ON CONSTRUCTIVE AXIOMATIC METHOD

ANDREI RODIN

Abstract. Formal axiomatic method popularized by Hilbert and recently defended by Hintikka (in its semantic version) is not fully adequate to the recent practice of axiomatizing mathematical theories. The axiomatic architecture of (elementary) Topos theory and Homotopy type theory do not fit the pattern of formal axiomatic theory in the standard sense of the word. However these theories fall under a more traditional and more general notion of axiomatic theory, which Hilbert calls constructive. A modern version of constructive axiomatic method can be more suitable for building physical and some other scientific theories than the standard formal axiomatic method.

1. Introduction

The modern notion of axiomatic method stems from works in foundations of mathematics starting in the late 19th century, most prominently from Hilbert’s Foundations of Geometry first published in 1899 [15], [17]. Recently this method has been systematically presented (with some special features, which I discuss below) by Hintikka [22]. In this paper I argue that this notion of axiomatic method is not fully adequate to the current mathematical practice. As a remedy I describe a different and in a sense more general version of axiomatic method that covers some important recent instances of axiomatic thinking as well as some older instances such as Euclid’s Elements. Since I don’t want to consider my argument as merely pragmatic I attempt to ground it also epistemologically through a critique of Hintikka’s paper. In this context I provide some further arguments in favor of the view

Date: August 14, 2014.

1The work is supported by Russian State Foundations for Humanities, research grant number 13-03-00384
that this more general axiomatic method can be more useful in physics and other sciences than the received version of this method.

The method that I have in mind has been already well known to Hilbert, who calls it constructive or genetic interchangeably ([21], p. 1a), and discussed in a later literature [4], [6], [44], [52]. My contribution consists in showing how the constructive method operates in the axiomatic Topos theory [42] (in a somewhat implicit form) and the recent Homotopy Type theory (HoTT) along with the closely related project of Univalent Foundations of mathematics [7] - and providing on this basis some new epistemological arguments in favor of this method.

The rest of the paper is organized as follows. After making a short exposition on the constructive axiomatic method after Hilbert and Bernays [20], [21] in Section 2, I demonstrate this notion using the example of Euclid’s Elements, Book 1 (Section 3). Although the main purpose of this paper is not historical, this historical example is very useful because it helps me to disambiguate the overloaded term “constructive” and illustrate my arguments with the familiar elementary geometry. In Sections 4 I recall some basic facts concerning Hilbert’s formal axiomatic approach and in Section 5 I introduce the notion of Curry-Howard correspondence, which I need in what follows. In Sections 6, 7 I treat two modern examples of axiomatic theories: Topos theory (Section 5) and HoTT (Section 6). In the concluding Section 8 I provide an epistemological discussion, which includes a critique of Hintikka’s view on axiomatic method [22].

2. CONSTRUCTIVE METHOD VERSUS FORMAL METHOD

In the Introduction to their [20], Eng. tr. [21] Hilbert and Bernays specify their intended notion of axiomatic method as follows:

The term axiomatic will be used partly in a broader and partly in a narrower sense. We will call the development of a theory axiomatic in the

2Piaget calls the genetic method operational.
broadest sense if the basic notions and presuppositions are stated first, and then the further content of the theory is logically derived with the help of definitions and proofs. In this sense, Euclid provided an axiomatic grounding for geometry, Newton for mechanics, and Clausius for thermodynamics. For axiomatics in the narrowest sense, the existential form comes in as an additional factor. This marks the difference between the axiomatic method [in the narrow sense] and the constructive or genetic method of grounding a theory. While the constructive method introduces the objects of a theory only as a genus of things, an axiomatic theory refers to a fixed system of things [..] There is the assumption that the domain of individuals is given as a whole. Except for the trivial cases where the theory deals only with a finite and fixed set of things, this is an idealizing assumption that properly augments the assumptions formulated in the axioms. We will call this sharpened form of axiomatics (where the subject matter is ignored and the existential form comes in) “formal axiomatics” for short. ([21], p. 1a)

The reader may notice in the above quote a slight terminological confusion, which in my view is not quite innocent but reflects a real conceptual problem. The authors first say that the constructive axiomatic method is co-extensive with the axiomatic method in the “broadest” sense of the word. When the authors call the existential form (which is a characteristic property of the formal method) an “additional factor” this adds to the impression that they consider the formal method as a special case of the constructive method. However further in the same passage they compare the constructive method with the axiomatic method as if the extensions of these two concepts were disjoint. One may try to explain this away by taking (in the latter case) “constructive” in the sense of “purely constructive”, i.e., “constructive non-formal” and “axiomatic” in the sense of “formal axiomatic” (assuming that the authors simply drop these qualifications). But this doesn’t solve the whole

\[\text{3The distinction “genetic versus axiomatic” is also found in much earlier Hilbert’s paper [16], (Eng. tr. [19]) and used in later discussions [4], [6], [44], [52]. At least in some cases this imprecise terminology leads to conceptual confusions. For example, Demidov in [6] describes Euclid’s geometrical theory (along with with}
problem because it remains unclear what (if any) is the constructive (in the broad sense of “constructive”) content of formal axiomatics. In what follows I shall try to show that this is an important conceptual question about axiomatic method but not just a question about words. I shall justify the view according to which the constructive method is indeed more general than the formal method and in Section 8 give an exact answer to the above question.

Let me now make my own terminological precautions. Hereafter I use terms “constructive” and “formal” in the same sense in which Hilbert and Bernays use them in the above quote. Namely, I count as constructive any method which involves some systematic procedure for “introducing the objects of a theory” (without trying to delimit the class of allowable procedures of this sort), and I distinguish the formal method by Hilbert and Bernays’ notion of “existential form”, which I explain shortly. When the risk of ambiguity becomes particularly high I remind the reader about these precautions.

3. PROBLEMS VERSUS THEOREMS

Let me now illustrate how a constructive theory “introduces the objects” at the example of a fragment of the geometrical theory presented in Euclid’s Elements, Book 1 [9], (Eng. tr. [8]). The same example will help us to see where the “existential form” comes from. Euclid’s theory is based on 5 Axioms (aka Common Notions), 5 Postulates and 23 Definitions. Even if there important differences between what Euclid call Common Notions and Definitions and what we usually call axioms and definitions today I shall not discuss them here and focus on Postulates. The first three Postulates read (verbatim after [8]) as follows:

[P1:] to draw a straight-line from any point to any point.

