What do you claim when you say you have a proof?

Claus Diem and Christoph Schulze

May 19, 2019

Prologue

In the common room

DISCIPLE: Your proof was quite beautiful.

SPEAKER: Thank you!

DISCIPLE: Actually, sometimes, I had a bit a problem following. There was quite a jump from the Fundamental Lemma to Theorem A.

SPEAKER: Well, yes, because the time was short. But I can give the details.

DISCIPLE: Rigorous?

SPEAKER: Yes, of course.

DISCIPLE: And Theorem B?

SPEAKER: The Fundamental Lemma also implies Theorem B; you can adapt the proof for Theorem A.

DISCIPLE: So, you have proven the theorems?

SPEAKER: Yes, I have a proof – otherwise I wouldn't say I have theorems.

DISCIPLE: Of course.

SKEPTIC: I've overheard your conversation. I would like to know what you mathematicians mean when you say that you "have a proof". That does not seem so obvious to me.

DISCIPLE: That we can demonstrate it rigorously.

SKEPTIC: This means in the basis of ZFC?

SPEAKER: Yes, you can say that "by default option" this means on the basis of ZFC.

SKEPTIC: So you say that you have proof of your theorems on the basis of ZFC?

SPEAKER: Yes ...

SKEPTIC: There was nothing remotely similar to proofs on the basis of ZFC on the blackboard.

SPEAKER: I outlined it. It requires work but you can convince yourself: One can give formal proofs on the basis of ZFC.

SKEPTIC: I couldn't, that's for sure. But I am sure that you couldn't do it either.

SPEAKER: Well, literally speaking maybe, at least I wouldn't do it. But that's just a practical issue, in principle it's possible. I can say how it could be done.

DISCIPLE: Of course, that's the claim. It should be clear that we do not literally write down formal proofs – it would be a waste of time and energy.

SKEPTIC: By saying that the alleged formal proof might be too long and in any case you do not intend to give it, you seem to acknowledge that you don't have a formal proof, and you don't even know if you could give one.

DISCIPLE: You know, formal proofs in ZFC are also abstract objects of mathematics; they can be studied just like finite geometries or groups. You should be familiar with this.

SKEPTIC: I am. But are you sure you want to say that the alleged proofs are abstract mathematical objects like groups?

DISCIPLE: Yes, like any other mathematical objects.

SKEPTIC: So by the "default option" you say that they are sets. Actually, I am not sure if it is justified to say that abstract objects called sets exist, but of course mathematicians act as if they would. So, let me act as a mathematician and also pretend that there exist these abstract objects called sets. Do you then say that saying "One can give a formal proof" shall actually mean "There exists a formal proof" in the sense of set-theory?

DISCIPLE: I would say that we do not just pretend that sets exists, but anyway: Yes, as I said, formal proofs are also mathematical objects, so yes, by default option they are sets. But I'm not so interested in formal proofs anyway.

SPEAKER: Well, I do think that it is important that mathematical statements are proven in a formal framework, or at least could be proven in such a framework. Otherwise, I wouldn't say that we have a result. So we really do want to have them, not just in some kind of abstract way.

SKEPTIC: But you don't.

SPEAKER: Well, not in a literal sense, yes. Any suggestions?

SKEPTIC: I guess it requires an analysis of what mathematicians actually do versus what they often claim to do, and then a careful analysis of claims of existence.

1 What do mathematicians claim on the weekend?

A saying among mathematicians, which unfortunately we, the authors, are unable to attribute to anyone in particular is: “From Monday to Friday, Mathematicians are realists, and on the weekend they are formalists”. The idea is of course that Mathematicians in their usual work appear to be realists and are also in a “realist mood” psychologically, but when they reflect on their endeavor they get their doubts and in order not to make unsubstantial claims, they withdraw a formalist position.

A statement of the form “I have proven Theorem A on the basis of Zermelo-Fraenkel set theory with the axiom of choice.” might then lead to a statement “One can give a formal proof of Theorem A on the basis of ZFC.”. At first sight, this seems to be an elegant solution to end any debate on the existence of abstract objects, but on a second thought this is not so clear anymore: As after all a formal proof is usually not given and what is written down is very far away from a formal proof, the question arises what is actually claimed here. Literally speaking the statement is doubtful, and because of this, one might be tempted to say that actually one just wanted to say: “There exists a formal proof of Theorem A on the formal basis of ZFC.” But in what sense is this statement then different from a Monday-to-Friday statement?

Evidently, a “weekend-claim” should neither be interpreted literally nor should it be interpreted in the sense of “usual” set-theoretic mathematics. Rather a “middle-ground-interpretation” should be given, an interpretation which is “just as fictional as necessary” and on the other hand “as literal as possible”. What could such an interpretation be?

A starting point is here the observation that it seems to be unreasonable to say that humans can establish infinitely many results, in mathematics or otherwise. So, we regard any statement that infinitely many results have been proven to be a priori invalid. To demonstrate this with an easy example: We have the evidently true statement on natural numbers “Every prime > 2 is odd.”.[1] One might, just for didactic reasons, reformulate this as follows: “For every prime $p > 2$ the following statement holds: p is odd.” As long as this is just understood as a reformulation of the first statement, this is of course unproblematic. But we hold that it is improper to say that here infinitely

[1] On our point of view on the interpretation of mathematical statements in general see the next section.

many statements are made or that infinitely many statements holds.

It is quite evident that a claim that "one can give a formal proof" should have, if not taken literally, at least a constructive meaning. This suggests that one might find an appropriate interpretation by considering ideas from constructive mathematics. Indeed, we want to argue that a particular, one might say "hypothetical interpretation" of constructive mathematics applied in this situation gives a appropriate interpretation.

In order to be able to even properly formulate the hypothetical interpretation of constructive mathematics, care has to be taken that no unsubstantial claims on the existence of abstract objects are made. For this, we take a view from the outside on mathematics, not preproposing any abstract objects like numbers or sets.

In order to highlight the informal character of what is usually called "proof" and in order to distinguish it from formal proofs, we call the former "convincing argument". We study the relationship between convincing arguments for a given mathematical statement and the claim that one has a formal proof, and with two thesis (Thesis 1 in subsection 5.1 and Thesis 2 in subsection 5.2), we suggest that from the one one can always obtain the other.

We particularly stress that for any given set-theoretic statement the claims that "one can give a formal proof on the basis of ZFC" (interpreted as suggested by us) and "one can prove set-theoretically that there exists a formal proof on the basis of ZFC" are – at least by current knowledge – substantially different. From this point of view, we in particular criticize, in subsection 5.6.8, a series of statements by G. Kreisel in [informal-rigour] in which this distinction is not made.

In addition, we argue that for any given set theoretic statement A the following holds: One can convincingly argue in Zermelo-Fraenkel set theory that there exists a formal proof of A on the basis of ZFC if any only if one can convincingly argue in Neumann-Bernays-Gödel set-and-class theory that A holds; see Theorem of Meta-Mathematics 3 in subsection 5.7. This gives a new (and as we think interesting) interpretation of what it means to establish a set-theoretic statement via Neumann-Bernays-Gödel set-and-class theory.

2 Interpreting mathematical claims

2.1 Realism versus anti-realism

According to adherents of Platonic realism, mathematics is (or should be if properly conducted) the study of certain "really existing" "abstract objects". Of course this is challenged. In what sense is, for example the statement

> There exists a unique positive real number whose square is $2 \uparrow\uparrow 6 - 1$,

in which Donald Knuth's arrow notation (see [coping]) is used, a statement on any kind of "object" after all?

Indeed, one might argue that as there is no such object, the statement is wrong. One might also raise the even deeper objection that notions in the statement (*real number*, *square* and maybe also $2 \uparrow\uparrow 6 - 1$) are void of meaning, and it therefore is not even reasonable to discuss whether the sentence is right or wrong. More generally, one might argue that all or at least a large part of mathematical existence statements which are considered to be true in mathematics are in fact wrong, because there simply are no objects of the asserted kind, or even nonsensical.

2.2 Empirical approach

Now, independent of metaphysical believes, mathematicians usually do indeed act as if the asserted imagined objects such as the square root of $2 \uparrow\uparrow 6 - 1$ were real objects. They talk about such object, they even get emotionally involved when they have found a new, as mathematicians often say, "beautiful" relationship in their imagined universe. Interestingly, throughout the history of mathematics, aesthetic judgments played a crucial role, maybe more than in any other science.

Let us take the observations of the previous paragraph as a starting point: Rather than asking what mathematics *is*, we focus on the *activity* of mathematicians in their role as mathematicians (which is meant as a self-identification). What is then this activity, an activity which might be called "doing mathematics", which shall be interpreted as one predicate.

There is then a purely *empirical* answer, following the above descriptions: Mathematicians act as if imagined "abstract" objects in an imagined "abstract universe" were real. For this, they claim to base their arguments on certain simple assertions called "axioms" and to apply (only) certain rules of inference. They try to find out what the characteristics of the imagined universe are, what the relationship between objects in the imagined universe are, and they try to convince others of their findings.

Moreover, mathematicians are not neutral with respect to the given axioms and rules. Indeed, usually the argue on the basis of some variant of set theory; such a foundation usually has a the appeal to mathematicians to be "particularly evident". Indeed, very often it is considered to be so evident that it is just assumed.

We note here that we use the word "imagined" for an empirical description of the activity of mathematicians. Throughout this note we stay neutral towards the fundamental question in how far the content of mathematical statements (in particular of statements in infinitely many objects) should be regarded as being "real" in the sense of being statements on "real abstract objects". It is, however, important to us that statements on infinitely many objects and also statements on large numbers, such as the one mentioned in the beginning of this section, cannot be regarded as statements on objects of human experience.[2]

2.3 Mathematics and fiction

The above thoughts are, of course, not new. Let us indicate some relationships: A quite similar problem occurs concerning talk on novels. A classical sentence is here "Anna Karenina is a woman". Literally speaking the sentence is nonsensical, so how can it be understood? Now, according to literary anti-realism, the proper way to understand the sentence is "According to the fiction of the novel Anna Karenina, Anna Karenina is a woman". This sentence is literally true, whereas the sentence "According to the fiction of the novel Anna Karenina, Anna Karenina is a man" is false. This idea has then be further developed to a *pretense theory* of talk on literary fictional entities. In the context of a pretend-as-if situation, one can then argue about the truth of statements according to the agreed-upon-fiction. Moreover, from an outside point of view, one can analyse whether people do indeed speak according to the agreed-upon fiction of the context, and one can criticize people for inconsistency if they fail to do so.

A good account of this line of thought is Section 2.1 of [fictional-entities]. For the further discussion, it is very helpful if the reader is familiar with this text.

This idea then suggests to interpret mathematical claims in a similar way, that is, with in-the-fiction-of operators or by pretense-theory. This line of thought has first been developed by Hartry Fields and is called "mathematical fictionalism" (see for example [real-math-mod]).

One can therefore say that we essentially follow the line of thoughts of mathematical fictionalism. As indicated above, we wish to make one difference between the point of view of this note and mathematical fictionalism and more generally anti-realism:

We stay neutral with respect to the existence of mathematical objects. Our goal is merely to not make any unsubstantiated claims of existence. In fact, in Section 5, we will study the relationship between what we call different "kinds of mathematics", and it would be quite natural for a person

[2] This still holds if one also counts imaginations as human experiences. One might say that for some mathematical claims, humans can imagine in a certain way corresponding objects, and some mathematical claims might refer to these objects. We hold, however, that humans cannot imagine infinitely many objects, one can just get a rather vague feeling about it.

to consider him- or herself a realist with respect to one kind of mathematics and not with respect to another kind of mathematics. For example, a person might consider natural numbers and arithmetic to be a priori given but set-theory a merely fictional human invention, another person might consider set-theory without the axiom of choice as a priori given but the axiom of choice to be merely fictional.

To say, as we do, that when mathematicians do a particular kind of mathematics (rather then talking about the kind of mathematics), they act as if the kind of mathematics was real is in any fact a proper statement on which, it seems to us, every mathematician can agree. From a psychological point of view, we can even go further: Particularly when doing the usual kind of mathematics, set-theory, mathematicians not only "act as if", when doing mathematics they also get into a mood of imagining that the objects they reason about "really exist".

2.4 Mathematics versus literary fiction

A reader might ask the following about the fictional interpretation of mathematical statements: Are you saying that doing mathematics is like telling novels?

