

PROOF-THEORETIC REDUCTION AS A PHILOSOPHER'S TOOL

1. PROOF-THEORETIC REDUCTION AND HILBERT'S PROGRAM

Hilbert's program in the philosophy of mathematics comes in two parts. One part is a technical part. To carry out this part of the program one has to prove a certain technical result. The other part of the program is a philosophical part. It is concerned with philosophical questions that are the real aim of the program. To carry out this part one, basically, has to show why the technical part answers the philosophical questions one wanted to have answered. Hilbert probably thought that he had completed the philosophical part of his program, maybe up to a few details. What was left to do was the technical part. To carry it out one, roughly, had to give a precise axiomatization of mathematics and show that it is consistent on purely finitistic grounds. This would come down to giving a relative consistency proof of mathematics in finitist mathematics, or to give a proof-theoretic reduction of mathematics on to finitist mathematics (we will look at these notions in more detail soon). It is widely believed that Gödel's theorems showed that the technical part of Hilbert's program could not be carried out. Gödel's theorems show that the consistency of arithmetic can not even be proven in arithmetic, not to speak of by finitistic means alone. So, the technical part of Hilbert's program is hopeless, and since Hilbert's program essentially relied on both the technical and the philosophical part, Hilbert's program as a whole is hopeless.

Justified as this attitude is, it is a bit unfortunate. It is unfortunate because it takes away too much attention from the philosophical part of Hilbert's program. And this is unfortunate for two reasons.

First, because it is not at all clear what the philosophical part of Hilbert's program comes down to. What of philosophical importance about mathematics would Hilbert have shown if the technical part of his program would have worked?

Secondly, because even though the technical part of Hilbert's original program is hopeless, there are modified versions of Hilbert's program that do not have the technical difficulties that Hilbert's original program has. In fact, for so-called relativized versions of Hilbert's program the technical parts are known to work.¹ However, the people working on the technical results related to relativized versions of Hilbert's program have



much less explicit or much less published views on the philosophical importance of these results than Hilbert. What is the philosophical upshot of these results? What of philosophical relevance follows from these technical results?

These two concerns are, of course, not independent. A better understanding of the philosophical part of Hilbert's original program will give rise to a better understanding of the philosophical significance of the proof-theoretic results coming from relativized versions of Hilbert's program, and the other way round. But what the philosophical upshot of such results is is not at all clear. Even though it is not clear what positive contribution these results can make to the philosophy of mathematics, it is pretty clear, I think, what contribution they cannot make. In the next section I will spell out more clearly what these proof-theoretic results are and what philosophical importance these results should not be taken to have. After that I will make a modest attempt to show what philosophical importance these results might have. This section will be modest in the sense that all I will argue for is that given certain views, to which I am sympathetic, these results would have a positive role in the philosophy of mathematics, if only a limited one.

Even though this paper is about the philosophical relevance of results closely related to and inspired by Hilbert's original program, I will not talk about Hilbert in this paper. What Hilbert's views on the matter were is not at all obvious, and I have nothing of interest to contribute to the on-going debate about the historical Hilbert. I refer you to Mancosu (1996), which contains a very useful survey of Hilbert's views and of debates that Hilbert was engaged in, as well as lots of references to works by and about Hilbert. The present paper is not historical, but merely concerned with the question: what might the philosophical importance of these technical results be?

When I ask what philosophical significance certain technical results have we have to distinguish two ways in which this might be understood. We can distinguish between foundational importance and large scale philosophical importance. The first concerns questions like what axioms are needed to prove certain theorems which theory is stronger in some sense than which other theory which parts of mathematics relate in certain ways to which other parts and the like. These questions concern the foundations of mathematics. Answering the foundational questions is not by itself answering the large-scale philosophical questions about mathematics. These latter questions are questions about the existence of mathematical objects, objectivity, the nature of mathematical truth, the nature of mathematical knowledge, and the like. Foundational questions are, of course, philosophical questions, too, but when I speak of the philosophical importance of

certain technical results I have in mind the large-scale philosophical importance that these results might have. I take it to be quite obvious that the technical results that Hilbert had in mind would have foundational importance. Whether or not these results also have philosophical importance in the other sense is what we will look at. It seems pretty clear that Hilbert had such importance in mind. After all, any program in the philosophy of mathematics worth the name should have implications for these questions.