[P2:] to produce a finite straight-line continuously in a straight-line.

Newton’s theory of fluxions) as axiomatic while Euclid’s arithmetical theory he describes as genetic. Thus this author uses the term “axiomatic” in a sense that does not coincide either with Hilbert and Bernays’ “broader” sense nor with their “narrowest” sense of this term. The author leaves it to the reader to guess the exact sense in which he uses word “axiomatic”.

[P3:] to draw a circle with any center and radius.

As they stand the three Postulates are not propositions and admit no truth-values. Hence they cannot be axioms in the usual modern sense of the word. In particular, the Postulates cannot be used as premises in logical inferences - if by logical inference one understands an operation that takes some propositions (premises) as its input and produces some other proposition or propositions (conclusion) as its output.

In fact Postulates 1-3 are themselves elementary operations, which take certain geometrical objects as input and produce some other geometrical objects as output. The table below specifies inputs (operands) and outputs (results) for P1-3:

<table>
<thead>
<tr>
<th>operation</th>
<th>input</th>
<th>output</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>two (different) points</td>
<td>straight segment</td>
</tr>
<tr>
<td>P2</td>
<td>straight segment</td>
<td>(extended) straight segment</td>
</tr>
<tr>
<td>P3</td>
<td>straight segment and one of its endpoints</td>
<td>circle</td>
</tr>
</tbody>
</table>

The three operations can be composed as follows: the output of P1 is used as input for P2 and P3; thus one gets complex operations, which can be denoted P1◦P2 and P1◦P3. The resulting system of operations is the core of a larger system of operations “by ruler and compass” assumed by Euclid in his *Elements*. Using this system of operations Euclid “introduces objects” of his theory. Such an introduction is systematic in the sense that it does not reduce to a simple act of stipulation: it is a procedure, which involves certain elementary operations (including P1-P3) and complex operations obtained through the composition of the elementary operation. As soon as the term *deduction* is understood liberally as a theoretical procedure, which generates some fragments of a given theory from the first principles of this theory, one can say that the object-building in Euclid’s is

---

4. This fact has been stressed by A. Szabo, see [54], p. 230.

5. In addition to P1-3 and other Postulates Euclid assumes implicitly certain other operations like the construction of intersection point of lines in appropriate positions. But this incompleteness has no bearing on my argument.
deductive. As we shall now see in Euclid’s theory the “object-oriented” deduction described above goes along with the more familiar propositional deduction.

Postulates and Axioms of are followed by so-called Propositions. This commonly used title is not found in Euclid’s original text where things called by later editors “Propositions” are simply numbered but not given any special title [8]. From Proclus’ Commentary [46], (Eng. tr. [47]) written in the 5th century A.D. we learn about the tradition dating back to Euclid’s own times (and in fact even to earlier times) of distinguishing between the two sorts of “Propositions” called Theorems and Problems. Euclid’s Theorems by and large are theorems in the modern sense of the word: propositions followed by proofs. But Problems are something very different: they are complex operations (or, if one prefers, complex constructions) built with elementary operations. Like Postulates Problems admit no truth-values.

As an example of Problem consider the (initial fragment of) Proposition 1, Book 1:

To construct an equilateral triangle on a given finite straight-line.

which is followed (skipping irrelevant details) by

- construction (complex operation) \( C \) based on P1-3 and some tacit primitive operations;
- proof \( P \) that \( C \) is equilateral triangle.

What I have said earlier about Postulates and Problems, on the one hand, and Axioms and Theorems, on the other hand, can make one imagine that Euclid’s theory splits into two independent parts one of which consists of constructions from Postulates while the other consists of proofs from Axioms. The above example shows that such a split does not occur. Euclid needs proof \( P \) in this Problem because construction \( C \) does not automatically speak for itself, i.e., does not produce a proposition saying that it has the wanted property \(^6\).

\(^6\)How in this case Euclid manages to get out of the mute construction any proposition at all? Since the circle concept makes part of P3 and P3 makes part of \( C \) Euclid applies his definition of circle and so gets propositions of the form “radii \( a, b \) of circle \( c \) are equal”. He uses these propositions as premises, then applies the First Axiom and thus gets the wanted proof.
More generally Problems and Theorems (resp. constructions and proofs) are related as follows:

- (solutions of) non-trivial Problems require propositional proofs as in the above example;
- (proofs of) non-trivial Theorems require constructions (conventionally called in such contexts “auxiliary”).

This explains why the logical order of Theorems and the genetic order of Problems in Euclid’s theory are closely intertwined and form a joint deductive order. For a more detailed analysis of this structure see [48], ch.2. In Section 5 below I explain how the Curry-Howard correspondence supports a similar structure, which combines a logical deduction with a genetic construction.

All Euclid’s Postulates and initial fragments (i.e., bare formulations) of Problems can be easily paraphrased into propositions. This can be done at least in two different ways which I shall call modal paraphrasing and existential paraphrasing. The following paraphrases of P1 are self-explanatory:

P1m (modal): Given two (different) points it is always possible to produce a straight segment from one given point to the other given point.

P1e (existential): Given two (different) points there exists a straight segment having these given points as its endpoint.

P1e instantiates what Hilbert and Bernays call the “existential form” used in the formal axiomatic method. The key logical feature of this paraphrase (which it shares with P1m) is the reduction of Euclid’s non-propositional Postulates and Problems to certain propositions (in case of P1e - to existential propositions). Such a reduction may look trivial and even purely linguistic but in fact it is not because none of the two ways of paraphrasing is sufficient for translating Euclid’s theory into a propositional form, i.e., into a theory consisting of axioms and theorems derived from the axioms according to certain fixed rules of logical inference. The problem is, of course, that the straightforward propositional paraphrasing does not translate proofs and constructions coherently, so in order to provide a reasonable reconstruction of Euclid’s theory in the above form one needs a lot of further
efforts [23]. For further references I shall call a procedure, which aims at replacement of all non-propositional content of a given theory by some suitable propositional content, the *propositional reduction* of this theory.

Thanks to Proclus we know that the idea of translating all Problems into Theorems dates back to Euclid’s own times and even to earlier times. From the same source we also know about the competing idea of translating all Theorems into Problems; this latter idea Proclus attributes to “the school of Menaechmus” 7.