To clarify, we discuss this question a bit: Above we pointed out an analogy between making mathematical statements and making statements pretending as if what was told in a novel was real. So, in particular, it is not claimed that "telling mathematics" is analogous to telling novels. If anything then a given novel would correspond to a description of an axiomatic system, not to a mathematical text written in a pretend-as-if way.

As no claim is made like "doing mathematics is like telling novels", one should also not object to the fictionalist interpretation by trying to refute this claim.

For further information on this discussion the reader might consult Section 2.4 of [fictional].

2.5 Problems with our neutral point of view

We note again that we take a neutral point of view towards the existence of abstract objects. So arguments that abstract entities exist in one way or the other are of no importance for us.

There is, however, a potential problem with our neutral position which we explain with an example:

Say a person is convinced that sets as described with the Zermelo-Fraenkel axioms with the axiom of choice "really" exist as abstract entities. This belief should be compatible with our neutral position and not cause any problems.

Now, say further that the person says: "I will argue for Theorem A assuming the Zermelo-Fraenkel axioms without the axiom of choice." For him, the axioms describe the really existing abstracts sets. He or she might now say:

> Whenever I talk about fiction, I of course assume that the usual reality is in place – at least as it does not outright contradict the fiction, in fact, in the case of literary fiction I have to do that. In this vein, I can now use everything which holds of sets, so I can also use the axiom of choice. And by the way: My statement on what I will be assuming was irrelevant because I don't have to assume anything; you might find that confusing, but still I won't make a mistake.

To circumvent this problem, we postulate that whenever in a "doing-mathematics"-situation a person argues "on the basis of" some particular axioms, this shall be interpreted as talk in a purely fictional context with no relationship to physical or abstract objects.

In a bold way, this can be realized by using another in the fiction-of operator as follows: Before the introduction to the in-the-fiction-of operator / pretend-as-if introduction for the particular kind of mathematics (in the above example Zerlemo-Fraenkel set theory without the axiom of choice) it is stated: "Let us pretend that abstract objects do not exist." or "According to the fiction of anti-realism the following holds:"[3]

This solution is analogous an anti-realist solution presented in [fictional-entities] to the following problem: How shall one understand, from an anti-realist point of view, a so-called external metafictional sentence like this one: "Anna Karenina is a fictional character." The solution presented in subsection 2.1.2 of the mentioned article is put the sentence into a fictionally realist context, for example by this introducing sentence: "According to the fiction of realism, the following holds:"

2.6 Convincing arguments, that is, "proofs"

An interesting aspect of doing mathematics (the activity of mathematicians) is that certain arguments (which by the nature of mathematics are always relative to "axioms", or – as suggested above –, can be regarded as being stated in a pretend-as-if context described by the axioms) are considered to be so rigorous that they are called "proofs". This is even more remarkable as there is no clear standard as to when an argument given by a mathematician shall be called a "proof". Indeed, throughout mathematical history,

[3] We recall that we use the in-the-fiction-of operator in a neutral way, so the sentence shall not implicitly imply that anti-realism is false.

mathematicians have differed greatly in their opinions as to what they consider a "proof", and if one now rereads older texts (which might be from the 19^{th} century), sometimes one obtains the impression that what is called a "proof" does not have the clarity and strictness one wishes to see nowadays. Nonetheless, mathematicians have usually been convinced that it is in praxis obvious what constitutes a "proof" and what not, and in case of a dispute mathematicians have usually been able to find – via discussions and further elaborations – a consensus as to what shall considered to be proven (again relative to particular axioms) and what not. These "proofs" together with discussion on them as well as the expressed conviction that they express a certainty with regards to truth form a major part of what is expressed in writing by mathematicians.

The word "proof" is a strong one, but as stated, more often then not in the actual praxis of doing mathematics, one can argue if arguments suffice for a "proof". Personal judgments play a strong role here. To emphasize this subjective aspect, we usually speak about a *convincing argument* instead of a *proof* when we refer to what is usually considered a proof. Occasionally, we also literally write *"proof"*. This also allows us to distinguish between what is usually called "proofs" by mathematicians (thus convincing arguments in our terminology) and (formal) proofs in formal systems, a distinction which will be of importance later.[4]

2.7 Convincing arguments versus formal proofs

Let us finally, as in Section 1, consider a situation in which some "Theorem A" is discussed, and let us consider the two claims stated already in Section 1:

1. "I have proven Theorem A on the basis of Zermelo-Fraenkel set theory with the axiom of choice."
2. "One can give a formal proof of Theorem A on the basis of ZFC."

We interpret the first claim in the sense that the person claims to *have a convincing argument* for the theorem in the fiction of or in the pretend-as-if situation given by (an informal description of) Zermelo-Fraenkel set theory with the axiom of choice.

To the second claim we wish to give a constructive meaning. We will suggest such a meaning in subsection 4.4. Then we will study the relationship between the two claims, and suggest three thesis for the relationship in Section 5.

[4] It is common to call what we call *convincing argument* an *informal proof* (see for example [informal-proofs]). Other researchers, who also emphasize the difference between what is actually expressed and formal proofs stick with the word "proof" (see for example [formal-natural].

We already note here that despite their similarity, the two phrases are interpreted quite differently: The first phrase is interpreted with the pretend-as-if situation for set theory, whereas the second sentence is interpreted with a pretend-as-if situation of constructive mathematics. So in particular in the first sentence the "on the basis of" refers to the fictional situation whereas in the second sentence it is part of what is claimed in the fictional situation.

3 Two crucial developments

Two rather recent developments in mathematics are of particular importance for us: *sets as fundamental objects of mathematics* and the *formalistic approach* to the foundations of mathematics.

3.1 Set theory

Set theory is so embedded to our doing of mathematics that it is hard to forget the specific, at one time revolutionary, point of view of this way of thinking: Not only does one regard collections of (imagined) "things" again as one thing, the "things" so obtained are then regarded as things in just the same manner, which makes it possible to repeat this process. Moreover, not just collections of finitely many "things" are considered, but also such of infinitely many "things".[5]

An early idea of the notion of set (Menge) was that every property should define a set. As this turned out to lead to logical contradictions, a separate notion of *class* was introduced, for which this is indeed the case (with an appropriate definition of "property") but which cannot necessarily be treated as "thing" in the way a set can.

We note that with respect to the use of proper classes, that is, classes which are not sets, the attitude of mathematicians is not always clear. Whereas statements about classes frequently appear in mathematical writings, often these can be reformulated via sets alone. It remains then the question of the use of proper classes is solely meant as a façon de parler or if proper classes are really considered as objects in addition to sets.[6] In order to be neutral to this question, we will speak about "set or class theory" rather than only about "set theory" or only about "class theory" (which of course would encompass a theory of sets).

[5]Richard Dedekind wrote in [Zahlen]: "Im folgenden verstehe ich unter einem *Ding* jeden Gegenstand unseres Denkens.", that is: "In the following, I mean by *thing* every object of our thinking."

[6]We stress that as stated in the previous section we view the doing-of-mathematics as speaking in pretend-as-if contexts. So we ask here what kind of objects are imagined or should be imagined for a consistent interpretation inside these contexts. In particular do not ask if sets are "more real" than classes.

Amazingly, via the idea of sets, statements and arguments about abstract or – as we see it in this note – fictional objects, numbers, (mathematical) geometries, forms, functions and much more be put into a common fictional context – the context of set theory. In this context of course then the approriate set-theoretic definitions have to be made. This unifying set-theoretic point of view has been so successful that nowadays the view is dominant that every object considered in mathematics is (or: should be seen as) a set or a class.

One should, however, be aware that there are other kinds of mathematics than set and class theory. We will clarify the notion of "kind of mathematics" and the role of different kinds of mathematics in subsection 3.4. Here we merely give some remarks of importance for the time being:

We first remark that the dominant set-theoretic view changes the intuition in such a way that perspectives which cannot be expressed or are difficult to express set-theoretically vanish nearly completely. To give an example, the usual definition of a function in set theory is such that a function is equal to the graph of itself. To a person already embedded into set-theoretic thinking, this seems to be natural, and such a person might laugh at descriptions involving "rule", "computation" or "dependent quantities". But the fact that "function" has an elegant set-theoretic definition does not mean that other ideas are without value, whether they might be also defined (maybe more difficultly) set-theoretically or are not definable set-theoretically at all.

A second remark we demonstrate with an example: A particular axiomatic system for synthetic geometry with basic objects being points and lines leads to a corresponding kind of mathematics. Now, one can define set-theoretically what one means by a geometry corresponding to the given axioms, consisting of points and lines, and then one can argue set-theoretically about these. One can then expect that every convincing argument on the basis of the geometric axiomatic system also set-theoretically leads immediately to a set-theoretically convincing argument. This does, however, not mean that the opposite also holds, that is, that a convincing argument on the bais of set-theory on corresponding objects (which are, of course sets) leads to a convincing argument on the basis of the geometric axiomatic system. We will discuss these aspects in subsection 5.6.

3.2 The formalistic approach

By "formalistic approach" we mean the following body of ideas: A formal language to express mathematics is rigorously described via (easy to follow) rules according to which certain expressions are called sentences;[7] certain

[7] One might also speak of "valid sentences"; we use the term "sentence" in a sense that "valid" would be redundant.

sentences or bodies of sentences are considered to be interpretable (again via easy to follow rules);[8] certain sentences or bodies of sentences (called "axioms") are a priori called "true";[9] there are (easy to follow) rules of inference to derive further "true" or "false" sentences from previously established "true" ones. Calling the language and the rules a *formal system*, a "proof in the formal system" is then a (physically given) (finite) sequence of sentences in the formal language each of the sentences is interpretable and "true" given the previous ones (on the basis of the axioms and the rules).

Rigorously following the formalistic approach would mean that all mathematics would be rewritten in formal systems, and then the outline of the formal systems used and these formal texts written would encompass all of established mathematics. Motivating and interpreting statements could be given, but there would be no need for them, and they would not be part of the body of established mathematics; in particular in this body of established mathematics, there would be no room for "universes of discourse" in which the sentences in the proofs are to be interpreted.

Following this line of thought, ideally then mathematicians would even agree on a single formal system, and there would be an agreement that mathematics is the production of formal proofs in this unique formal system. The rules for the formal system could then be seen as defining criteria for the notion of mathematics, but would themselves lie in meta-mathematics rather than in mathematics. Needless to say that this vision is very different from mathematics as the activity of mathematicians at this point of time.

This rigorous approach should be distinguished from the usual process of doing mathematics, which always involves the effort to find appropriate precise terms and convincing arguments written in a language which keeps possible misunderstandings at a minimum.

We note here that the development and the use of a formalistic approach as described need not go along with a particular attitude with respect to the "nature" of mathematical statements. In particular, it is independent of the acceptance or rejection of Platonic realism.[10]

[8] By "interpretability" we mean a formal criterion on texts, i.e. finite sequences of sentences. In propositional logic all sentences are interpretable, but formal languages more in line with conventional languages are conceivable in which this is not the case. An example might be a formal language with a sentence like "Let x be an element of X.".

[9] We deliberately write "called 'true' ", because we want to emphasize that here "true" and "false" are merely expressions assigned to certain statements, which might be replaced by any other expressions. The fact that the emotionally strong words "true" and "false" (which are also used in statements on formal systems in their usual meaning) are used here is of course not without problems. A particular reader of a particular system might be of the opinion that the system with a particular interpretation does really establish true and false statements; but this does not need to be so.

[10] See also the last sentence of footnote 9.

3.3 The formalistic set theoretic foundation

The formalistic approach can then in particular be applied to set or class theory, which leads to various (related) formalistic approaches to set or class theory as the foundation of mathematics. With these approaches, it seems that the foundations of mathematics as it is usually conducted have reached a long-lasting nearly stable stage.

3.4 Other kinds of mathematics

Up to now we identified mathematics with some set or class theory which gives a unified foundation of a lot of mathematical practice. In general, when speaking of some *kind of mathematics*, we always expect that there is some (description of) collections of rules of inference and of axioms.[11] This condition gives a very general setting of mathematics. However, usually mathematicians have a more restricted view which things are really worth being considered as mathematics and opinions may differ in this issue.

Given an informal description of a kind of mathematics, it should be possible to write down a formal system about which it can be said that it gives the kind of mathematics formally. In the following, we assume that such a formal system is actually given.[12]

A kind of mathematics is not to be confused with a mathematical theory, like number theory or K-theory, considered inside some set or class theory - they do not comprise a separate foundation. Also, group theory considered as a theory of particular sets called groups is not a kind of mathematics, but group theory as considered in abstract algebra (that is, as a theory on elements called group elements with certain operations satisfying certain axioms) is a kind of mathematics.