In addition, we should distinguish negative and positive roles that such technical results might have for large-scale philosophical debates. A technical result would have a negative role if it would help to limit or refute a certain large-scale philosophical view. Gödel's theorems might be taken to play a negative role in the discussion about Hilbert's original program. A technical result would play a positive role, on the other hand, if it would play an important part in the defense of a certain large-scale philosophical view. In this sense, it would not be used to refute or limit other views, but make a constructive contribution to the formulation or defense of a certain position in the philosophy of mathematics. It seems that Hilbert had something like this in mind for the technical results he wanted to have. So, the main question that we will address in this paper is whether or not proof-theoretic reductions have any philosophical importance, and whether or not they can play a positive, constructive role in the defense of a large scale philosophical position in the philosophy of mathematics.

2. WHAT PHILOSOPHICAL IMPORTANCE PROOF-THEORETIC REDUCTIONS DO NOT HAVE

The technical part of Hilbert's program, as well as relativized versions of Hilbert's program involve the notion of a proof-theoretic reduction. Proof-theoretic reductions are results from proof-theory which are closely related to relative consistency proofs. We will have to look a bit more closely at what these results are. In particular, we will have to have a look at in what sense they can be called reductions. There are at least two very common ways in which the word 'reduction' is used. On the one hand it is used to talk about a relation between two theories: theory reduction. On the other hand it is used to talk about a relation among things, or entities: ontological reduction. As we will see, proof-theoretic reductions aren't reductions in either one of these senses of the word 'reduction'.

2.1. *Proof-Theoretic Reduction*

Simply put, a theory T_1 (formulated in language L_1) is said to be proof-theoretically reducible to a theory T_2 (formulated in L_2) if the following conditions hold:

- The shared part of the languages of L_1 and L_2 is sufficiently large (i.e., contains at least primitive recursive arithmetic).
- Any proof in T_1 of an equation of closed terms of the shared part of the languages of T_1 and T_2 can be effectively transformed into a proof of the same formula in T_2 .
- The above fact is provable in T_2 (i.e., T_2 proves that any proof in T_1 of an equation of closed terms of the shared part of the language etc. . . .).

(For a more precise definition, see Feferman (1998). The one above should suffice for our purposes.) If T_1 is proof-theoretically reducible to T_2 then T_2 will prove that if T_2 is consistent then T_1 is consistent, too. This is so because of the following. Assume we have a proof-theoretic reduction of T_1 to T_2 . If T_1 is not consistent then it proves '0=1'. This is an equation of closed terms in the shared part of the languages of T_1 and T_2 . Thus there is an effective way of transforming this proof of '0=1' into a proof of '0=1' in T_2 . In addition, T_2 proves that this is so. Thus T_2 proves that if there is no such proof in T_2 then there is no such proof in T_1 . In other words, T_2 proves that if T_2 is consistent then T_1 is consistent, too. Thus if a theory T_1 is proof-theoretically reducible to another theory T_2 then we have a relative consistency proof of T_1 in T_2 . In particular, if T_2 is a formalization of finitist mathematics, and T_1 is a formalization of all of mathematics then a proof-theoretic reduction of T_1 to T_2 would complete the technical part of Hilbert's original program.

There are a number of successful proof-theoretic reductions. One example is the reduction of Peano Arithmetic augmented by certain restricted forms of second order quantification (for example, ranging only over sets of natural numbers definable in first order arithmetic) to Peano Arithmetic.² Another, more tricky, example is Feferman's system **W** which is expressive enough to allow the formulation of a substantial part of modern analysis, in particular the part that is used in formulating scientific theories, but still is proof-theoretically reducible to Peano Arithmetic.³

To see what philosophical importance such proof-theoretic reductions might be taken to have we should first note what importance they should not be taken to have. Most importantly, even though proof-theoretic reductions are called reductions, they are not reductions in two commonly used senses of the word. They are neither theory reductions, nor ontological reductions, at least not in the sense in which these are commonly understood.

Let us look at each one of them, and how they relate to proof-theoretic reductions.