Both proposals make equally strong echoes of some more recent approaches. In particular Menaechmus’ constructivism makes an echo of *calculus of problems* proposed by Kolmogorov in 1932 [26]. In the present discussion I want to consider this old controversy critically and pay a special attention to constructive mathematical thinking. This is a

---

7Here is the relevant passage:

Some of the ancients, however, such as the followers of Speusippus and Amphinomus, insisted on calling all propositions “theorems”, considering “theorems” to be a more appropriate designation than “problems” for the objects of the theoretical sciences, especially since these sciences deal with eternal things. There is no coming to be among eternals, and hence a problem has no place there, proposing as it does to bring into being or to make something not previously existing - such as to construct an equilateral triangle [...]. Thus it is better, according to them, to say that all these objects exist and that we look on our construction of them not as making, but as understanding them [...]. Others, on the contrary, such as the mathematicians of the school of Menaechmus, thought it correct to say that all inquiries are problems but that problems are twofold in character: sometimes their aim is to provide something sought for, and at other times to see, with respect to determinate object, what or of what sort it is, or what quality it has, or what relations it bears to something else.” ([47], p. 63-64).

Speusippus [408 - 338/9 B.C.] is Plato’s nephew who inherited his Academy after Plato’s death. Amphinomus is otherwise unknown (except few other references found in the same Commentary) Speusippus’ contemporary. Menaechmus [380 - 320 B.C.] is mathematician of Plato’s circle. The dates of these people suggest that the these ideas had been already around for quite a while before Euclid completed his *Elements* at some point between 300 and 265 B.C.
sufficient reason for not taking the idea of propositional reduction (of mathematical and other theories) for granted.

4. Hilbert’s Views on Axiomatic Method

Since presently there exists an extensive literature, which analyses Hilbert’s work on axiomatic method in historical and theoretical perspectives, in this Section I shall not try to say anything new about Hilbert but only recall some basic features of his axiomatic approach, which I later use for contrasting against them some new features of more recent axiomatic approaches. At the early stage of his life-long work on axiomatic foundation of mathematics (1893-1894) Hilbert describes his notion of well-founded axiomatic theory as follows:

Our theory furnishes only the schema of concepts connected to each other through the unalterable laws of logic. It is left to human reason how it wants to apply this schema to appearance, how it wants to fill it with material. This can happen in manifold ways. But whenever the axioms are satisfied, then the theorems must apply too. ([14], p. 104)

The two key features of this notion of theory (which are by and large realized in Hilbert’s Foundations of 1899 [15]) are (i) its schematic character and (ii) its logical grounding. Let me first focus on (ii). In the above quote Hilbert apparently refers to the “unalterable laws of logic” as something definite and somehow known. A weaker assumption which we may attribute to Hilbert and which still allows us to make sense of his words is that the laws of logic are known better than mathematical (and perhaps some other) theories, which can be construed in a schematic form and grounded with these logical laws. This latter assumption was moreover plausible at the time when geometry had already ramified into its traditional Euclidean and multiple non-Euclidean branches but a similar development in logic, which began later in the 20th century, did not yet happen.
The fixity of logic is important for understanding the schematic character of Hilbert’s axiomatic theories (i). Axioms and theorems of non-interpreted formal theory are propositional schemes, which admit truth values and thus turn into propositions through an interpretation. An interpretation amounts to assigning to certain terms like “point”, “straight line”, etc. certain semantic values, which can be borrowed from another (usually informal) mathematical theory or from some extra-theoretical sources like intuition and experience. However this game of multiple interpretations does not concern all terms of a given theory. Some terms, namely logical terms, have a fixed meaning, which (at least in the early versions of Hilbert’s axiomatic approach) is supposed to be self-evident. The different treatment of logical and non-logical terms reflects the epistemological assumption according to which logical concepts (such as implication) are generally better known than geometrical or any other non-logical concepts.

In the *Foundations* of 1899 and other Hilbert’s early axiomatic theories the “laws of logic” are taken for granted but not specified explicitly and precisely. Hilbert addresses this problem in 1917 saying that “it appears necessary to axiomatize logic itself” ([18] p. 1113). Hintikka quite rightly, in my view, stresses the fact that between axiomatizing geometry (or another non-logical theory) and axiomatizing logic there is no continuity. Indeed, since the axiomatization in the above sense involves an application of logic to the given subject-matter (say, geometry), the axiomatization of logic requires an application of logic to itself (at least if we are always talking about one and the same logic as certainly does Hilbert), which is, generally, problematic ([22], p. 80). Hintikka further argues that a recursive enumeration of logical truths, called by Hilbert an axiomatization of logic, is called so improperly because such a procedure doesn’t allow for studying various models of logic in anything like the same way in which one studies various models of any other formal theory. As a remedy Hintikka uses his original system of IF-logic, which allows for self-application. In what follows (Sections 6, 7) I describe a different solution, which blurs the distinction between logical and non-logical terms by allowing logical terms to have further non-logical interpretations.
Hilbert’s mature axiomatic method reinforced by symbolic logic [20] has some important features, which are wholly absent in Hilbert’s early conception of this method described above. While in the early version a non-interpreted axiomatic theory is understood as a “scheme of concepts” devoid of any intuitive content the later symbolic version includes an additional assumption according to which this abstract scheme comes with its proper concrete representation, namely the symbolic representation. The symbolic representation involves a special sort of intuition, which Hilbert calls the “logico-combinatorial intuition” ([13], p. 179). A mathematical study of symbolic calculi (which include logical calculi proper and symbolic representations of formal theories based on these calculi) Hilbert isolates into a special area of mathematics, which he calls metamathematics. Hilbert perfectly realizes that treating the metamathematics with the same formal axiomatic method leads to a hopeless infinite regress. So his foundational project at this point becomes different and in certain respects more modest (albeit in some other respects more radical) than earlier: now he aims at isolating a limited area of elementary (and as he really hoped - only finitary) mathematics developed constructively and then treat the rest of mathematics on this constructive basis using appropriate non-constructive “idealizing existence assumptions” 8. Thus in the advanced symbolic setting the formal axiomatic method is no longer seen by Hilbert as self-sustained: it needs a support of constructive methods operating at the metatheoretical level. This is perhaps a reason why Hintikka [22] elaborates rather on the early version of Hilbert’s method. In Section 8 I argue that this does not solve the problem because in this case the formal method requires a constructive support anyway.