Famous examples for set or class theories are Zermelo-Fraenkel (with or without the axiom of choice) and Neumann-Bernays-Gödel set-and-class theory, where only the second one uses proper classes. These constitute different kind of mathematics, even though model theory indicates that one should expect basically the same results in these two kinds of mathematics. But there are also weaker, stronger or different set or class theories. For example one may use a set theory without the axiom of choice or some weaker or stronger version of this axiom. As another example, one can demand the existence of inaccessible cardinals, or equivalently of Grothendieck universes.

[11]Since this is an informal description of mathematics, we do not restrict the language or the languages for the doing of the kind of mathematics in any way. Following what is common, we expect that several natural languages combined with some symbolism are used.

[12]In the words of subsection 4.3 of the next section, the system shall be algorithmically given.

These changes have real consequences: A text might constitute a convincing argument in one kind of mathematics but not in another, related one.

Kinds of mathematics different from set or class theories may be found in mathematical history or, as mentioned before, as a theory studied in model theory. Synthetic (euclidean) geometry, propositional logic, Peano arithmetic, group theory or some theory of (maybe further specified) fields as in model theory or even some theory of Sudokus may serve as examples. For these kinds of mathematics it would not be unusual to be formalized inside some set or class theory, but we would not consider them as such but – as stated – as defined by their own axioms.

Given a mathematical text, in most cases mathematicians would interpret it inside some set or class theory. However, usually it is considered to be irrelevant to give a precise listing of the axioms. Mathematical texts (with statements and convincing arguments) can then often be interpreted in different kinds of mathematics.

A class or family[13] of particular kinds of mathematics of importance for us is constructive mathematics. Constructive mathematics has, however, its very own flavor to it which makes it different from the "usual" kinds of mathematics based on first order (classical) logic. In the next section we will explain in detail why we consider constructive mathematics to be the right framework for claims that statements can be proven formally, for example with respect to ZFC but also with respect to other formal systems.

4 Constructive mathematics

As the reader might not be familiar with constructive mathematics, we now first provide some information on it. We then discuss how we intend to use constructive mathematics in arguments on the existence of formal proofs of theorems one wishes to establish. Finally, we give a particular interpretation of statements of constructive mathematics, which might be called "hypothetical interpretation", and which we view as being particularly suitable for our applications.

4.1 What is constructive mathematics?

Constructive mathematics (including its meta-theories) is a body of ideas on what statements and arguments are proper in mathematics, how these statements should be interpreted together with actual mathematical results following these ideas. It emerged from the "Gundlagenstreit" in the beginning of the 20[th] century, and according to this, it is often associated

[13]We use these notions purely informally here.

L.E.J. Brouwer's intuitionism. From the current point of view, constructive mathematics is broader than intuitionism, or to say it differently, intuitionism is one of the schools inside of constructive mathematics. As shall be made clear below, it is not this school we are interested in for our purposes.

As constructive mathematics is not commonly studied, a first question is: What ideas are alerted to by the term "constructive mathematics?

According to the opening sentence of the entry on constructive mathematics in the Stanford Encyclopedia of Philosophy ([constructivism]), written by the constructive mathematician Douglas Bridges:

> Constructive mathematics is distinguished from its traditional counterpart, classical mathematics, by the strict interpretation of the phrase "there exists" as "we can construct".

For us, it is important to note that the phrase "we can construct" is not taken literally. Rather in constructive mathematics, also pretend-as-if statements are made, albeit statements different from the ones in classical mathematics.

In constructive mathematics, the starting point of reflection are the natural numbers, which are a priori assumed to exist.[14] Fundamentally, the mathematical statements are on algorithmic operations on natural numbers and furthermore on algorithmic operations on algorithms operating on natural numbers and so on. For example, a sequence of natural numbers is an algorithm taking natural numbers and outputting natural numbers too. A function from the sequences of natural numbers to sequences of natural numbers is then an algorithm taking and outputting such an algorithm.

Care has to be taken, however, because constructive mathematics is not just computational mathematics with natural numbers. Indeed, according to the philosophy that everything has to be constructed, the algorithms also have to be constructed. This means that algorithms claimed to exist have to be output from previously defined algorithms.

This still does not give a clear criterion what one must do to establish that an algorithm can be constructed. What kind of arguments are allowed for this?

The solution is centered around a self-limitation already on the level of the underlying logic with restrictions around existence statements and negations. In particular, it is not allowed to use implications of the form

$$(\neg\forall x : A(x)) \to (\exists x : \neg A(x)) \,. \tag{1}$$

Another aspect of constructivism is: It is emphasized that it is only reasonable to speak about the truth value of a mathematical statement if

[14] See however our interpretation of statements of constructive mathematics in subsection 4.3.

one can convincingly argue for its truth or falseness. The pragmatic rule is that to make a statement A shall mean exactly the same as to make the statement "A can be proven."

This general principle is then applied when sentences made with logical connectors are to be interpreted.

So a statement $A \vee B$ shall not only mean that $A \vee B$ can be proven but also (by applying the rule internally) that A can be proven or B can be proven.

Likewise, in constructive mathematics, a statement of the form $A \rightarrow B$ has the same meaning as "Every (potential) proof of A can be converted into a proof of B", "One can give a procedure that every proof of A can be converted into a proof of B.". Moreover, in all the three statements, one can add without modifying the meaning of the statement the initial phrase "It is known that" or "One can convincingly argue[15] that". A negation of a statement, say $\neg A$ of A, means then that it is impossible to prove A, that is, every try of a proof of A leads to a contradiction. Finally, an existence statement, say $\exists x : A(x)$ means that one can show that an x_0 can be constructed for which $A(x_0)$ holds.

We note that it is then clear why the implication (1) shall not be used, just in the same way that $\neg\neg A \rightarrow A$ and $A \vee \neg A$ shall not be used.

These intuitive rules or reasoning are made precise in the formal system "Intuitionistic Predicate Calculus" or IQC, developed by Arend Heyting. In up-to-date expositions on constructive mathematics, such as in [constructivism], these logical foundations are stressed. Interestingly, the founder of intuitionism, Brouwer, emphasized the preliminary role of mathematics over logic, but as is often the case, an intellectual system has been created which now has its own philosophy, independently of its historical origins.

Building on intuitionistic logic, there are different schools of constructivism. There is agreement on the use of the natural numbers and the importance of the notion of algorithm. However, there are variation concerning what principles of reasoning are allowed, on the style of presentation and on possible interpretations of statements. Information on the different schools can be found in the books [foundations] (Chapter III) and [constr-in-math] (Chapter I, Section 4), which we also recommend independently of this.

The different schools allow ways of reasoning which go beyond a pure algorithmic construction. Most strikingly, Brouwer allowed for the "possibility to use our free will to decide at each state what the next number in the sequence will be", as expressed by Michael Beeson in [foundations], Chapter III, Section 4. Another principle was used by Andrey Markov Jr. He ar-

[15]In usual terminology: prove

gued: If it is not true that a particular algorithm does not terminate (which means by the principles of constructivism that one can refute every attempt to prove that the algorithm does not terminate), then it terminates.[16]

An important extension of purely algorithmic constructive mathematics is the introduction of the notion of set. As stated in Chapter VIII of [foundations], there are two approaches: One can add the notion of set (or class) or one can postulate a "real" constructive set theory. Moreover, often a constructive axiom of choice is used, for example by Bishop.[17]

Besides different basic principles, the schools of constructive mathematics differ concerning the style or writing and the way of arguing. There is the "dry" "Russian school" which is essentially recursive function theory with intuitionistic logic.[18] In opposite direction, there is Errett Bishop's book Foundations of Constructive Analysis (with a second edition with Douglas Bridges called Constructive Analysis [constructivism]). These books highlight the spirit of constructive mathematics, that is, the construction, while not putting too much emphasis on foundational questions and ignoring discussions on foundational formal systems altogether. The first book, published in 1967, is of historical importance because it showed that one can really "do" constructive mathematics, and one can speak of a "Bishop school of constructive mathematics"; cf. [constructivism]. For us it is of importance that Bishop's book is intuitively (in contrast to (overly) formally) written and is based on intuitionistic logic and the claim that the natural numbers exist a priori.

4.2 Our use of constructive mathematics

For our purposes, that is, for arguments on the existence of formal proofs of theorems one wishes to establish, there is no place for free-choice sequences. As said above, one can integrate set-theoretic arguments into constructive mathematics. We do not envision this for our applications.

Interesting is now Markov's principle. We do not want to allow this principle either because otherwise we could argue for the existence of formal proofs of mathematical statements by contradiction. We regard this as being too weak.[19]

[16]See [constr-in-math], Section 4.6 and note for comparison also [foundations] Chapter III, Section 1 with the exercises.

[17]It might seem that in constructive mathematics there is trivially always a choice function when one desires one because everything that is claimed to exist must come along with a construction. As explained in [constr-theories, I,4.7] this is not so.

[18]This evaluation follows [constructivism].

[19]An abstract version of Markov's principle says $(\forall n : (\phi(n) \vee \neg\phi(n)) \wedge \neg\forall n : \neg\phi(n)) \rightarrow \exists n : \phi(n)$; cf. [constr-theories, I,7]. We reject this principle in our application for arguments on formal proofs for the reason given.

One aspect has not been addressed so far: Is it reasonable to really allow all thinkable algorithms? This question can be answered along just as it is usually answered in constructive mathematics. To illustrate the answer, let us first consider constructive elementary number theory. Here, finitely many algorithmically defined functions[20] are used to make statements on natural numbers, and then these statements are analyzed, however, there are no "higher algorithms" producing algorithms. The important aspect for us is that it should be obvious that the domain of the functions is total. This suggests to only consider functions defined by loop-algorithms, or, what amounts to the same with respect of equality of functions, functions defined by primitive recursion. This idea is formalized in the formal system of *Heyting arithmetic*, HA. This system is built on IQC with a function term and corresponding axioms for each presentation of a primitive recursive function.

One can then "go up" and iterate the idea of operation by primitive recursive function by considering primitive recursion operating on algorithms. For this, to keep track on what the algorithms actually operate on, one should work with *finite types.*[21] One then naturally obtains the notions of *finite type recursion* and – as a generalization of a primitive recursive function – *primitive recursive functional.*[22] Interestingly, in this way one can also obtain new functions from $\mathbb{N}$ to $\mathbb{N}$: As explained in [constr-in-math, 9, 1.4], the premier example of a totally recursive but not primitively recursive function, the Ackermann function, can be given in this way. The ideas presented here have been formalized in a typed formal system called *finite type arithmetic*, HA^ω.[23]

In another direction, an important question is how formal an argument for the existence of a formal proof of a theorem shall be. Would it, for example, be reasonable to demand that it be written in HA^ω (with meta-

[20] From the point of view of constructive mathematics, functions are the same as algorithms, however, when we speak about functions, we have an extensional notion of equality in mind and when we speak about algorithms are more refined one. Note here that in contrast to "usual" set theoretic mathematics, "equality" in constructive mathematics is not assumed to be a priori given; rather different notions of equality are used depending on the context.

[21] A type is an expression built following these rules: 0 is a type and if σ, τ is a type then $\sigma \times \tau$ and $\sigma \to \tau$ is a type. Algorithms of type 0 are the algorithms which do not have an input and output a natural number. The function associated to such an algorithm can be identified with the output, that is, one can say that the function is a natural number. For the interpretation of the statements given in the next subsection it is, however, important that an algorithm of type 0 is not a natural number – it is an algorithm producing a natural number.

[22] These ideas were developed by Kurt Gödel in [dialectica].

[23] Stictly speaking there are at least two distinct systems with this name. The "most basic" system is given in [constr-logic&-math], which also gives a nice exposition to constructive mathematics. The definition of HA^ω in [constr-in-math] includes a "combinator" which is not present in the definition in [constr-logic&-math].

variables, particularly for the function symbols)?

We can give a clear “no” to this question. Just as set-theoretic mathematics, constructive mathematics is never done on this level of formality (apart from small calculations), and clearly it is not reasonable to demand that it should be.

Is it then reasonable to demand that at least it “should be possible” to reformulate / rewrite an argument in this system? Again, we can answer this with a clear “no”. As such a rewriting realistically would not be carried out and cannot be carried out either and as our starting point was the question what such a claim of “should” could mean in such a situation, it would be ironically inappropriate to carelessly introduce such a demand here.

