbb{R})$ those triples $\langle a, b, c\rangle$ such that $c = a + b$.

[11]Here, '+' and '·' are the addition and multiplication of the elements of F, and should not be confounded with the addition of vectors and the multiplication of vectors by scalars (the elements of F); mathematics is "the art of giving the same name to different things", said Henri Poincaré.

[12]Roughly speaking, in a transportable formula there cannot be imposed restrictions whatever on the principal sets occurring in the formula. The formula must hold in all cases of substitutions. So, in the case of vector spaces, we should not use as the field a particular

case as the field $\mathscr{R}$ of the real numbers but, as we have done, a general field $\mathscr{F}$. The case of real vector spaces would become a particular case, a structure of *that* species (of (general) vector spaces). A field is characterized by an abstract structure of the following species: we have a non-empty domain F and two binary operations over F, namely '+' and '·', obeying well-known postulates (Maclane and Birkhoff 1968, p.133).

[13]The reason for this to be an abuse of notation is easy to explain. The operation of multiplication of a vector by a scalar (a complex number) is defined to be *multiplication to the left*, that is, it is a function from $\mathbb{C} \times \mathscr{H}$ to $\mathscr{H}$. If we want that this also represents a function from $\mathscr{H} \times \mathbb{C}$ to $\mathscr{H}$, we need to say it explicitly.

[14]These are operators such that there is some basis for the vector space so that the matrix representing the operator is a diagonal matrix.

[15]For details, see Redhead (1987), p.8.

[16]Below we shall question whether this collection can be considered as a *set* of ZFC.

[17]But see the Note Added in Proof at the end.

[18]For details, which do not interest us here, see Penrose (2005), chap.17.

[19]These are particular subsets of the real number line, suitable for the axiomatics to hold, but whose definition is not relevant for us here.

[20]This idea was also posed by Dalla Chiara and Toraldo di Francia (1993).

[21]A *bounded* operator T is a linear operator over the Hilbert space so that there exists a natural number N such that for all vectors α, we have that $\|T(\alpha)\| \leq N\|\alpha\|$. If T is not bounded, it is unbounded.

[22]It is enough to pay attention to the way the structure was constructed within ZFC.

Acknowledgments

I would like to thank the anonymous referee for correcting so many typos in the previous version. Those that remain are of course my fault. I would like to thank also the editor of *Principia* for letting me add the above Note Added in Proof. Partially supported by CNPq.