8“When we now approach the task of such an impossibility proof [= proof of consistency], we have to be aware of the fact that we cannot again execute this proof with the method of axiomatic-existential inference. Rather, we may only apply modes of inference that are free from idealizing existence assumptions.” ([21], p. 19)
5. Curry-Howard Correspondence and Cartesian Closed Categories

The present Section is a preliminary to the following two Sections 6, 7 where I treat modern examples of axiomatic theories. In the nutshell the idea of Curry-Howard correspondence is given in Kolmogorov’s 1932 paper [26] where the author establishes that his newly proposed calculus of problems has exactly the same structure as the intuitionistic propositional calculus published in 1930 by Heyting \(^9\). It turns out that this correspondence is extendible onto a large class of symbolic calculi including those, which have been developed independently and apparently for very different purposes. So the were established a number of correspondences (i.e., of more and less precise isomorphisms) between (proof-related) logical calculi (propositional, first-order or higher-order), on the one hand, and computational calculi (the simply-typed lambda calculus, type systems with dependent types, polymorphic type systems), on the other hand. For a substantial introduction (which however does not fully reflect the present state of the art) see [53].

Even if in the textbooks the Curry-Howard correspondence is often described as unexpected and even “amazing” ([53], p. 5), it is not a unexplained mathematical phenomenon but rather a case of genuine convergence of different approaches in foundations of mathematics. This convergence leads to the so called “proofs-as-programs and propositions-as-types” paradigm in logic and Computer science, which can be called constructive in the broad sense specified above. When Hilbert and Bernays distinguish between the constructive and the formal versions of the axiomatic method (Section 2 above) they firmly assume that propositions and objects of some given (interpreted) theory belong to distinct domains of things. They don’t treat this distinction formally but simply take it for granted. Within the propositions-as-types paradigm this distinction is made formally and explicitly: propositions are represented (not as objects of sort but) types, namely types of proofs. Recall that

---

\(^9\)I am not making any priority claim for Kolmogorov here but simply use his 1932 article as a convenient reference, which states epistemological ideas behind this development more clearly than do many later more technical studies. A relevant historical material can be found in [3] and [51] but a more focused historical study on the idea of Curry-Howard correspondence still waits to be written.
in the (advanced symbolic) Hilbertian setting propositions and proofs of formal theory \( T \) (or, more precisely, their syntactic expressions), are objects of the corresponding metatheory \( M_T \). The propositions-as-types paradigm puts proofs and other objects of given theory \( T \) at the same theoretical level, so proofs here are objects of sort among other objects of different sorts belonging to the \textit{same} theory. Thus this paradigm fits the constructive notion of axiomatic method according to which building proofs is a special case of building objects in general.

In 1963 Lawvere observed that “cartesian closed categories serve as a common abstraction of type theory and propositional logic” ([36], p. 1), [34]; in other words Cartesian closed categories (CCC) capture the structure behind the Curry-Howard correspondence. In 1968-1972 this observation has been developed by Lambek in a series of papers [27], [28], [29] into what is now known under the name of Categorical Logic; for a systematic exposition of these results see [37]. The three way correspondence between (i) logical calculi (propositions), (ii) computational calculi (types) and (iii) (objects of) CCC and some other appropriate category) is sometimes called in the literature the \textit{Curry-Howard-Lambek} correspondence.

Let me now tell a bit more about a relevant part of Lawvere’s work that touches upon the issue of axiomatic method directly. Lawvere discovered CCC during his work on alternative axiomatization of Set theory \(^{10}\). His idea was that the non-logical primitive of standard axiomatic Set theories like \( ZF \), namely the binary membership relation \( \in \), was wrongly chosen. Lawvere suggested to use instead the notion of function and the binary operation (i.e., ternary relation) of composition of functions. This choice has been motivated by the idea to axiomatize Set theory in the form of a theory of \textit{category} of sets \( \text{Set} \). The resulting axiomatic theory is known today under the acronym ETCS (Elementary Theory of Category of Sets) [30], [35], [42]. This proposal may appear radical to one who has habituated oneself to \( ZF \) and its likes but as it stands this proposal does not require any modification of the standard formal axiomatic method; the suggested modification concerns

\(^{10}\)Lawvere also had other important motivations for introducing CCC, see [31].
only the choice of non-logical primitives. The unusual choice of primitives allowed Lawvere to see that the condition of being CCC makes part of the wanted axiomatic description of $\text{Set}$. In this context the further observation that CCC provides a structural description of Curry-Howard correspondence looks like an unexpected bonus.

Let us see more precisely what happens here. The standard axiomatic approach requires to fix logic first and then use it for sorting out intuitive ideas about sets, numbers, spaces and whatnot. Trying to use the fixed logical rules for axiomatic reasoning about the chosen subject-matter one may discover that these logical rules are not quite appropriate for the task. Then one may fix these rules and try again. But unless one is axiomatizing logic (see Section 4 above) one still always expects to get as the outcome of axiomatization some non-logical theory, not a new axiomatic theory of logic. However in the case of ETCS something similar happens: it turns out that $\text{Set}$ is equipped with its proper internal logically-related structure, namely CCC. This structure is not simply transported from the background logic but emerges as a specific feature of $\text{Set}$ among other categories. This fact does not exclude the possibility of building ETCS as standard formal axiomatic theory [42] but it nevertheless suggests a reconsidering of place and role of logic in this theory. In the next Section we shall see how the notion of internal logic is made precise in Topos theory and how it helps Lawvere to axiomatize this theory.