There is finally the idea for the even weaker demand that one can argue constructively (as envisioned by us) that one can rewrite the argument in a formal system of constructive mathematics (like HA^ω). But then this new rewriting procedure would realistically also not be given formally, and one would be required to have a new rewriting procedure, and so on, leading to a never ending regression.

In summary and positively expressed, for arguments on the existence of formal proofs of mathematical theorems, we envision an informal presentation which is “purely algorithmic” along the lines just outlined.[24] The presentation might be similar to that of Bishop’s book with possibly a presentation of algorithms given by pseudo-code. (But as said, Bishop uses principles we would not like to use, namely sets and the constructive axiom of choice.)

We note that with this approach has the nice feature that one can always “go up” with algorithmic arguments.[25] This means that if we envision a meta-analysis of arguments on the existence of formal proofs, a meta-analysis of this and so on, we never have to leave the framework outlined.

4.3 Our interpretation of constructive statements

We now come to the *interpretation* of constructive statements for our application in arguments on the existence of formal proofs of theorems, along with a corresponding suggestion for the use of language.

Following the citation in the beginning of subsection 4.1, “constructive mathematics is distinguished from [...] classical mathematics by the strict

[24] In the terminology given by Solomon Feferman in [constr-theories] (which is also given in [foundations, V,3]), our assessment is that the system HA^ω as defined in [constr-in-math] is directly adequate and directly in accordance with the envisioned body of arguments on the existence of formal proofs of mathematical theorems, where texts on a fixed alphabet are identified with natural numbers.

[25] This corresponds to the formal feature of HA^ω that for two types σ, τ, there is always the type $\sigma \to \tau$.

interpretation of the phrase 'there exists' as "we can construct". More precisely, it seems that there is a general consensus of mathematicians who regard themselves in the constructive tradition that one imagines the natural numbers and strives to construct everything from this basis – as was programmatically expressed by Leopold Kronecker (see [on-Kronecker]): "Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk."

For our applications, the evident approach would then be to say that texts based on a fixed alphabet are represented by natural numbers and vice versa. Statements as described in the previous subsection then become statements on transformations of texts via algorithms, transformations of algorithms on such algorithms and so on. Above, we have tacitly already used this identification of natural numbers and texts.

Let us reflect on the meaning of the statements on texts which are then made: We first note that the concept of text is not the same as the concept of a natural number, or to say it differently, the phrases "text" and "natural number" have a prescientific meaning which is clearly not identical. Texts have the evident property that they can be given literally; what one writes down *is* a text and does not just denote a text. In contrast, the idea of a natural number is always abstract. Surely, one can represent numbers by texts, for example using the usual decimal representation or the "elementary" unitary representation. Nonetheless, if for example somebody writes down "123" she or he surely writes down a unique text independently of the meaning of the symbols, but she or he only gives a number if a meaning is attached to the text.

So, texts can be given physically-literally,[26] but in our application we talk about texts which are constructed by algorithms without this ever being the case physically.

The goal is now to limit the imaginary aspect to a minimum. Particularly, we do not want to make statements on the transformation of an *infinite amount* of texts via algorithms. Rather we want to make statements on the transformation of *individual texts* via algorithms. Also, because algorithms can also take algorithms as inputs and output algorithms, we want to make statements on the transformation of *individual algorithms.* This will then correspond well with the statement made in Section 1 that we regard a statement that infinitely many results have been established to be a priori invalid.

For this, we use what we have already used in the previous subsection: We can give algorithms by writing down pseudo-code, and we can iterate this process by using already given algorithms. We can then in particular

[26] One can then discuss when two texts should be regarded as being equal and also if a physically given text denotes something like an "abstract text"; in any case it is evident that texts can be given physically and literally.

consider algorithms which take no input. We then say that the output *is algorithmically given* or *can be constructed*. As a special case of this, we can consider algorithms which output texts. Again, we then say the output text is *is algorithmically given* or *can be constructed*.

We iterate and stress again that when making these statements, we are again making statement which should not (and in many applications cannot reasonably) be interpreted physically, so we are again making statements in a pretend-as-if context. The phrase "is algorithmically given" makes this clearer than the phrase "can be constructed", but both phrases are unproblematic if one realizes the fictional context. We also note that any literally given text is also algorithmically given (as always in the fictional context).

As already stated, we do not want to make statements on infinitely many input instances. Rather, concerning the input instances, we merely want to make *hypothetical* statements of the form "Suppose that some x_0 is given. Then $A(x_0)$ holds.".

But what shall it mean that the statement is "hypothetical"? Our idea is that the reader always only imagines a particular algorithmically given x_0. But this raises again a question as it is not clear what algorithmic constructions shall be allowed. (It does not help to say "all possible constructions", because we want to discuss what we shall deem as being possible.) The strictest possible solution, the most narrow interpretation which seems to be reasonable, is: Such a hypothetical statement is on all x_0 which have been algorithmically given or will be algorithmically given by any conscious being shall be allowed. Wider, "more fictional" interpretations are possible.

These hypothetical interpretations lead to a corresponding interpretation of the forall-quantifier: We interpret a statement of the form "$\forall x : A(x)$" (including corresponding statements in natural language) as:

> Suppose some x_0 is algorithmically given. Then $A(x_0)$ holds.

Similarly, for "$\exists x : A(x)$" one might then not only say "Some x_0 with $A(x_0)$ can be constructed, but also:

> Some x_0 with $A(x_0)$ is algorithmically given.

Is this interpretations always possible, that is, can one reasonably interpret statements of constructive mathematics as outlined in the previous subsection in this "hypothetical" way? In particular, is the interpretation consistent with (typed) Intuitionistic Predicate Calculus? The author's answer is that this is so; after all the only occurrence of the forall-quantifier in the axiom scheme is the scheme $(\forall x : A(x)) \to A(t)$, and this is nicely consistent with our interpretation.

With the outlined "hypothetical" interpretation of forall-statements, there is no need to imagine infinitely many objects. For this reason, by itself it would seem to be reasonable to call this interpretation "finitism". However, as the term is already used (as far as the authors can tell with rather different meanings) in a similar way as "constructivism" (based on natural numbers), we refrain from doing so. We note that our interpretation is different from ultrafinitism which rejects natural numbers deemed to be "too large".

In our applications, we want to talk about texts and not natural numbers, but to complete the picture and to deepen the link to constructive statements as they are usually made (including in the previous subsection), let us come back to natural numbers now. As said above, the idea of a natural number is always abstract. Moreover, as it is based on the idea of counting, an algorithmic aspect is already present in the notion of natural number. So it is not at all clear what a "non-algorithmically given" natural number should be. For this reason we suggest to just speak of a "given natural number", not an "algorithmically given natural number" and to also say that the number "has been constructed", not just that it "can be constructed". The "hypothetical" interpretation can then also be applied to natural numbers.

To given an example of our interpretation, by saying that there is the rule or algorithm $\mathcal{A} : (a, b) \mapsto a \uparrow\uparrow b$, one says:

> Suppose a tuple of natural numbers (a, b) is given. Then the algorithm $\mathcal{A}$ gives / constructs a particular number (which is then denoted by $a \uparrow\uparrow b$).

If the reader then imagines two given numbers a and b, say $82479824 \cdot 3798217$ and 2418^{92741}, she or he can construct the corresponding number $a \uparrow\uparrow b$, in this case $(82479824 \cdot 3798217) \uparrow\uparrow (2418^{92741})$, and then this number is given / has been constructed. The statement is here that whatever two numbers the reader imagines, she or he can construct the resulting number by application of the algorithm. We note that again that the statement does not have a reasonable physical interpretation.

4.4 Application

We now discuss how we want to apply the previous thoughts to the problems we have posed. For simplicity we explain this for Zermelo-Fraenkel with the axiom of choice set-theory and the corresponding formal system ZFC, but the following can easily be adapted to other kinds of mathematics.

We fix a set-theoretic statement A (in usual language, maybe based on some definitions which themselves are based on the notion of set).

Futhermore, we suppose we are given a statement in set-theoretic language, A^{formal} for which one can convincingly argue that it is a formal reformulation of A.

Let us now consider these statements:

1. One can give a convincing argument for A on the basis of Zermelo-Fraenkel set theory with the axiom of choice.

2. One can give a formal proof for A on the basis of ZFC.

Let us first recall that the "one can give" shall mean that an argument is in fact given, so this statement shall not be made in a fictional sense.

Note that these two statements correspond to the two statements in subsection 2.7. The first statement we do not want to discuss further, because we do not want to enter the discussion what should be considered to be convincing or not. We just recall that we follow the fictionalist approach and interpret the "on the basis of Zermelo-Fraenkel set theory with the axiom of choice" as an introduction to a pretend-as-if situation. The same applies to all other on-the-basis-of statements referring to a kind of mathematics.

So let us come to the second statement. Here we wish to apply what we wrote above. We now postulate that the statement shall mean:

First, a formal set-theoretic statement A^{formal} is algorithmically given for which one can convincingly argue that A is a formal reformulation of A. And then in the sense stated above, a formal proof of A^{formal} in the formal system ZFC is algorithmically given. This means that the "can give" shall be interpreted as "can be constructed". So, here the "on the basis of" refers to the formal system, not the pretend-as-if-situation (as is the case when it refers to a kind of mathematics).

We note and stress that really nothing "real" is being constructed here. Rather, the statements made in the predend-as-if situation of a particular form of constructive mathemtics.

In the next section, we will study the relationship between statements 1. and 2. and the analogous statements for other kinds of mathematics as well as further statements.

In subsection 2.4 we have said that the fictionalist interpretation of mathematical claims does not mean that "doing mathematics is like telliing stories".

Having said this, doing constructive mathematics for our purposes (on "texts") is psychologically close to arguing or "telling stories" about what one "could" (but in fact cannot) do, like humans often do in everyday conversations.

We therefore fix this (meta-mathematical) definition for the next section:

Definition 1 We call a text of constructive mathematics about "formal texts" a *formalistic story.*

We note that this is merely a definition which seems to be pychologically fitting to us; we do not want to enter a discussion on how far such a text really "is" a story or not.

We do, however, stress again that the texts which are allegedly constructed in fact do not exist in any physical sense, and to stress this, one can say that one "merely tells a story about such texts".

5 Mathematics and meta-mathematics

There is an interesting relationship between the use and the theory of formal systems: Formal systems were first introduced (for example by Gottlob Frege) to study and clarify mathematical arguments, and as already stated in subsection 3.2, the idea of formal reasonings (formal proofs) is foundational to the present day doing of mathematics. Now, the idea of formal systems by itself invites a mathematical study on formal systems, and quickly the study of formal systems was integrated into mathematics.

On the one hand now, the mathematical study of formal systems has its motivation in meta-mathematical considerations. Conversely, the mathematical study of formal systems has lead to mathematical results on formal systems which have an impact on how mathematics is viewed and conducted. One example is the application of (not immediately obvious) statements from propositional logic or first-order logic in a convincing argument.

The (historically surprising) statements on the foundations of set theory (incompleteness, independence of axioms, etc.) have a particular, one might say philosophical, impact on the way mathematics is seen and is conducted. (Explicitly, they have lead to a certain "liberal attitude" towards different axiom schemes, where no such scheme is regarded to be "absolutely true".)

It is, however, not obvious in what sense such results make statements *on* mathematics (in particular on the limitations of mathematical endeavors) rather then *in* mathematics.

To study this question, more generally, we study the relationship between arguments ("proofs") as such and arguments on formal proofs. For the latter we consider formalistic stories and model-theoretic proofs. We treat the relationship between arguments and formalistic stories and call the corresponding directions of inference "formalization" and "deformalization". Then we discuss the relationship between arguments or formalistic stories and model-theoretic proofs, calling the corresponding directions "going down" and "going up".

Finally, we study set-theoretic interpretations of kinds of mathematics, using also model-theoretic results. In order to do so, we need the previous results on formalization, deformalization and going down for all kinds of mathematics. For this reason, in comparison to the previous section of this work, we broaden the scope and start immediately with a kind of mathematics.

As mentioned in subsection 3.4, for a kind of mathematics under consideration, we assume that a corresonding formal system is actually given. We can then speak about formal reformulations of statements for the kind of mathematics and formalistic stories.

Let a formal system, say S, be given. This means that an alphabet, a grammar, possibly a criterion on bodies of sentences being interpretable[27], axioms and rules of inference are given. This in turn means that algorithms to judge whether an expression is a sentence, a body of sentences is interpretable, a sentence is an axiom and rules of inference are applied correctly shall be given following our rules of constructive mathematics. Moreover, we demand that the system is so "humanly oriented" that trained mathematicians can – "by hand" – check on small examples whether the rules of the formal system are used correctly and can write down some illustrating examples of formal proofs corresponding to convincing arguments.