2.2. Theory Reduction

Theory reduction is a reduction of one theory to another in the sense that one theory can be understood as being contained in, or being a special case of, another theory. A classic example of this would be the reduction of a specialized physical theory to a very general physical theory. The former is merely one that looks at a limited case of phenomena, and maybe because of that was discovered first. Once the general theory comes along the special one might be reduced to the general. The former is contained in the latter. Besides examples of different physical theories one might think that chemistry can be reduced to some branch of physics in this sense.

What exactly would constitute a successful reduction of, say, chemistry to physics is not at all clear. What it would take for one scientific theory to reduce to another is a controversial and complicated issue in the philosophy of science that cannot be appropriately addressed here. The notion should be intuitively clear enough for our purposes here, though. We will have to focus here on how certain technical results relate to theory reduction, in the intuitive understanding.

There is a technical notion which *prima facie* does seem to have some relevance to the question of whether or not a theory can be reduced to another. This notion is that of a relative interpretation. If one theory is relatively interpretable in another then, in some sense, the second can be seen as containing the first as a special case. More precisely, a theory T_1 (formulated in language L_1) is said to be relatively interpretable in a theory T_2 (formulated in language L_2) iff the following holds:

- we can associate with each predicate symbol of L_1 a predicate of L_2 , with each closed term of L_1 a closed term of L_2 , and with each function symbol of L_1 a function of L_2 , such that the “translation” of the formulas of L_1 into L_2 that is determined by this association is such that:
- The translation of the axioms of T_1 into L_2 (with all quantifiers restricted to some predicate of L_2) each follows from the axioms of T_2 .

This captures the idea of one formal theory being only a special case of another in a precise way.

There is a bit of a tricky relation between relative interpretations and proof-theoretic reductions. It is a well known result from proof theory that any axiomatizable theory T can be relatively interpreted in Peano Arith-

metic + the statement that T is consistent.⁴ If we have a proof-theoretic reduction of T_1 to T_2 then we saw above that T_2 proves that if T_2 is consistent so is T_1 . Thus T_2 together with the statement that T_2 is consistent proves that T_1 is consistent, and assuming that T_2 contains Peano Arithmetic we will have that T_1 is relatively interpretable in T_2 + the statement that T_2 is consistent. Whether or not T_1 is relatively interpretable in T_2 alone is another matter, and gives rise to some subtle issues that we will not get into here.⁵

But even if we would have a relative interpretation of the kind that we can hope to get through these general results, would this constitute a theory reduction in a sense of the word in which it would have philosophical implications?

We can assume for now that for one theory to be reducible to another the former has to be relatively interpretable in the latter. But can any relative interpretation be considered a reduction of one theory on another? I think it is clear that it can not. Relative interpretation is at most a necessary condition for theory reduction, but it is not a sufficient one. Simply because one theory can be associated with another in the way that relative interpretation requires does tell us something about the complexities involved in these theories, but it does not *per se* tell us something about reduction in a relevant sense. Whether or not the interpreted theory really is a special case of the interpreting theory is a further question. Even if the above association can be specified, it is a further question whether or not it is the right association. Formal interpretability alone will not do. What else is required is, as I mentioned above, a controversial issue in the philosophy of science. All that we have to agree on for now is that proof-theoretic reduction, or relative interpretation, alone is not sufficient for theory reduction, in a philosophically relevant sense of the word.

To make this clearer, and to see that relative interpretation alone does not provide a theory reduction (in the present sense of the word), consider the following: take a first order formulation of some physical theory. By the result mentioned above it is relatively interpretable in some arithmetic theory. But, of course, the former theory can not be reduced, in the intuitive sense of the word, to the latter. The former is about physical objects, the latter about numbers. The former is not just a special case of the latter. So, reduction requires more than just an association of the relevant formal languages in the right way. There is a further question which one, if any, is the right association. So, even if we have a relative interpretation, it would not mean that we have a reduction of one theory to the other in a philosophically relevant sense of the word.