6. Topos theory

The concept of topos first appeared in Algebraic Geometry in the circle of Alexandre Grothendieck around 1960 as a far-reaching generalization of the standard concept of topological space and didn’t have any special relevance to logic. In his seminal paper [33] Lawvere provided an axiomatic definition of topos called today the definition of elementary topos $^{11}$. Lawvere’s axiomatization of Topos theory is a great success story of

$^{11}$The title “elementary” reflects the fact that Lawvere’s definition (unlike Grothendieck’s original definition) straightforwardly translates into the standard first-order formal language [42]. According to this definition an (elementary) topos $T$ is CCC equipped with a subobject classifier, which plays in a general
the axiomatic method in the 20th century mathematics. It provides a very significant simplification of Grothendieck’s ideas making them available for anyone who wants to study the subject. Like ETCS the axiomatic theory of elementary topos does not bring by itself any new notion of axiomatic theory. However any systematic exposition of topos theory contains a chapter on internal logic of topos. In standard textbooks [42], [38] this notion is introduced as an extra feature after the basic topos construction is already done. The construction of internal logic has a syntactic part (Mitchel-Bénabou language) and a formal semantic part, which interprets the syntax of this language in terms of constructions available in the base topos (Kripke-Joyal semantics aka natural semantics). By semantics in that case one should understand something else than the usual interpretation of non-logical terms of the given formal system like in the case of models of ZFC. Kripke-Joyal semantics assigns to symbols and syntactic expressions, which have an intuitive logical meaning (logical connectives, quantifiers, truth-values, etc.), explicit semantic values, which otherwise can be called geometrical (if the given topos is seen as a generalized space). This is not something wholly unprecedented in the history of the 20th logic: think, for example, of Tarski’s topological of (Classical and Intuitionistic) propositional logic [55]. However when such an interpretation is considered in the context of Hilbert’s ideas about axiomatic method it looks very strange. Giving logical symbols and expressions non-logical interpretations blurs the distinction between logical and non-logical elements of syntax and thus undermines the basic Hilbert’s idea of clarifying non-logical contents by means of logic. However as long as internal logic $L_T$ of topos $T$ is construed as an extra gadget associated with $T$ one doesn’t need to revise the axiomatic method by which one introduces $T$ itself. One may also use the standard axiomatic method for building with $L_T$ (namely, taking it as a background logic) some further theories “internally” in $T$.

The concept of elementary topos is slightly more general than that of Grothendieck topos: there are elementary toposes, which are not Grothendieck toposes.

does the role similar to that played in $Set$ (which also qualifies as a topos in the sense of Lawvere’s definition) by the two-element set. The two-point set classifies subsets of a given set $S$ in the sense that if one asks whether a given element $p \in S$ belongs to subset $U \subseteq S$ there are just two possible answers: yes and no.
I shall not explore epistemological implications of the idea of “doing mathematics internally in a given topos” because I believe that notion of internal logic also plays in Topos theory a different and more fundamental role. In the beginning of his [33] (where the axioms for elementary topos first appeared in the press) Lawvere writes:

[A] Grothendieck “topology” appears most naturally as a modal operator, of the nature “it is locally the case that”, the usual logical operators, such as $\forall$, $\exists$, $\Rightarrow$ have natural analogues which apply to families of geometrical objects rather than to propositional functions, and an important technique is to lift constructions first understood for “the” category $\mathbf{S}$ of abstract sets to an arbitrary topos. [...] In a sense logic is a special case of geometry. ( [33], p. 329)

The observed analogy serves Lawvere not just for providing a given topos with an extra construction called internal logic. It serves him first of all for formulating axioms of Topos theory! Namely, Lawvere observes that the internal logic of general topos and the internal logic of $\mathbf{Set}$ share the same CCC structure and thus the wanted axiomatic Topos theory is obtained through a simple generalization of ETCS. I submit that this analysis that allowed Lawvere to axiomatize Topos theory qualifies as genuine logical analysis - even if this sort of analysis is very unlike Hilbert’s logical analysis used in his Foundations of Geometry [15]. The internal logical structure of topos is the logical structure, which allows for a logically transparent axiomatic presentation of topos concept. 12.

One may argue that the feature of Lawvere’s approach to axiomatizing Topos theory, which I try to highlight here, may matter when we are talking about the way and the context in which this axiomatization has been first achieved but not when we are talking about the final result. This depends on what counts as final result. I can see that Lawvere’s

12As Lawvere makes it explicit in this and some other of his works he uses a Hegelian approach to logic as his philosophical motivation. In this paper I attempt to explain what is going on here in my own words without referring to Hegel. For an analysis of the impact of Hegel’s philosophy on Lawvere’s mathematical work see [48], Section 5.8.
theory of elementary topos can be presented as a standard axiomatic theory à la Hilbert [42]. However I doubt that such a presentation reflects the relationships between logic and geometry in this case properly. If the internal logic of a given topos is seen as its true logical foundation (rather than an extra feature) then Lawvere’s theory qualifies as broadly constructive because this logic regulates the construction of topos directly (namely, through its natural geometrical semantics) rather than indirectly through certain further “idealizing existence assumptions” à la Hilbert. Since CCC makes part of the topos structure, the above remarks about the constructive character of Curry-Howard-Lambek correspondence (Section 5) also apply to Topos theory 13. A constructive theory of topos remains a work in progress. However in the next Section I describe a more recent theory that already uses a constructive axiomatic architecture officially.

7. Homotopy Type theory and Univalent Foundations

Homotopy Type theory (HoTT) is a recently emerged field of mathematical research 14, which has a special relevance to philosophy and logic because it serves as a basis for a new tentative axiomatic foundations of mathematics called the Univalent Foundations (UF). The most comprehensive exposition available to the date is [7]. In this paper I do not attempt to review UF systematically but only describe a special character of its axiomatic architecture.

HoTT emerged through a synthesis of two lines of research, which earlier seemed to be quite unrelated: geometrical Homotopy theory and logical Type theory. The key idea is that of modeling types (including the type of propositions) and terms (including proofs) in

13For a discussion on constructive aspects of Category theory and Topos theory see [43].
14To the present date it’s official age is 8 years, see [7] p.4. However HoTT has a longer pre-history. The earliest direct hint I know is in Lawvere’s 1970 paper [32] where the author writes: “For deductions over X, one may take provable entailments or one may take suitable “homotopy classes” of deductions in the usual sense. One can write down an inductive definition of the “homotopy” relation, but the author does not understand well what results.” (pp. 34). Other milestones in this pre-history are [12] (on the geometrical side) and [24], [25] (on the type-theoretic side).
Type theory by spaces and their points in Homotopy Theory. Beware that along with basic spaces Homotopy theory also considers *path spaces* where “points” are paths in the basic spaces, spaces of “paths between paths” called *homotopies*, spaces of “paths between paths between paths” and so one. All these higher-order spaces are also used for interpreting types.