We stress that we not want to speak about the language as a set; in fact in line with Section 4, we do not want to use set theory at all when we do meta-mathematics.

Nonetheless, given such a datum, we can easily introduce set-theoretic objects with the same name as the ones for the formal system (like "alphabet", "sentence", "axiom", "rule"), in such a way that one can say that the objects "correspond" to the objects defined by the formal system. Then one can argue set-theoretically about these objects. Of course, all this takes place in the pretend-as-if context of set-theory.

Concretely, there then is a set Σ called *alphabet*. Attached to this set we have the set Σ^* of tuples of arbitrary length on Σ, the elements of which are called *expressions* on Σ; we have the *language* $\mathcal{L}$ (consisting of elements called *sentences*), which is a subset of Σ^*; we have the set of *axioms* Γ, which is a subset of $\mathcal{L}$; there is a set of *rules of inference* and maybe also a set to define interpretable bodies of sentences. This datum is the set-theoretic description of the formal system; we denote by S^{set}. It is then defined, using these sets and set-theoretic constructions, that certain elements of $\mathcal{L}^*$ (the set of tuples of arbitrary length on $\mathcal{L}$) are called *formal proofs* (or *model-theoretic proofs*).

We shall always only consider a single formal system S. For this, we use the following terminology and notations: We call a sentence in the language of S an S-*sentence*. We remark that we only make constructive and never set-theoretic statements on S-sentences. In a set-theoretic pretend-as-if situation or when reflecting on such a situation, we might want to reason about a set associated to a given S-sentence or, more generally, expression in the language of S. In this case, for an expression E, we denote the set by

[27]See subsection 3.2.

E^{set}. We apply this in particular to S-statements A^{formal} which are formal reformulations of statements A in a given kind of mathematics. We note that then A has a meaning in the pretend-as-if context of the given kind of mathematics and $(A^{\mathsf{formal}})^{set}$ is an object in the pretend-as-if context of set theory, that is, a set. To this set we then also assign a meaning.

5.1 Formalization

Recall that in subsection 4.4 we said that to claim that a mathematical statement A is formally proven on the basis of ZFC shall mean that on the basis of constructive mathematics (as discussed in Section 4) one can construct a formal proof of a formal reformulation of A in ZFC. Now, when a mathematician claims that he or she has proven a result, he or she usually just writes down his arguments, which are by default based on Zermelo-Fraenkel set theory with the axiom of choice, but he or she does not say that there is a formal proof in ZFC on the basis of the arguments. With Definition 1 we can say that no formalistic story is told here.

Can one a priori say that every statement convincingly argued for on the basis of Zermelo-Fraenkel set theory with the axiom of choice can be proven on the basis of ZFC? This question is non-trivial, because mathematicians in their judgment if an argument is convincing to them do usually not explicitly consider the question of rewriting.

We answer a the corresponding general question for all kinds of mathematics affirmatively with the following thesis:

Thesis 1 Let some kind of mathematics with a corresponding formal system S, an (informal) statement A in the kind of mathematics and a convincing argument for A (again in the kind of mathematics) be algorithmically given. Then one can construct a formal proof of a formal reformulation of A on the basis of S.

This thesis is in nature similar to the Church-Turing thesis. We stress that it should not be taken as a definition; in fact, the phrase "convincing mathematical argument" (what other authors call "informal proof") already has a meaning, so it cannot be taken as a definition.

We do however want to suggest that if somebody ever came up with a counterexample, the mathematical intuition on when a mathematical argument is seen as convincing (a "proof") or the axiomatic basis would change and afterwards the thesis would again be correct.

We note that a related inference as the one in the thesis is present in Gödel's argumentation for the First Incompleteness Theorem: In the fictional context of set theory, he starts with a formal system S (a set) containing Robinson Arithmetic and uses this implication: If a formal sentence α (also a set) follows from S, a sentence whose interpretation is "α can be proven from S" also follows from S.

Despite this similarity, there are, however, differences between Gödel's technique and what we argue for here: The argument by Gödel is purely mathematical (set theoretic), it is an inference in the fictional context of set theory. Our thesis is however meta-mathematical, it is about the human endeavor concerning mathematics.

5.2 Deformalization

Let us now consider the opposite question which we will study in the general setting from the beginning. So let a mathematical statement A in some kind of mathematics and a construction of a proof of a formal reformulation of A be given. Can one always suppose that this construction (which is – by constructive mathematics – the same as the argument for the construction) leads to a convincing argument for A?

Considering such constructions in full generality, at first view one may doubt that this is always the case: The imagined "output" of a construction, would, if it can be computed at all, lead to an accumulation of symbols which would not be readable by humans. But conversely, the construction has to be readable and understandable by humans in order to be a convincing. Therefore the construction should convince the attentive reader that statement A holds. This leads to the following thesis:

Thesis 2 Given a kind of mathematics and a statement A in the given kind of mathematics, a construction of a formal proof of a formal reformulation of A is a convincing argument for A in the given kind of mathematics.

5.3 Going-down

By the theses of the previous subsections 5.1 and 5.2, for every given statement, convincing arguments should lead to constructions of proofs of formal reformulations, and conversely, constructions of proofs of formal reformulations should immediately be convincing arguments.

Suppose such a construction is given for some statement A of some kind of mathematics – as always with a given formal system. Let A^{formal} be the given formal reformulation of A. Now the given construction of a formal proof gives rise to a convincing argument in Zermelo-Fraenkel set theory to show that there exists a formal proof for $(A^{\mathsf{formal}})^{set}$. This means that the algorithms of the construction of the proof are transferred to set-theoretic constructions which carry out the algorithm. Since classical set theory is more powerful than constructive mathematics as described, such a transformation is always possible. Summarizing, we have:

> Any construction of a formal proof of A^{formal} leads to a convincing argument that there exists a model-theoretic proof of $(A^{\mathsf{formal}})^{set}$.

Assuming Thesis 1 we then obtain the following. As it is potentially weaker than Thesis 1, we formulate it as a separate thesis.

Thesis 3 Let some kind of mathematics with a corresponding formal system, a statement A in the kind of mathematics and a convincing argument for A be algorithmically given. Then one can construct a formal reformulation A^{formal} of A and give a convincing argument for the set-theoretic statement for: There is a model-theoretic proof of $(A^{\mathsf{formal}})^{set}$.

5.4 Going up for existence of arguments

Suppose now that a kind of mathematics (as always with a corresponding formal system S for which in turn we fix a set theoretic realization) and a statement A in the kind of mathematics is given. Using the notation from above, suppose now a mathematician does the following:

1. He or she gives a formal reformulation A^{formal} of A.
2. He or she argues ("proves") set-theoretically (!) that there exists a model-theoretic proof of $(A^{\mathsf{formal}})^{set}$ on the basis of the given set-theoretic realization of S.

One might then be tempted to say that A "is proven" (has been convincingly argued for) on the basis of Zermelo-Fraenkel set theory especially if the kind of mathematics being considered is Zermelo-Fraenkel set theory itself. This would, however, contradict our general attitude that in order to say that a result is formally proven (on the basis of some set-theoretic formal system) one has to constructively argue that there is a formal proof. Explcitly, this would contradict the postulate in subsection 4.4. We note that this critique is independent of the use of the axiom of choice.

We will come back to the relationship between the two statements 1. and 2. in subsection 5.6.

5.5 Going up for non-existence of arguments

We now consider a statement on the non-existence of formal proofs for statements. Examples for this are many, let us state one: One can argue model-theoretically in Zermelo-Fraenkel set theory without the axiom of choice that if $(\mathsf{ZF})^{set}$ is consistent, then $(\mathsf{ZF})^{set}$ with $(\neg\mathsf{C})^{set}$ is also consistent, that is, that the axiom of choice does not follow from $(\mathsf{ZF})^{set}$. This result is particularly interesting as the known arguments for it rely on non-trivial set theory.

We want to "go up" and conclude something on the limitations of humans and more generally on conscious beings of doing mathematics. Let us stick with the given example. Here we would like to conclude: Any convincing argument on the basis of Zermelo-Fraenkel set theory that the axiom of choice holds would lead to a convincing argument that ZF is inconsistent.

It would be a misconception to immediately conclude from a mathematical statement on a (seemingly) corresponding statement on the limitations of humans. One can however argue by "going down":

Assume that someone gave a convincing argument for the axiom of choice on the basis of Zermelo-Fraenkel set theory. Then assuming that Thesis 3 holds for ZF, it would be proven on the basis of Zermelo-Fraenkel set theory that there is a model-theoretic proof of the axiom of choice. This would be a contradiction. So we would then derive a contradiction in Zermelo-Fraenkel set theory, in one word, we could convincingly argue that Zermelo-Fraenkel set theory without the axiom of choice is inconsistent. Moreover, if Thesis 1 holds in this situation, this would then mean that one can construct a formal proof of a contradiction on the basis of ZF, showing that the formal system ZF is inconsistent.

5.6 Kinds of mathematics interpreted in set theory

We now consider the relationships between statements in kinds of mathematics and corresponding set-theoretic statements. For this, throughout this subsection, we fix some kind of mathematics, based on a formal system S, and a statement A inside this kind of mathematics.

5.6.1 Convincing arguments and formalistic stories in a kind of mathematics

We have the following statements:

(Arg) One may give a convincing argument for A.

(FS) One may give a formalistic story showing A.

(Arg) is a statement on human knowledge concerning the given kind of mathematics and (FS) is a statement on human knowledge concerning constructive mathematics. The "one may give" shall be taken literally and not be interpreted in a fictional way. So we expect in (Arg) and (FS) that such stories are actually given.

Since these statements refer to different kinds of mathematics (if the considered kind of mathematics is not constructive mathematics itself), we do not have a common formal basis to establish formal relationships between them. Nevertheless, by theses 1 and 2 in subsections 5.1 and 5.2, (Arg) and (FS) may be considered as informally equivalent.

5.6.2 Models and validity

We now consider what is called *semantic of formal systems* or *model theory* and with this relationships between statements on models of the given kind of mathematics and similar objects.

We first note that by doing model theory we argue in a set-theoretic pretend-as-if situation. We stress in particular again that we do not want to consider set-theoretic arguments when doing meta-mathematics, and in fact consider it inappropriate to use set theory in meta-mathematics. Nonetheless, we can consider from an outside point of view the doing of set theory and therefore in particular of model theory.

For this, we suppose that the formal system S, on which the kind of mathematics under consideration is based, is a system of first-order (classical) logic.[28]

We keep the notation introduced in the introduction to this section, so σ is the set interpreted as the alphabet, $\mathcal{L}$ is the set interpreted as the language and Γ is the set of axioms.

We consider models of Γ (and Σ), which we also call *models of* S in order to emphazise the relationship with the kind of mathematics defined by S.

In concrete examples, the models are objects as one expects them to be. For example, a model of ring theory is a ring.[29] [30]

[28]We are confident that the following considerations can be adapted to other kind of formal systems, but for accuracy and because such a generalization would be a topic for itself, we stick to first-order systems.

[29]One might ask here: Is a model of a ring not "a bit more than just a ring"? After all, before one even defines a model, one has a set of function symbols and a set of relation symbols. And then, to define a model, one not only fixes functions and relations but also maps from the sets of function symbols to the set of functions and from the set of relation symbols to the set of relations. This is, however, not the case, as can be seen with the definition of a ring: A ring can be defined as a tuple $(R, \cdot, +, 0, 1)$, so in partulcar the functions $+$ and $\cdot$ as well as the constants (0-arity functions 0 and 1) have to be ordered; it is incorrect to just consider just a set $\{+, \cdot 0, 1\}$ (or two sets $\{+, \cdot\}, \{0, 1\}$). So in fact, in the definition of a ring as a 5-tuple, the functions are indexed by natural numbers. Now, the set of function symbols can be seen as an index set. In our situation, as S is constructively given, there are only at most countably many function and relation symbols, and then it is natural to identify these with consecutive natural numbers starting with 1. Then models of rings are essenially rings defined as indicated. In any case, a model of rings does not contain superfluous data.

There is, however, a difference between the usual definition of a ring and the definition as a model for Γ: For rings, one writes down the axioms directly, that is, one does not consider a set of axioms. But this difference is not relevant concerning the extension of the definitions. In general, as the axioms are – by assumption – algorithmically given by the formal system S, one can get rid of the set of axioms if the axiom scheme in S is finite.