To be sure, there is a somewhat subtle relation between one theory reducing to another, and claims about whether or not these two theories are about different objects. One might think that a successful reduction shows that these two theories really are about the same objects, and that I am thus begging the question. This is partly right. There are controversial cases where it is not clear whether or not two theories are about the same objects. Theories about the mind and about the brain, for example. But there are also cases where it is clear that the theories are not about the same objects. Physics and number theory are of this kind. It is simply beyond question that physics is not about numbers. This is a case where it is not, or surely should not be, controversial whether or not these two theories are about the same objects. But physics is relatively interpretable in a certain number theory. Thus this is a case that shows that relative interpretation is not theory reduction, since physics can not be reduced to number theory. And one of the reasons why it cannot be is that physics is not about numbers, but number theory is.

Relative interpretations only show something about the interpretability of certain formal languages, and formal theories. But for theory reduction more is required. The informal, interpreted theories that are modeled by these languages have to stand in a certain relation to each other. What this further relation is is not clear, but it is clear that some such further requirements are necessary for a relative interpretation to be called a successful theory reduction. And from the technical results about interpretability alone no conclusions about reduction in a philosophically relevant sense can be drawn.

2.3. *Ontological Reduction*

Also, proof-theoretic reductions are not ontological reductions. By ontological reduction I simply mean a demonstration that certain entities really are identical to certain entities that are apparently of a different kind. For example, if one could show that mental events are just physical events then one would have an ontological reduction of mental events to physical events. To give an ontological reduction of things of kind *A* to things of kind *B* is to show that things of kind *A* really are things of kind *B*. Therefore we only have one kind of things, not two.

From the fact that one theory is proof-theoretically reducible to another it does not follow that what the one theory talks about is just the same as what the other theory talks about. As we saw above, the reduced theory might involve a form of second order quantification, and simply from the fact that this form of quantification together with the axioms that govern it and the rest of the theory do not help in proving more about the closed

equations in the shared language does not mean that the objects in the domain of these quantifiers really are identical to certain ones of the objects in the domain of quantification of the reducing theory. It only means that with the axioms that we have in this theory about such quantification we cannot prove anything more about a certain class of formulas than with another theory that does not involve such quantification.

2.4. *Conservative Extension*

There is one further consideration that one might think gives rise to the philosophical importance of proof-theoretic reductions. It comes from the fact that proof-theoretic reduction guarantees conservative extension with respect to the relevant class of formulas. If T_1 is proof-theoretically reducible to T_2 then T_1 does not prove anything about equations of closed terms the shared part of their languages over and above what T_2 proves (and on top of this, T_2 knows this). More generally, if T_1 is proof theoretically reducible to T_2 with respect to a certain class of formulas, then T_1 forms a conservative extension over T_2 with respect to that class of formulas. If we only consider proof-theoretic reductions with respect to equations of closed terms (which is enough to get the connection to relative consistency proofs), then we are only guaranteed conservative extension with respect to closed equations, but in practice one will often get more. Since we are concerned here with the question whether proof-theoretic reductions can get their philosophical importance through their connection with conservative extensions we will put aside subtleties about for which class of formulas we get a conservative extension, and assume for the moment that we are dealing with a case where we get a full conservative extension.

One might now think that having a conservative extension result justifies, or at least supports, the following attitude: Suppose a certain part of mathematics is axiomatized with a theory T_1 , and another part with a theory T_2 , and that T_1 is proof-theoretically reducible to T_2 . Because of this T_1 is a conservative extension of T_2 . Thus all the parts of T_1 that are not in T_2 do not help to prove more about the shared language of the two theories, we can take an instrumentalist attitude towards these parts. That is to say, we can interpret them as being merely useful tools we take recourse to in achieving results about the shared language of T_1 and T_2 . So, we might think that a proof-theoretic reduction justifies having an instrumentalist attitude towards the parts of T_1 that go beyond T_2 . And the purpose for which we take talk about these parts to be instruments is to get (maybe easier or more convenient) proofs about the shared part of T_1 and T_2 .