Like in the case of Topos theory in HoTT geometry and logic are glued together with some category-theoretic concepts. The central categorical concept used in HoTT is that of \(\omega\)−*groupoid*. In the standard category theory a groupoid is defined as a category where all morphisms between objects are reversible, i.e., are isomorphisms. \(\omega\)-groupoid is a higher-categorical generalization of this concept (called in this context 1-groupoid) where usual morphisms are equipped by morphisms between morphisms (called 2-morphisms), further morphisms between 2-morphisms (called 3-morphisms) and so on *ad infinitum* or more precisely up to the first infinite ordinal \(\omega\). An analogy between spaces equipped with paths between points, homotopies between those paths, etc. is straightforward. It allows for mixing the geometrical and the categorical languages and talk interchangeably, e.g., about “spaces of paths”, “groupoids of paths” and “groupoids” *simpliciter*. The “Homotopy theory construed categorically” and the “theory of \(\omega\)-groupoids construed homotopically" turn to be one and the same thing (provided both concepts are adjusted appropriately [12]).

The axiomatic HoTT uses resources of Type theory, namely the Constructive (aka Intuitionistic) Type theory with depended types due to Martin-Löf [41] (MLTT), for turning the tables at this point. The notion of \(\omega\)-groupoid aka space aka homotopy type is taken as primitive while the notions of proposition, set, (one-dimensional) groupoid, category, etc. are construed as derived notions with MLTT. For this end types (and, in particular, propositions) and terms (and, in particular, proofs) in MLTT are interpreted, correspondingly, as \(\omega\)-groupoids (aka spaces) and points of these spaces (which, as I have already explained, can be paths and “higher paths” of other spaces). I skip further details. The obtained interpretation of MLTT in the categorically construed Homotopy theory directly translates
all constructions in MLTT into geometrical constructions. Then one may consider some additional axioms on the top of MLTT such as the Axiom of Univalence (AU), which gives its name to the Univalent Foundations. Although AU has been first motivated geometrically, it also has a logical meaning being a generalized version of Church Extensionality Axiom for propositions \(^{15}\).

The fact that MLTT admits a geometrical interpretation and thus allows for a further geometrically motivated development, is a great mathematical discovery, which I cannot, of course, fully elucidate here. However the above short description of the idea of HoTT is sufficient for making my point about the changing notion of axiomatic method. In this paper I consider HoTT (with or without AU) only as an axiomatization of modern Homotopy theory. The idea of UF according to which HoTT can be used as a basis for developing the rest of mathematics has no bearing on my argument (but suggests that the special character of axiomatic approach in HoTT may be of general significance for logic and mathematics).

Let us now see how HoTT looks like against the standard notion of axiomatic theory stemming from Hilbert (Section 4). When one says that (the appropriately construed) Homotopy theory provides a model of MLTT, one uses the word “model” not in the same sense in which one usually talks, say, about various models of Lobachevskian geometry (like Poincaré model, Beltramy-Klein model, etc.). Interpreting a given formal theory of Lobachevskian geometry \((T)\) amounts to giving all the non-logical terms and expressions of this theory (such as “point”, “line, etc.”) some semantic values borrowed from another mathematical theory or elsewhere. The logical terms and expressions of \(T\) don’t take part in this semantic game. But in the case of HoTT (like in the case of Kripke-Joyal semantics for topos logic) all symbols and well-formed syntactic expressions, and all rules of MLLT on

\(^{15}\)Church Extensionality Axiom states that logically equivalent propositions are equal, in symbols \((\phi \rightarrow \psi) \land (\psi \rightarrow \phi) \rightarrow (\phi \equiv_p \psi)\) where the index \(p\) expresses the fact that the equality is typed and so \(\equiv_p\) is the equality of propositions. See abstract and video of talk given by T. Coquand at Bourbaki Seminar in Paris, June 21, 2014, available at www.bourbaki.cns.fr/seminaires/2014/Progruin14.html
the top of their intended logical and computational interpretation admit also a geometrical interpretation - and this is precisely what is called “model” in this case. It is clear Hilbert’s general understanding of relationships between geometry and logic in his *Foundations* of 1899 [15] in case of HoTT doesn’t apply. This means the standard notion of axiomatic theory does not apply to HoTT either.

At the same time HoTT (provided with the $\omega$-groupoid model) strikingly resembles Euclid’s geometry as it is presented in the First Book of *Elements*: HoTT makes up its geometrical universe starting with a point; then through an inductive definition further basic types/spaces are introduced: the type of propositions (having two terms *true* and *false*), sets, (1-)groupoids, 2-groupoids and so on. Each of these “levels” allows for further constructions of growing complexity [57]. The analogy with Euclid is quite precise in the following respect: in both cases the axiomatic procedure consists of a systematic introduction of theoretical objects (which can be propositions but generally are not). Recall that this is the precise sense in which Hilbert and Bernays call Euclid’s theory constructive. Like Euclid and unlike Hilbert and his modern followers Voevodsky does not use the propositional reduction as a means of axiomatic theory-building. 16. Saying that HoTT is a constructive theory one must a particular care of disambiguating the term “constructive” properly. MLTT is a constructive theory in a strong sense, which makes this theory computable. It is not known to the date whether or not AU is constructive in the same strong sense (prima facie it is not) 17. However HoTT with AU qualifies as constructive in the broader sense of “constructive” specified above even if it is not computable.

---

16Voevodsky remarks that the inadequacy of the standard set-theoretic foundations of mathematics becomes obvious in the context of HoTT and UF because these standard foundations for no good reason reduce the whole world of homotopy types to just one such type, namely that of sets [56]. A similar argument can be used against the idea of propositional reduction: even if in the axiomatic architecture of mathematics propositions play a special role, there is no good reason to reduce all types of objects to propositions using Hilbert’s device of “idealizing existence assumptions”.