[30]There are often competing "natural" set-theoretic definitions concepts of an intuitive

We can now apply the usual model theoretic definition to the set of sentences Γ and the sentence $(A^{\mathsf{formal}})^{set}$.

Let us clarify an important point here: To do model theory means to do set theory. So it always takes place in a set-theoretic pretend-as-if context. But this is not all: one also uses inside of set-theory set-theoretic definitions to mimic the already given notions of formula, truth, implication etc. Care should be taken not to confuse the internal definitions with the a notions which are already present by introducing the pretend-as-if context.

We have the following statement inside Zermelo-Fraenkel set theory:

(MT-Validity) $(A^{\mathsf{formal}})^{set}$ is logically valid with respect to Γ.

Let us clarify this statement: $(A^{\mathsf{formal}})^{set}$ is the set in $\mathcal{L}$ associated to A. For the definition of logically validity, it is necessary to define which statements are "true" and which are "false". This is defined by a set-theoretic recursive definition over the length of the formulas. That is, one defines in a recursive manner a function on the set of statements with values in a set of cardinality two, whose values are interpreted as "true" and "false".

Every given S-sentence S rives rise to a ZFC-formal property of models of Γ. Let us denote this property by S^{MT}. (The models correspond to a free parameter in the statement.) $(A^{\mathsf{formal}})^{\mathsf{MT}}$ is then a formal property of models of Γ, and (MT-Truth) can be reformulated as:

(MT-Truth) $(A^{\mathsf{formal}})^{\mathsf{MT}}$ holds for any model of S.

The equivalence of (MT-Validity) and (MT-Truth) might on first thought seem to be obvious, but maybe on second thought it is not so obvious anymore. For this reason and also because we will consider a variant of the equivalence later, we give a detailed argument in the next subsection.

Furthermore, given A, in mathematical praxis one may observe that one has a corresponding informal property A^{MT} of models which may literally coincide with A. A requirement is here that A^{MT} shall be an informal formulation of $(A^{\mathsf{formal}})^{\mathsf{MT}}$.

It is then natural to consider:

(Informal-MT-Truth) A^{MT} holds for any model of S.

One should expect that (Informal-MT-Truth) is equivalent to (MT-Truth).

concept. So it is anyway not so important if what is called "model" (for example model of a ring) is exactly the same as one obtains with a particular set-theoretic definition (for example a particuar set-theoretic definition of the notion of ring). We note here that it is difficult to describe and analyse these choices with a theoretical framework, and we will also not do so.

Example 2 We illustrate these considerations by the example of groups. We start with a statement A on groups which are considered as abstract objects independent of set theory in their proper kind of mathematics. This informal statement A gives rise to a formal statement A^{formal} on groups in the corresponding formal system S. Since S may be considered in set theory, we obtain a set $(A^{\mathsf{formal}})^{set}$ as described above. $(A^{\mathsf{formal}})^{\mathsf{MT}}$ is a statement on models of groups, but this means that it is a statement on groups. Furthermore statement A as a statement on abstract groups may also be formulated as a statement A^{ST} on set-theoretic groups. As mentioned above, $(A^{\mathsf{formal}})^{\mathsf{ST}}$ is now a formal reformulation of A^{MT}.

5.6.3 Result on truth and model theoretic validity

In this section, we argue for the equivalence of (MT-Validity) and (MT-Truth) claimed in subsection 5.6.2.

For this, we first recall some usual set-theoretic definitions (definitions in the pretend-as-if context of set theory) of model theory.

Definition 3 Let an alphabet (which is a set), a language $\mathcal{L}$ and a set of axioms Γ be given.[31] Let M be a model of Γ. Then we define the *truth value* function t_M: For a property / formula (also being a set!) φ with free variables contained among $x_1, \ldots, x_n$ and a tuple $(m_1, \ldots, m_n)$ from M, $t_M(\varphi; m_1, \ldots, m_n)$ is defined as 0 or 1 by induction on the length of φ (with variable $m_1, \ldots, m_n$) in the obvious intuitive way.

We give two examples for the inductive definition, on which we concentrate also in the rest of this subsection.

1. Let $\varphi = \forall x_{n+1} : \psi(x_{n+1})$.

   Then we define

$$t_M(\varphi; m_1, \ldots, m_n) := \min_{m_{n+1} \in M} t_m(\psi; m_1, \ldots, m_{n+1}) \ .$$

2. Let $\varphi = \varphi_1 \vee \varphi_2$.

   Then we define

$$t_M(\varphi; m_1, \ldots, m_n) := \max\{t_M(\varphi; m_1, \ldots, m_n), t_M(\varphi; m_1, \ldots, m_n)\} \ .$$

If φ is a closed formula, that is, a formula without variables, we then have the truth value $t_M(\varphi)$.

[31] Here "be given" is merely a phrase in the considered kind of mathematics, which is set theory; it does have a constructive meaning. It just means: "For all ..., the following holds:".

Definition 4 Let φ be a formula (again a set) with free variables among $x_1, \ldots, x_n$. We say that φ is *logically valid* with respect to Γ if for all models M, and all $m_1, \ldots, m_n \in M$, $t_M(\varphi; m_1, \ldots, m_n) = 1$.

Let us now suppose that a formal system S and an S-property φ is algorithmically given. (An S-property (or formula) is defined like an S-sentence only that it can have free variables.)

Let the free variables of φ occur under the first n variables, $x_1, \ldots, x_n$.

We then obtain a ZFC-property φ^{MT} on models M and elements $m_1, \ldots, m_n$. (Formally, φ^{MT} has a free variable for the model and free variables for the elements, but we can interpret as indicated).

Theorem of Meta-Mathematics 1 *Let a first order formal system* S *and an* S*-sentence* φ *be algorithmically given, where the free variables of* φ *are among* $x_1, \ldots, x_n$*. Then in the fictional context of set theory (and with the notation from above) the following are equivalent:*

(1) *(MT-Validity)* φ^{set} *is logically valid with respect to* Γ*, that is, the following holds:*
For every model M *of* Γ *and all elements* $m_1, \ldots, m_n$ *of* M*,* $t_M(\varphi^{set}; m_1, \ldots, m_n) = 1$*.*

(2) *(MT-Truth)* φ^{MT} *holds for every model of* S*, that is, the following holds:*
For every model M *of* S *and all elements* $m_1, \ldots, m_n \in M$*,* $\varphi^{\mathsf{MT}}(M; m_1, \ldots, m_n)$ *holds.*

We remark that we make here only a statement on (constructiely) given S and φ, so in particular we do not make infinitely many statements, avoiding the possible mistake mentioned at the beginning of the note.

Argument for the Theorem. Let S be algorithmically given. For any given φ with free variables only occuring among $x_1, \ldots, x_n$, we have to show the following set-theoretic statement:

For all models M of Γ, and $m_1, \ldots, m_n \in M$,

$$\varphi^{\mathsf{MT}}(M; m_1, \ldots, m_n) \longleftrightarrow t_M(\varphi^{set}; m_1, \ldots, m_n) = 1 \,.$$

We argue by induction on the length of φ, where the (meta-mathematical) induction hypthesis is:

For all given S-properties ψ of length less than the length of φ, the corresponding statement for ψ has been proven.[32]

[32]The reader should note that this is a valid meta-mathematical statement in our constructive framework.

Let some φ as described by given.

Let us consider as examples the two cases already introduced.

1. $\varphi = \forall x_{n+1} : \psi$.

   By induction hypothesis, the corresponding statement for ψ has been proven.

   Let M be a model of S and let $m_1, \dots m_n \in M$. Then the following are equivalent:

   - $\varphi^{\mathsf{MT}}(M; m_1, \dots, m_n)$
   - $\forall m_{n+1} \in M : \psi^{\mathsf{MT}}(M; m_1, \dots, m_{n+1})$
   - $\forall m_{m+1} \in M : t_M(\psi^{set}; m_1, \dots, m_{n+1}) = 1.$
   - $t_M(\varphi^{set}; m_1, \dots, m_n) = 1$

2. $\varphi = \varphi_1 \vee \varphi_2$.

   By induction hypothesis, the corresponding statements for φ_1 and φ_2 have been proven.

   Let M be a model of S and let $m_1, \dots m_n \in M$. Then the following are equivalent:

   - $\varphi^{\mathsf{MT}}(M; m_1, \dots, m_n)$
   - $\varphi_1^{\mathsf{MT}}(M; m_1, \dots, m_n)$ or $\varphi_2^{\mathsf{MT}}(M; m_1, \dots, m_n)$
   - $t_M(\varphi_1^{set}; m_1, \dots, m_n) = 1$ or $t_M(\varphi_2^{set}; m_1, \dots, m_n) = 1.$
   - $t_M(\varphi^{set}; m_1, \dots, m_n) = 1$

5.6.4 Set-theoretic realizations

The consideration of all models is a particular set-theoretic realization of a kind of mathematics: One might use other set-theoretic definitions still describing intuitively all the objects under consideration. This was already addressed in Footnote 30. A special case is to consider just a particular class of models. This means that all objects under consideration are models. We discuss only this special case.

Example 5 Let a kind of synthetic geometry be algorithmically given, for example basic affine geometry. This can be described by a formal system with one relation symbol of arity 1 (“point or line”) and one one of arity 2 (“contained in”) and no function symbol. The underlying set of the model then consists of points and lines, and there is an incidence relation between these.

Now, one does not need to consider such (set-theoretic) geometries in full generality but one can assume that a line is a set of points, and indeed one often does. A statement holds for all these structures if and only if it holds for all models. (The reason for this is that every model is isomorphic to such a model with a appropriate notion of isomorphism, but this is not relevant for the following.)

In addition, instead of considering models as first described, one would usually start with a set of points and a set of lines. Similarly, if one considers lines as sets of points, one would start with the set of points and not consider the set of points and lines. So, one would consider different set-theoretic structures capturing the idea of an affine geometry. These modifications are not considered in the following, because they seem to be hard to put in a common framework.

Motivated by this we define (meta-mathematically):

Definition 6 By a *set-theoretic realization* of a kind of mathematics (based on a formal system) we mean a description, call it D, of further specified models of the kind of mathematics such that for every statement A of the kind of mathematics, A^{MT} holds for every model of S if and only if A^{MT} holds for every model described by D.

We now consider such a realization. We then consider what we call (STR):

(STR) A^{MT} holds in any model in the realization.

We might abbreviate this statement as:

(STR) A holds for the given realization of S.

By definition, (STR) is equivalent to (Informal-MT-Truth), which in turn should be equivalent to (MT-Truth).

5.6.5 Arguments in kinds of mathematics and set-theoretic arguments

We can now consider the corresponding statements on convincing arguments and formalistic stories. For (STR) these are:

(Arg-STR) One may give a convincing argument for (STR).
(FS-STR) One may give a formalistic story for (STR).

Here in (FS-STR), one uses the formal reformulation $(A^{\mathsf{formal}})^{\mathsf{MT}}$ of A^{MT}.

The statements (Arg-MT-Validity) and (FS-MT-Validity) as well as (Arg-MT-Truth) and (FS-MT-Truth) are defined in an analogous manner. By the considerations of 5.1 and 5.2, the statements (Arg-STR) and (FS-STR), (Arg-MT-Validity) and (FS-MT-Validity) as wel as (Arg-MT-Truth) and (FS-MT-Truth) may be considered to be informally equivalent.

As said in the previous subsection, we stipulate that (STR) and (MT-Truth) are equivalent in the fictional context of set theory, so that (Arg-STR) is equivalent to (Arg-MT-Truth) and to (Arg-MT-Validity). There should also be an equivalence between (FS-STR), (FS-MT-Truth) and (Arg-MT-Validity) since it should be possible to implement this correspondence in formalistic stories.

One has two more implications which one usually would accept in mathematical practice: The statement (Arg) implies (Arg-MT-Truth) and (Arg-STR). We illustrate this again by the example of basic affine geometry.

Example 7 Suppose we have some argument that a statement on points and lines in affine geometry (independently of set theory) holds. Then usually one would say that this statement also holds for all models of such geometries, which are structures with an underlying set partitioned into a set of points and a set of lines, that is, they are point-line-geometries. This statement on geometries is then equivalent to a corresponding statement on geometries where the lines are given by sets of points.

These considerations on arguments carry over to formalistic stories. Given a formalistic story on points and lines (based on a formal system for affine geometry), one also expects that one may modify it such that one obtains a formalistic story for affine geometries in set theory. This supports that (FS) implies (FS-MT-Truth) and (FS-ST).