This is a suggestion that seems to be onto something, but it cannot be right as it is. It could only be correct if the reduced theory really only were

a tool to achieve further results about the shared part of the reduced and the reducing theory. But consider, for example, Hilbert's original proposal. Hilbert wanted to proof-theoretically reduce all of math to finitistic math, i.e., have a relative consistency proof of all of math on the basis of finitistic math. But if the present interpretation of the philosophical importance were right then all of math would just be a tool to achieve further results about finitistic math. But that is not right. To be sure, if there were such a result then one could use all of math as such a tool. But whether or not it in fact is such a tool is a different question from the question whether or not it can be used as such a tool. The question whether or not it is such a tool is a question about what the function of all of math is, how it is in fact used, what in fact its purpose is. And it seems that there are lots of parts of mathematics that have their goals in quite different things than achieving results about finitistic math. Take, for example, the theory of real numbers. To be sure, this theory might give rise to interesting results in finitistic math, but that certainly is not its goal. It would be a mistake to think that real number theory got developed and has as its goal to have more results about finitistic math. That there is such a proof-theoretic reduction is necessary for this kind of instrumentalism to get off the ground, but it is not sufficient to establish it.

Once we distinguish 'can be used as a tool to do X ', 'should be used as a tool to do X ' and 'is a tool to do X ', we see that a technical result alone could only establish the first, but something of philosophical significance would only follow from the second or third. This form of instrumentalism thus is too simplistic. Even if a proof-theoretic reduction of one part of math to another is possible, it is in no way clear that the purpose of that part of math is only to achieve results about the reducing part of math. In fact, it seems clearly false. Simply that we can have this attitude does not mean we should have it, and for many cases it seems clear that we should not have it.

I would like to point out that in this section I was looking at the case where we have a proof-theoretic reduction of one theory that is taken to be the axiomatization of one part of mathematics to another theory which is taken to be the axiomatization of another part of mathematics. What I have said here might not directly apply to other cases, like the case where we have a proof-theoretic reduction of a theory taken to be the axiomatization of a certain part of mathematics, to a proper part of that same theory. We will look at this soon.

The possible candidates for the philosophical significance of proof-theoretic reductions that we have seen so far are not satisfactory. But other ones might do better. Let us now look at some more promising candidates.

3. WHAT PHILOSOPHICAL IMPORTANCE PROOF-THEORETIC REDUCTIONS MIGHT HAVE

3.1. *Indispensability Arguments*

One way in which proof-theoretic reductions might have philosophical importance, apparently quite independently from any large scale philosophical program like Hilbert's, has been discussed in the literature, in particular by Feferman. According to this consideration proof-theoretic reductions have philosophical importance because they show how much of mathematics is affected by the indispensability arguments. If all of mathematics that plays a role in science is proof-theoretically reducible to a certain small part of mathematics, then this is relevant in determining which parts of mathematics are indispensable for science, and which ones are not.

This line of reasoning assumes something like an instrumentalist attitude of the following kind: if there are certain parts of a theory T such that they do not give us any greater deductive power than the theory without these parts then an instrumentalist attitude towards these parts of the theory is justified (or might be justified in certain cases). In particular, if T_2 is contained in T_1 , but nonetheless, T_1 is proof-theoretically reducible to T_2 then we are justified in taking an instrumentalist attitude towards the parts of T_1 that are not in T_2 . This form of instrumentalism is somewhat different from the one we looked at above. Now we are looking at the relation of two parts of one and the same theory. Above we looked at two different theories that had overlapping parts. This form of instrumentalism does not seem to suffer from the same deficits as the above one. But for it, too, we are faced with the question whether we simply can take such a stance, or whether we should take it, i.e., whether this stance is the correct one. Again, the technical results are necessary for the instrumentalism to get off the ground, but they alone are not sufficient to establish it. In fact, all that strictly seems to be necessary is that T_1 forms a conservative extension over T_2 . What additional work the stronger proof-theoretic reduction is doing needs to be established separately.

I will not get into this discussion here, neither into the one about the indispensability arguments, nor about this form of instrumentalism.⁶ I would like to point out, though, that whatever the outcome, the above line at most would show that a certain argument, the indispensability argument, only has a limited application. The importance of proof-theoretic reduction in this sense is only to limit the results we would get from the indispensability argument, assuming we get one at all. This is, of course, not negligible, but certainly much less ambitious than Hilbert's original goals might be

understood. In and by itself the above line of reasoning does not provide a positive account of reasons for any specific position in the philosophy of mathematics, but only a negative one.