17As to 2013 Voevodsky and collaborators describe this question as the “most pressing” ([7], p. 11).
In addition to the homotopical $\omega$-groupoid model discussed earlier in this Section HoTT with AU has some other “natural” models [1]. Studying and comparing these models from an unified viewpoint largely remains an open research problem. However Voevodsky makes it clear that he considers the the $\omega$-groupoid model as basic.\textsuperscript{18} Even if Voevodsky’s reason for preferring this model is prima facie pragmatic, it has strong epistemological consequences. The idea of reconstructing the world of everyday today’s mathematics using homotopy types as building blocks represents what Marquis calls the “geometrical point of view” in foundations of mathematics [39], [40]. However since the syntax of MLTT and AU also have a logical semantics such a reconstruction also qualifies as logical. However impressive is this new synthesis of logical and geometrical reasoning I cannot see that it represents an entirely new point of view in logic and mathematics because in Euclid’s constructive axiomatic approach logic and geometry are intertwined similarly. By Friedman’s word:\textsuperscript{19}

> Euclidean geometry [...] is not to be compared with Hilbert’s axiomatization, say, but rather with Frege’s *Begriffsschrift*. It is not a substantive doctrine, but a form of rational representation: a form of rational argument and inference. ([11], p. 94)

If we replace in the above quote Frege’s *Begriffsschrift* by Voevodsky’s Univalent Foundations the analogy with Euclid becomes even more close.

\textsuperscript{18}After expressing his misgivings about the standard set-theoretic foundations of mathematics Voevodsky writes:

Univalent foundations seeks to improve on this situation by providing a system, based on Martin-Löf’s dependent type theory whose syntax is *tightly wedded* to the intended semantical interpretation in the world of everyday mathematics. In particular, it allows the *direct formalization* of the world of homotopy types; indeed, these are the basic entities dealt with by the system. ([56], p. 7, emphasis mine)

\textsuperscript{19}Friedman attributes this view on Euclid to Kant without claiming that this view is his own.
8. What is Constructive Axiomatic Method?

In the preceding Sections I considered some recent trends in the axiomatic method in a historical perspective. I tried to show that in this perspective these recent trends appear as a revival of an older form of axiomatic method, which Hilbert and Bernays call genetic or constructive (Section 2). In the present final Section I would like to evaluate these new developments from a more critical and more theoretical viewpoint. Are such developments justified logically and epistemologically? Do they change the standard Hilbert-style axiomatic method in a right direction? What (if any) are logical and epistemological advantages of the new constructive axiomatic approach?

As my point of departure I take Hintikka’s recent paper [22] where the author defends a modern version of Hilbert’s formal axiomatic method. I shall try to show that this formal method is not self-sustained and needs to be supported by a constructive method. I shall also defend a stronger claim according to which the constructive axiomatic method is more fundamental and more basic than the formal method: while it is possible (but not necessarily desirable) to do some axiomatic mathematics “purely constructively” without using formal methods, it is strictly impossible to do any axiomatic mathematics “purely formally”.

Hintikka:

What is crucial in the axiomatic method [...] is that an overview on the axiomatized theory is to capture all and only the relevant structures as so many models of the axioms. Once a complete axioms system for some particular science is established, it literally becomes possible to investigate the phenomena that are the entire subject matter of that science in one’s study, instead of laboratory or observatory with pencil and paper. Of course, in our day and age pencil and paper are being replaced by a computer. But even so, using a computer is easier and cheaper than building an accelerator. ([22], p. 72)
Even if the above quote can be read as a proposal to replace real scientific experiments by theoretical calculations and numerical experiments it also allows for a more charitable reading: an axiomatically organized science allows for checkable theoretical predictions, which are logically inferred from the first principles of this science put into the form of axioms. This feature obviously has a great epistemic value and thus indeed provides a strong argument for using the axiomatic method in science more widely. Asking then where the axioms come from Hintikka says that they are obtained through an analysis of their intended models. And where the models come from? Hintikka gives the following answer, which could be equally heard from Hilbert (if one replaces Hintikka’s “class of structures” by Hilbert’s “systems of things” [15]):

The class of structures that the axioms are calculated to capture can be either given by intuition, freely chosen or else introduced by experience (ib., p. 83)

One may wonder how a mathematical structure (or a structure of some different sort) can be directly given by intuition, free choice, experience or whatnot without being construed axiomatically (or otherwise) beforehand? Should we assume at this point a Platonistic hypothesis according to which mathematical structures exist independently of our axiomatic descriptions of these structures? An answer implied by Hintikka’s understanding of mathematical intuition is different. He explicitly rejects the notion of intuition as an intellectual analogue of sense-perception and insists that the intuition (along with the free choice and experience) plays rather an active role. Defending his semantic view on logical inference (on which I comment below), Hintikka says

[N]ew logical principles are not dragged [..] by contemplating one’s mathematical soul (or is it a navel?) but by active thought-experiment by envisaging different kinds of structures and by seeing how they can be manipulated in imagination. The relations that can be so revealed are model-theoretical rather than proof-theoretical. Maybe such thought-experiment are examples of what is meant by appeals to intuition. But if so, mathematical

...
intuition does not correspond on the scientific side to sense-perception, but to experimentation. (ib., p. 78)

So what Hintikka wants to say is rather this. Building an axiomatic theory is a complicated two-way process; it is a game with Nature (and perhaps also with Society) where raw empirical and intuitive data effect one’s axiomatic construction in progress while this construction in its turn effects back one’s choice of further data, which become in this way less raw and more structured. Asking where the process starts exactly is the chicken or the egg kind of question. Hintikka’s IF logic with its intended game-theoretic semantics provides a precise mathematical model for games of this sort [45].

My concern is about the kind of games that we need to play with Nature for doing science and mathematics. Although yes-no questioning games indeed play an important role in science and perhaps also in mathematics I claim that this is only the top of an iceberg. The main body of this iceberg is filled by mathematical and empirical constructive activities such as designing new experiments. If we consider applications of mathematics outside the pure science we may also point to the mathematical design of new technologies and new industrial products. In order to design a bridge or a particle accelerator one usually helps oneself with mathematical models of that thing, not with formal axioms.

Since such activities qualify as instances of Hintikka’s “active thought-experiment” I don’t diverge from Hintikka up to this point. The divergence comes next. I don’t grant Hintikka’s view according to which the mathematical thought experimentation is, generally, a spontaneous ruleless activity, which should be studied by “empirical psychologists” rather than logicians, mathematicians and epistemologists (ib. p. 83). I observe that constructive axiomatic theories like Euclid’s geometry, Newton’s mechanics, Lawvere’s axiomatic Topos theory and Voevodsky’s HoTT-UF greatly increase one’s capacities of mathematical thought experimentation by providing basic elements (points and straight lines in Euclid) and precise rules for it. I can see that the spontaneity and the rulelessness may play a creative role in mathematics and science but I claim that typically a constructive axiomatic organization of science and mathematics makes the thought experimentations in these fields
richer and more powerful. Such an organized but not simply spontaneous mathematical thought experimentation is typically used for making theoretical predictions in sciences, designing bridges, accelerators etc.