5.6.6 Model theory

We now consider the well known theorem on soundness and completeness of first-order logic, which are at the center of model theory. Of course, this is a theorem in the usual set-theoretic ficational context.

> Let $\mathcal{L}$ be a first-order language. For all sets Γ of sentences (of $\mathcal{L}$) and all sentences α (of $\mathcal{L}$) the following are equivalent:
>
> (1) There exists a formal proof (within some proper deductive system) which shows that Γ implies α.
>
> (2) The formula α is logically valid with respect to Γ.

The implication (1) $\rightarrow$ (2) is called *soundness* and the reverse implication (2) $\rightarrow$ (1) is called *completeness*. By these equivalences, (MT-Validity) is equivalent to:

(MT-Proof) There exists a formal proof which shows that Γ implies $(A^{\mathsf{formal}})^{set}$.

Here we claim the abstract existence of a formal proof, which – as already stated – is regarded as a set. This is crucial, so in particular one cannot reformulate the statement in a similar way as one can reformulate (MT-Validity) by (MT-Truth) and then by (STR).

We now consider again the corresponding statements on convincing arguments and formal proofs:

(Arg-MT-Proof) One may give a convincing argument for (MT-Proof).
(FS-MT-Proof) One may give a formalistic story for (MT-Proof).

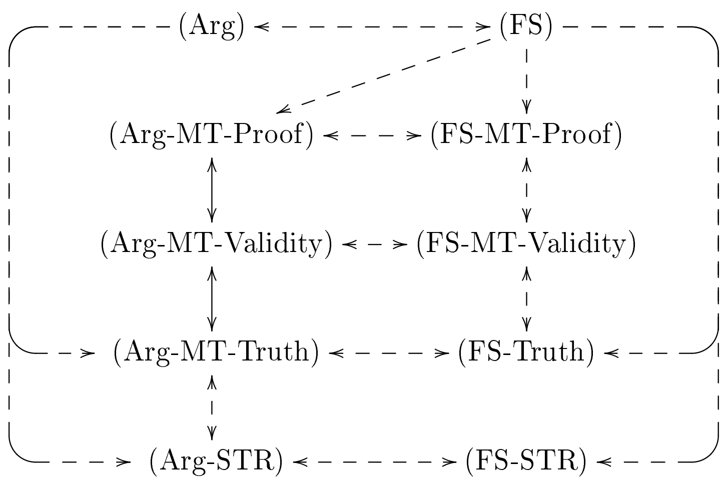
Now, an application of the theorem of soundness and completeness of first-order logic gives an equivalence between (Arg-MT-Validity) and (Arg-MT-Proof). Since this application is a fixed step, not depending on any translations, this equivalence may be regarded as a stronger relationship than most of the former correspondences.

By using subsection 5.6.1, we obtain that (Arg-MT-Proof) and (FS-MT-Proof) should be considered informally equivalent.

Moreover, if one has access to a formalistic story for the (set-theoretic results of) soundness and completeness, one can show that (FS-MT-Validity) and (FS-MT-Proof) are (directly and not just via the theses) equivalent. We argued in subsection 5.3 that clearly (FS) implies (Arg-MT-Proof). In the same vein, (FS) implies (FS-MT-Proof).

5.6.7 Summary on kinds of mathematics interpreted in set theory

We sum up the direct implications mentioned so far in the following diagram:



The drawing of a solid arrow line between (Arg-MT-Proof) and (Arg-MT-Truth) reflects that this relies solely on the fact that a particular set-theoretic result has been established. A solid arrow is used between (Arg-Validity) and (Arg-MT-Truth) is used because of Theorem of Meta-Mathematics 1.

For all the other implications we have strong arguments, but we can differentiate here: The strongest dotted lines are arguably from (Arg) to (Arg-MT-Truth) or (Arg-ST). In fact, essentially nothing seems to be necessary here in order to obtain this transformation.

The dotted lines between the (FS...)'s and from (FS) to (Arg-MT-Proof) also seem to be unproblematic, but some work to really write down the algorithms would have to be done. Here one can imagine that the arrows from (FS) to (FS-MT-Proof) and between (FS-MT-Proof), (FS-MT-Validity) as well as between (FS-MT-Validity) and (FS-Truth) as "nearly solid" because it does indeed seem to be possible to write down the corresponding algorithms.

The middle dotted lines are the weakest inferences; the statements seem to be equivalent as expressed by Thesis 1 and Thesis 2.

5.6.8 Models of ZFC and existence of proofs

We close this section by considering the implications of the previous considerations if applied to ZFC as a formal system.

So now models of ZFC are considered. In this application we have as a special case this informal inference: An argument for A in Zermelo-Fraenkel set-theory gives an arguments for $(A^{\mathsf{formal}})^{\mathsf{MT}}$ being logically valid in all models of ZFC and also for A^{MT} being valid in all models of ZFC. Further, to have such an argument is equivalent to having an argument for: There exists set-theoretically a formal proof of $(A^{\mathsf{formal}})^{set}$.

Here we are back at the warning in subsection 5.3. We stress again the importance to not apply the "convenient" implication from the latter three statements to the former.

Critique of a statement by G. Kreisel. Corresponding to the statement above, we criticise an opposite statement made by G. Kreisel in [informal-rigour]: On page 154 he argues exactly as we deem to be incorrect.

Let us examine closer what Kreisel does and where the problem lies.

For a formal sentence α in some formal system (in which all sentences are interpretable, but not necessarily a system of first order), he defines three notions:

- Val α: α is intuitively valid
- $V\alpha$: α is valid in all set-theoretic structues

- $D\alpha$: α is formally derivable by means of some fixed (accepted) set of formal rules.

He states that “implicitly”, $D\alpha$ implies Val α.

He also refers to Val α as α being “logically valid” and states that “the moment one takes it for granted that logic applies to mathematical structures, Val α implies $V\alpha$”.

He also says that for a set-theoretic statement α, one does not immediately conclude that $V\alpha$ implies α.

But then he argues for a statement of first order that by the completeness theorem, $V\alpha$ implies $D\alpha$, which then means that $V\alpha$ also implies Val α. So in total, according to Kreisel, all three asserions are equivalent.

What is going on here?

Upon second look one notices that it is not so clear what the three assersions Val α, $V\alpha$ and $D\alpha$ shall actually mean.

Two consistent interpretations suggest themselves:

1.
   - Val α means: In the fictional context of the considered kind of mathematics, α holds.[33]
   - $V\alpha$ means: In the fictional context of set theory, α holds in all models of the given kind of mathematics; this is our (MT-Truth).
   - $D\alpha$ means: In the fictional context of set theory, there exists a formal proof of α; this is our (MT-Proof).

2.
   - Val α means: In the fictional context of the considered kind of mathematics, one can convincingly argue that α holds; this is our (Arg).
   - $V\alpha$ means: In the fictional context of set theory, one can convincingly argue that α holds in all models of the given kind of mathematics; this is our (Arg-MT-Truth).
   - $D\alpha$ means: One can construct a formal proof of α; this is our (FS).

It seems to us that for Val α and $V\alpha$, Kreisel had more or less the first interpretation in mind (maybe with a “platonistic flavour”, which is, however, irrelevant for our purposes.). But there is some indication that for $D\alpha$ he had the second interpretation, i.e., (FS), in mind. An indication is the statement that “implicitly”, $D\alpha$ implies Val α. Another one is a discussion on the use of rules by Bourbaki.

[33]For this, it would be better to consider an informal reformulation of α.

If one follows the first interpretation, $D\alpha$ is indeed equivalent to $V\alpha$. But this takes place in the fictional context of set theory, and to go to Val α one would possibly have to jump to another context. In the special case that the formal system itself is ZFC, one can stay in the set-theoretic context, but then – as discussed above – there is no reason to think that one can conclude as indicated. This is then an open problem.

If one follows the second interpretation, by our theses, Val α should be equivalent to $D\alpha$. Moreover, Val α should imply $V\alpha$, but from $V\alpha$, one then cannot – by current knowledge – obtain $D\alpha$.

Critique of a statement by H. Field. We also criticise a corresponding statement by H. Field in the Introduction to [real-math-mod] (in Part Two, Section 5, called "Logical Implication").

Field asks "What should a fictionalist say about such metalogical notions of logical implication and logical consistency?"

He starts with "a set of sentences", Γ, and "a sentence", α.[34] Here, it is implicitly assumed that these are sentences in a formal system. We note that already this set-up is problematic for an analysis of logical implication from a fictionalist point of view, because sets are presupposed. As argued, one should rather say that Γ and α are algorithmically given.

Field is interested in the question of what it shall mean to say that "Γ implies α".

We note here that one can reformulate this question by "moving Γ to the formal system"; it then becomes the question discussed by Kreisel (and also by us): When shall it mean to say that one has proven α?

Field mentions Tarki's definition of logical implication, which is based on models. (Γ Tarski-implies α if and only if α is logically valid in any model of Γ.) Obviously, this is a definition in a set-theoretic pretend-as-if situation. In particular, it is not discussed what it shall mean that one has established such an implication.

Field discusses Kreisel's work, in particular Kreisel's claim of the equivalence of Val α, $V\alpha$ and $D\alpha$.

He then writes (with "B" changed to "α"):

> What the completeness theorem for a given formalization does is prove (platonistically) that if Γ Tarksi-implies α, then Γ derives α; this together with the previous chain shows that all three notions coincide extensionally for first order logic.

Without the insertion "(platonistically)", the first statement is unproblematic inside the fictional context of set theory. But the second part of the

[34] α is denoted by B in [real-math-mod].

statement indicates that not such an "internal" set-theoretic point of view shall be taken.

Then already the predicates "Tarksi-implies" and "derives" are problematic, just as the phrase "is valid" would be problematic. After all, the discussoion does not take place inside a fictional context; rather it is a discussion on statements in fictional contexts.

What then shall the inserted "(platonistically)" mean here? To argue "platonistically" about mathematical objects is in practice the same as to argue in some fictional context. This immediately rises the question: in which fictional context?

Given that one talks about formal proofs here, it would be natural if the fictional context of Peano arithmetic (building on classical logic) was used (where a "text" simply was the same as a number.) We note that because of the use of classical logic this context would already allow for arguments which are not valid with our constructive approach. For example, in contrast to the constructive approach, they would allow for an argument of the form "We saw that the assumption that there is no proof leads to a contraction, therefore we have a proof.".

But Peano arithmetic is also not sufficient here: one has to presuppose set-theory. How shall this then be combined with the given kind of mathematics? A clean possibility is to say that "everything is a set", that is, one just does set-theory. But then Val α is actually identical to $V\alpha$.

We note here the following general point: a lot can be hidden in a "platonistic" argument. Likewise, inside the "right" fictional context, one can argue for "just about anything". And as stated, to argue on mathematical objects in some fictional context is the same as to argue platonistically from a particular Platonistic point of view.

Rhetorically, it does, however, make a big difference if one writes "platonistically" or "fictionally", as with the latter it is immediately clear that one should ask: in which fictional context? The reader can test this him- or herself by considering Field's sentence quoted above and substitute "platonistically" with "fictionally".

5.7 Existence of proofs and class theories

A proof attempt. Inspired by the ideas of subsection 5.6.3, we now make an attempt to argue for this statement:

> For any algorithmically given ZFC-sentence α, in the fictional context of Zermelo-Fraenkel set theory, the implication
>
> $$(\exists \text{ Proof of } \alpha^{set}) \to \alpha$$
>
> holds.

Here is the attempt:

In the fictional context of set theory, we do the following:

We define a "truth-value"-function t with values in $\{0,1\}$ which assigns to every formula φ (which is now a set) with free variables among $x_1,\ldots,x_n$ and all sets $m_1,\ldots,m_n$ the truth value of φ evaluated at $m_1,\ldots,m_n$. Just as in 5.6.3, we proceed by induction on the length of the formula.

We now have (still inside the fictional context of set theory):

For every formula φ (a set) with variables among $x_1,\ldots,x_n$ for which there exists a proof of $\forall x_1,\ldots,x_n : \varphi$, the following holds:

$$\forall m_1,\ldots,m_n : t(\varphi; m_1,\ldots,m_n) = 1$$

We leave the fictional context of set theory and argue meta-mathematically.

Let a ZFC-formula φ be algorithmically given, where the variables are among $x_1,\ldots,x_n$. Then we have in the fictional context of set theory:

1. If there exists a proof of $(\forall x_1,\ldots,x_n : \varphi)^{set}$, then for all $m_1,\ldots,m_n$, $t(\varphi; m_1,\ldots,m_n) = 1$.
2. For all $m_1,\ldots,m_n$,

$$\varphi(m_1,\ldots,m_n) \longleftrightarrow t(\varphi^{set}; m_1,\ldots,m_n) = 1\ .$$

We have just established the first item. The argument for the second one is just as the one for Theorem of Meta-Mathematics 1.