3.2. *Foundational Reductions*

One other way of looking at proof-theoretic reductions is also found in the writings of Feferman (see Feferman 1998) and concerns their having a close relation to so-called foundational reductions. Foundational reductions arise from the relation that one foundational framework has to another. For example, classical Peano Arithmetic might be taken to be based on a certain foundational framework. Certain basic philosophical views will be necessary to justify reasoning about numbers with classical logic. The intuitionistic Heyting Arithmetic, on the other hand, can be seen to be based on a different foundational framework. Certain proof-theoretic results, like Gödel's translation of Peano Arithmetic into Heyting Arithmetic, can then be seen as providing a foundational reduction of the foundational framework of classical arithmetic to the foundational framework of intuitionistic arithmetic.

Looking at it this way, one can argue that the foundational framework of classical arithmetic was not really necessary to justify classical arithmetic. The framework of intuitionistic arithmetic is sufficient to do this. In this sense, the foundational framework of classical arithmetic can be reduced to the foundational framework of intuitionistic arithmetic. But looking at it this way also makes clear that these results will only have a negative role. They would only show that what some people take to be required to justify a certain theory, or certain forms of reasoning, is not really required. One doesn't have to adopt a certain classical foundational framework to justify classical arithmetic. But from that alone it is not clear what positive philosophical role such results can play. To be sure, there seems to be something there, but it is not clear how one can get more than a negative result from these considerations.

3.3. *Consistency (and Existence)*

To see what the philosophical importance of proof-theoretic reductions might be, over and above limiting the indispensability arguments or the adoption of certain foundational frameworks, it will help to see what the importance of the closely related relative consistency proofs might be. It was the main technical goal of Hilbert's program to have a relative consistency proof for all of mathematics in finitistic mathematics. At least Hilbert thought that such a relative consistency proof would have great philosophical importance. But what precisely the philosophical importance

of relative consistency proofs is not clear, whether or not they are proofs in finitistic mathematics or in other parts of mathematics. Why should we, as philosophers, care about relative consistency proofs? This can be divided up into two questions:

- why should we care about having a consistency proof?
- why should we care about having a consistency proof within a certain theory?

Let us address these two in this order.

3.3.1. *Consistency proofs*

It might seem that the answer why we would want to have a consistency proof is obvious: we want to be sure that a certain theory is consistent, so to be sure that no paradoxes or contradictions can be derived from certain axioms. This, of course, would be nice to know. But it is at first not so clear why knowing this, or being able to prove this, should be taken to be of much philosophical significance. After all, a consistency proof shows that *the axioms of a certain axiomatization* of a branch of mathematics are consistent. It does not thereby necessarily show anything about the branch of mathematics, only about the axiomatization of the reasoning within that branch of mathematics. If the axioms were inconsistent, would this reflect on the branch of mathematics, or on the axiomatization of the reasoning within that branch? Take the real numbers as an example. Suppose someone proposes an axiomatization of (a certain part of) our reasoning about real numbers, and it turns out that this axiomatization is inconsistent. Would this lead us to conclude that there are, after all, no real numbers, and that real number theory is based on a mistake? Well, it depends on the inconsistency found and it is conceivable that we would react this way. But it is much more likely that we will have one of the following reactions:

- we will conclude that the axiomatization proposed incorrectly describes our (informal) reasoning about real numbers, or
- we will conclude that we had false beliefs about real numbers.

Even if we grant that an axiomatization correctly captures our reasoning about real numbers, our belief that there are real numbers will not directly depend on our reasoning about them being consistent. We will be more inclined to give up that we reasoned correctly about them than to say that a whole part of mathematics is bankrupt. And that is exactly what happened when we found out that our reasoning about sets was inconsistent.

In the light of this, what philosophical importance can consistency proofs have? It might seem that all they can do for us is to assure us

that our reasoning in a certain branch of mathematics will not lead to contradictions. That is, of course, comforting to know, but is that really of much philosophical importance?

In addition, it seems that there is no real cognitive need for a whole variety of consistency proofs. Take the example of ZF set theory. It is a very strong theory, and it would have been the grand prize for Hilbert to accommodate it in his program. But not many people doubt that this theory is consistent. Even philosophers of mathematics who oppose a certain set theoretic approach to the foundations and philosophy of mathematics usually do not believe that ZF set theory is inconsistent. So, why would it be so great to show that it is consistent?