The real question is not how liberally one can use the word “axiomatic” but how exactly an axiomatic theory controls its contents. Formal theory $T$ motivated by certain intended model $M$ controls (the content of) $M$ through the truth-evaluation of its axioms and theorems in $M$. For the sake of my argument I now understand (after Hintikka) the relation of logical inference in $T$ as the semantic consequence relation. Having granted this I claim that this method of axiomatic control is not self-sustained because the concept of semantic consequence relation is highly sensitive to one’s basic semantical setting. When the semantic consequence relation is discussed with respect to intuitive structures coming from the air (intuition, free choice, experience) it remains itself a very imprecise intuitive notion. I agree with Hintikka that this fact does not mean that one gets here a choice between appealing to irrational resources and giving up the semantical view on logical consequence (ib., p. 77-78). In order to construe the relation of semantic consequence with a mathematical precision one should fix some formal semantics, which allows for doing the truth-evaluation properly (as this is done, for example, in the case of Kripke-Joyal semantics for topos logic) $^{20}$. In other words one needs to build a basic mathematical model (in the sense of “model” used in science rather than Model theory) of the intended structure (or rather a sufficiently large class of such structures) suggested us by intuition, experience and what not. The mere yes-no questioning games cannot solve this problem because, recall, we are looking now for a mathematical framework allowing us to do the truth-evaluation properly. Unless we have got such a framework we are not in a position

---

$^{20}$In order to construe a notion of semantic consequence for given formal language $L$ one needs to
- fix a formal semantics $M$ for that language
- take a collections $A_T$ of well-formed formulas of $L$ (which may express axioms of a given theory $T$ formalized with $L$)
- specify the class $M(A_T)$ of models of $W_T$ by evaluating formulas from $A_T$ in $M$

Formula $\phi$ is called a semantic consequence of $A_T$, in symbols $A_T \models \phi$, iff $\phi$ is a tautology in $M(W_T)$. 
to give to yes-no questions definite answers. The wanted setting cannot come from the air but can be built by constructive methods. 21.

Recall that Hilbert realized the need to support his formal axiomatic approach by constructive methods considering this issue from a very different perspective: he meant to apply constructive methods for proving the (formal) consistency of axiomatic theories syntactically. Now we can see that Hintikka’s semantic approach to axiomatization does not allow one to avoid using constructive methods either.

The above analysis suggests a view on the truth-valuation as an advanced rather than basic feature of mathematical and other theories. Unlike Topos theory HoTT in its existing form has no resources for doing truth-valuation internally. It is however conjectured that such an internal truth-valuation for HoTT can be construed within a higher-order topos structure in which HoTT would play the role of internal language. 22. This example demonstrates my thesis that the constructive axiomatic method is more general and more basic than the formal version of this method, which requires the truth-valuation. This is hardly surprising since mathematics and science not only seek for truths and logical relations between those truths (knowledge-that) but also for effective methods of doing this and that knowledge-how [10]. From a historical viewpoint it is obvious that the knowledge-how is a more primitive form of knowledge, which can exist outside any scientific context. When science is brought into the picture there is an unfortunate tendency to isolate the

---

21In the last quote Hintikka hints that the relevant notion of logical (=semantic) consequence can be made more precise by introducing some “new logical principles” suggested by the intuitive thought-experimentation with mathematical structures. If the application of such new principles in its turn involves the truth-valuation then we get into a circle and make no progress. If these principles regulate the thought-experimentation directly without using the truth-valuation then these rules qualify as constructive in the appropriate sense. If the second guess is correct then my view on the axiomatic method is not so different from Hintikka’s after all. Still the following difference remains: I don’t think that constructive principles so obtained are necessarily logical because they can be specific for a given theory (just as Euclid’s Postulates are specific for his geometrical theory).

22In [7], p. 12, the authors state that there is a “general consensus” that this conjecture is correct albeit the details remain to be worked out.
relevant knowledge-how either in a special domain of applied science (and applied mathematics) or in social, psychological, educational, pragmatic and other contexts of doing science, which do not include scientific (and mathematical) theories as such. In this paper I have shown that this cannot work for axiomatic mathematical theories because in this case the two types of knowledge are interlaced already at the atomic level of theoretical reasoning.

The case of experimental natural sciences prima facie appears similar but obviously requires a separate study. Leaving this subject for a future research I would like to conclude this paper by a remark concerning applications of axiomatic method in empirical sciences. Such prospective applications play a significant role in Hintikka’s conception of axiomatic method [22]. The axiomatization of physics also always remained in the focus of Hilbert’s work [5]. Nevertheless during the past century Hilbert’s project of axiomatizing physics didn’t show any significant progress. In spite of efforts to axiomatize some ready-made physical theories made mostly by logicians and philosophers, the axiomatic approaches in physics today remain quite far from the mainstream developments in this discipline. Theoretical predictions in physics and other sciences are commonly made without using formal logical methods. I believe that this apparent failure of the formal axiomatic method in physics can be explained by the fact that this method leaves the constructive mathematical thought-experimentation (which matters in physics and other experimental sciences at the first place as Hintikka also recognizes it) without a proper control and proper regulation. Earlier in this Section I explained why the formal approach cannot provide the needed regulation by itself. Since the constructive axiomatic method provides this wanted regulation, one may expect that this version of axiomatic method can be more useful in today’s physics.

A historical observation that all earlier successful applications of the axiomatic method in physics used constructive (rather than formal) version of this method (recall Hilbert’s examples: Newton, Clausius) makes this expectation only more reasonable. An obvious reason why the constructive axiomatic method has been never applied in the 20th century physics is that during this period of time this method was known only in its technically
outdated form. This is why in the beginning of the 20th century it seemed that the only possible strict axiomatic treatment of then-recent theories such as Lobachevskian geometry could be only a *formal* axiomatic treatment. However as I tried to show in this paper later technical developments changed this perspective. This is why today the prospects of (constructive) axiomatic method and other related logical methods in physics and other experimental sciences are much better [2], [49], [50].

References