Putting 1. and 2. together and applying it to a sentence α, we obtain the desired result.

Critique. There is the following problem with the approach: The alleged "truth-value"-function t would be – if it existed – not a function in the usual sense of set theory because the domain would not be a set. Now, one can indeed give assignments which are not functions just by algorithmically giving the formula. However, here we want to define the "function" t by induction. This means that inside set theory we want to argue that there exists this "function". This does not square.[35]

A meta-mathematical variant. What one can do is to argue meta-mathematically: For every algorithmically given formula φ, one can give a ZFC-formula t_φ such that in Zermelo-Fraenkel set theory, t_φ defines an

[35] In the Introduction to [real-math-mod], Section 5 (on page 31) H. Field tries to argue as we criticize here. Indeed, he writes "Suppose for instance that Γ is the set of all truths of set theory." The question is: How does one define the alleged set?

assignment and the following holds: If the free variables of φ are among $x_1, \ldots, x_n$:

$$\forall m_1, \ldots, m_n : \varphi(m_1, \ldots, m_n) \longleftrightarrow t_\varphi(m_1, \ldots, m_n) = 1 \, .$$

If one uses this formula and its set-theoretic interpretation instead of the "complete" "truth-value"-function t, one obtains the following result on formalistic stories. (We use formalistic stories instead of convincing arguments because we want to avoid Thesis 2; with the thesis the result can be reformulated for convincing arguments.)

Theorem of Meta-Mathematics 2 *Let a set-theoretic statement A be given, and let furthermore a formal reformulation A^{formal} be algorithmically given. Suppose*

1. *One can algorithmically give a constant C.*
2. *One can construct a formal proof that there exists a proof in Zermelo-Fraenkel set theory with the axiom of choice that for the given C that there exists a ZFC-proof of $(A^{\mathsf{formal}})^{set}$ of length at most C.*

Then one can construct a formal proof of A^{formal} and therefore give a formal proof of A.

Actually, in 2. the immediate condition is that the length of all formulae should be bounded by such a constant C. But this is equivalent to the given condition.

A variant in Neumann-Bernays-Gödel set theory. Another possibility is to use Neumann–Bernays–Gödel set-and-class theory. So let us argue in the fictional context of this theory:

We define by induction on a natural number k a class $\mathcal{T}_k$ such that for a formula (a set) φ of length k and variables among $x_1, \ldots, x_n$ as well as sets $m_1, \ldots, m_n$, the condition $(\varphi; m_1, \ldots, m_n) \in \mathcal{T}_k$ captures the intuitive condition that φ applied to $m_1, \ldots, m_n$ holds.

The definition of $\mathcal{T}_k$ is completelely analogous to the definition of t_M for a model M.

Now, by the union axiom, we can form the class $\mathcal{T} := \bigcup_{k \in \mathbb{N}} \mathcal{T}_k$. One can then work with the class $\mathcal{T}$ as a substitute for a "truth-value"-function. Like this, one obtains:

Theorem of Meta-Mathematics 3 *Let a set-theoretic statement A be given and let a formal reformulation A^{formal} of A be algorithmically given. Then the following statements are equivalent:*

(1) *(Arg-MT-Truth) One can convincingly argue that A^{MT} holds for every model of* ZFC.

(2) *(Arg-MT-Proof) One can convincingly argue in Zermelo-Fraenkel set theory with the axiom of choice that there exists a* ZFC*-proof of A^{formal}.*

(3) *One can convincingly argue in Neuman-Bernays-Gödel set theory with the axiom of choice that there exists a* ZFC*-proof of A^{formal}.*

(4) *One can convincingly argue in Neuman-Bernays-Gödel set theory that A holds if interpreted for sets.*

(5) *One can convincingly argue in Zermelo-Fraenkel set theory that A^{MT} holds for every model of* NBG *(where again the quantifiers are restricted to the objects called "sets" in the model).*

We have just established (3) $\rightarrow$ (4). The implication (1) $\rightarrow$ (2) was already argued for, the implication (2) $\rightarrow$ (3) is trivial. The statement (4) $\rightarrow$ (5) is an application of the general principal, argued for above, that (Arg) $\rightarrow$ (Arg-MT-Truth).

The implication (5) $\rightarrow$ (1) is an application of the well-known result that NBG is a conservative extension of ZFC, which is the following statement in the fictional context of set theory:

> For every ZFC-sentence α (a set), the following holds: If α holds for every model of NBG, then it also holds for every model of ZFC.

We stress here that this implication is really the statement of the conservative extension theorem. So, one cannot conclude from (4) that A holds.

Of course, if one assumes that Thesis 1 and Thesis 2 hold, the statements can be substituted by their corresponding formalisitic statements.

A variant in Morse-Kelley set theory. Morse-Kelley set-and-class theory differs from Neuman-Bernays-Gödel set-and-class theory by a more general class comprehension axiom: Classes exist also for formulae whose free variables range over classes and not just sets. One then obtains:

Theorem of Meta-Mathematics 4 *Let a set-and-class-theoretic statement A be given and let a formal reformulation A^{formal} of A be algorithmically given. Then the following holds in Morse-Kelley set-and-class theory: If there exists an* MK*-proof of A^{formal}, then A holds.*

6 Discussion: Is "constructive" strong enough?

We have argued that claims that formal proofs "are given" for mathematical statements should be interpreted with constructive mathematics, and we have given an interpretation of the statements of constructive mathematics for which one does not need to imagine infinitely many natural numbers or (in our application) texts.

However, the alleged formal proofs often cannot be written by humans. So the question arises: Is the constructive interpretation of such a claim strong enough?

The answer then depends on whether one regards the statements of constructive mathematics in our interpretation as being "less imaginary" than set-theoretic statements.

In any case, it seems to the authors that with the approach presented, the fictional setting is the closest to actual physical writing that one might hope for if one wants to uphold claims that formal proofs "are given" even in cases when they in fact cannot be written or it would be very inconvenient if they were written.

References

[fictional] M. Balaguer, Fictionalism in the Philosophy of Mathematics, Stanford Encyclopedia of Philosophy, 2011

[foundations] M.J. Beeson, Foundations of Constructive Mathematics, Metamathematical Studies, Springer-Verlag, 1985

[constr-ana] E. Bishop and D. Bridges, Constructive Analysis, Springer-Verlag, 1985

[constructivism] D. Bridges and E. Palmgren, Constructive Mathematics, Stanford Encyclopedia of Philosophy, 1997

[formal-natural] M. Carl, Formal and Natural Proof – A phenomenological approach, 2015, Preprint, http://www.math.uni-konstanz.de/ carl/Paper/FormalNaturalProofFinal2015.pdf

[Zahlen] R. Dedekind, Was sind und was sollen Zahlen?, second edition, Vieweg, 1893, https://archive.org/details/wassindundwasso00dedegoog

[constr-theories] S. Feferman, Constructive theories of sets and classes, in: M. Boffa, D. van Dalen and K. McAloon (eds.)

Logic Colloquium '78, pp. 159 – 224, North-Holland, 1979

[real-math-mod] H. Field, Realism, Mathematics & Modality, Basil Blackwell, 1989

[dialectica] K. Gödel, Über eine bisher noch nicht benutzte Erweiterung des finiten Standpunkts, dialectica, **12**, pp. 280 – 287, 1958

[formalized] J. Harrison, Formalized Mathematics, Technical Report 36, Turku Centre for Computer Science, Technical Report 36, 1996

[informal-rigour] G. Kreisel, Informal rigour and completeness proofs, Studies in Logic and the Foundations of Mathematics, **47**, 138 – 186

[fictional-entities] F. Kroon and Alberto Voltolini, Fictional Entities, Stanford Encyclopedia of Philosophy, 2018

[informal-proofs] B. Larvor, How to think about informal proofs, Synthese, **187**, 715 – 730, 2011

[IUTT] S. Mochizuki, Inter-universal Teichmüller Theory I – IV, Preprints, 2016

[coping] D. Knuth, Aathematics and Computer Science: Coping with Finiteness", Science **194**, 1235 – 1242, 1976

[example] A. Naumowicz, An example of formalizing recent mathematical results in Mizar, Journal of Applied Logic **4**, pp. 396 – 413, 2006

[constr-logic&-math] T. Streicher, Introduction to Constructive Logic and Mathematics, manuscript, 2001, `http://www.mathematik.tu-darmstadt.de/~streicher/CLM/clm.pdf`

[constr-in-math] A. Troelstra and D. van Dalen, Constructivism in Mathematics, An Introduction, Volume 1, Springer-Verlag, 1988

[on-Kronecker] H. Weber, Leopold Kronecker. In: Jahresbericht der Deutschen Mathematiker-Vereinigung **2**, 19 – 25, 1893

Supplement: Inspirations and the process of writing

What inspired the authors to write this note and to articulate the perspective on mathematics presented here?

In 2012, Shinichi Mochizuki made four papers about "Inter-universal Teichmüller Theory" ([IUTT]) public which contain allegedly a - yet unconfirmed - proof of the abc conjecture, a central conjecture from number theory. In the last part of the fourth paper, Mochizuki introduces what he calls the "language of species". The first author assigned to the second author the task for his "Diplomarbeit" to explain "what is going on here", and the second author agreed to this task.

It turned out that "species" are nothing but formulae describing categories, and there are also "mutations" describing functors. For example, a species defines the underlying class of a category via a property which the objects shall satisfy; this seems to be similar to the possible introduction of "fake" classes for individual properties in a higher language which can be rewritten in ZFC. This lead to numerous – sometimes hours long discussions. The rather concrete question discussed was: Should the "species" be seen as objects themselves or does Mochizuki merely advocate a particular point of view, maybe a way to do and to write down mathematics? What is or what should be the ontological status of "species"? The authors then realized that to even discuss these question, they have to consider deeper underlying questions on formalisms in mathematics in general and on the ontological status of mathematics in general.

The second author expressed a formalistic point of view: If mathematicians claim to have a proven a result, they actually claim that they can write down such a proof in a formal system, which is by default ZFC, everything else would be imaginary. The first author countered: This does not correspond to what you are doing yourself if you are doing mathematics. In fact, you do not formally write down your "proofs in a formal system", and you are also not able to do so. So your "formalistic proofs" are also purely imaginary. Moreover, your view on mathematics does not correspond to your behavior: I can see you getting excited when you talk about "mathematical objects". One can indeed argue that the objects are purely imaginary (as one can also for the "formalistic proofs"), but when expressing their thoughts, mathematicians do indeed act as it the objects were real; this should be expressed.

The first author stressed then the point of view that questions of realism should be sidestepped by simply looking at what mathematicians do, and this is, he argued talking about non-real objects as if they were real. So the ideas of mathematical fictionalism should be considered. The second author emphasized that be that as it may, mathematicians in their quest for certainty want to be sure that their arguments follow formally from clearly outlined axioms.

This stress on the formalistic foundations of mathematics became of importance when the authors wanted to write about the proper interpretation of statements on formal sentences. Here, however, a crucial problem came to light: What does a claim "The argument could be written in ZFC." actually mean? What if such a writing cannot be done by humans? What if the writing can (or realistically could) be done by computers but then the formal proof could not be checked by humans?

A statement like "There is a formal proof of Fermat's last theorem in ZFC" does indeed "feel" more real than the very starting point of arithmetic: Counting never stops, or formulated more "platonistically": Every natural number has a successor, and certainly more real than set theory. The challenge than was to find an approriate framework which one the one hand is still mathematical (rather than physical) and on the other hand captures the idea of construction. With the idea of construction in mind, the authors then started to study constructive mathematics (which they did not know well before) and found that (with an approriate interpretation) it provided exactly the kind of framework they searched for.

It still remains the question in what sense, if any, statements on the existence of formal proofs in this framework should be seen as being "less imaginary" than the usual set-theoretic statements. Should such a qualification not at least require some kind of argument in favor of the possibility for humans to obtain such a proof in a physical sense?

In any case, statements of rewriting and on formal proofs in the form of "it could be done" are made by mathematicians, and it is a challenge to come up with a meaning of such statements. The authors are convinced that for an appropriate answer one should turn to constructive mathematics, whether or not one then regards the claims then as being "less imaginary" than set-theoretic claims.

They invite everybody to reflect on the challenge for themselves.