Take the philosophy of mind as a further example. We have many ways to talk and reason about the mind, and there are many philosophical problems about the mind. Suppose now that someone came along to show that all of our talk and reasoning about the mind can be captured in a nice formal system and that furthermore this system can be proven to be consistent. What would follow from this for the philosophy of mind? Well, basically nothing, it seems.

Consistency proofs would have a greater importance if there was a closer connection between the axioms capturing a certain part of mathematical reasoning and what the mathematical discipline is about. In the above reasoning, it was implicit that the objects of a mathematical theory were independent of the reasoning of mathematicians about them. And that is not unreasonable, since it also holds of our reasoning about, say, tables or dogs. They are independent of our reasoning about them, and if we make mistakes in reasoning it is our fault, not theirs. But mathematics might be different. In mathematics there might be a closer connection between our reasoning in a certain branch of mathematics and the objects that this branch is about. And if so, then a consistency proof might do more than giving us comfort in a certain axiomatization.

There are (at least) two approaches that make a much closer connection between the axioms of a theory and what the axioms are about than the above one. One of them has been called *full-blooded platonism*.⁷ The basic idea of this view is that there are lots and lots of abstract objects, in fact, for every consistent theory, there are objects that satisfy that theory. Furthermore, for every consistent theory there are objects that can reasonably be regarded as the objects of that theory, and of no other theory. Simply put, there are so many abstract objects that every consistent theory has its very own objects that this theory, and no other theory, is about. In particular, different set theories (for example, ZF + the continuum hypothesis, and ZF + the negation of the continuum hypothesis⁸) are really not two competing

theories about the same objects, namely the one and only sets, but rather two non-competing theories about different objects. Both of them will be true.⁹

Such a view will have as a consequence that there is a much closer connection between consistent axioms and the objects the axioms are about. Any consistent set of axioms is about its own domain of objects. Thus, given full-blooded platonism, as long as we have consistency we have the existence of the objects the theory is about. Simply put: consistency implies existence.

A related, but different, view is one that might be called *stipulative instrumentalism*. This view is, roughly, that certain entities are stipulated to, or pretended to, exist in order for talk about them to fulfill a certain purpose. For example, we might have a need for talking about things that have certain properties. We might want to use talk about them to model certain features of the world. Then we can say precisely what properties these things are supposed to have, how they are supposed to relate to each other, and the like, and finally stipulate, or pretend, or simply assume, that there are things that satisfy the description we just gave. When we then talk about these things, such talk will not be literally true, but will only be true given the stipulation, or within the pretense that there are such things as we stipulate, or assume, or pretend, there are.

According to this view, too, consistency of the axioms will have an important role. The axioms here have to be taken to be a precise form of the stipulations that we made about the things that we want to talk about. And if we have consistency then what we stipulate will be well-defined. After all, if we stipulate that these things have properties that lead to contradictions then we have not given a coherent description of what we wanted to stipulate. And without that it is hard to make sense that we stipulated anything. As soon as we have consistency everything goes smoothly, though. Again, simply put: consistency implies (stipulated) existence.

So, if one holds certain philosophical views, like full-blooded platonism or a form of stipulative instrumentalism about all or a certain part of mathematics then consistency proofs will be of a more central importance than just an assurance that your reasoning about the objects won't lead to contradictions. It will guarantee that there are such objects, in the relevant sense. So, there might be some importance to consistency proofs to philosophers who have such views. But why worry about relative consistency proofs?

3.3.2. *Relative consistency proofs*

For a relative consistency proof to be of special philosophical importance we have to believe something like all of the following:

1. Consistency proofs are of philosophical importance (see above).
2. The theory in which the relative consistency proof is given somehow is from a philosophical point of view distinguished in a way that the theory whose consistency is proven (at least *prima facie*) is not. Or in terms of proof-theoretic reductions: the reducing theory has some philosophically special feature that the reduced theory (at least *prima facie*) does not have.
3. A successful proof-theoretic reduction (or relative consistency proof) 'transfers' some of the philosophically special features that the reducing theory has to the reduced theory.

Now, if you are a full-blooded platonist then you will usually believe that full-blooded platonism is true of every mathematical theory. You do not have to believe this, of course, but usually, if you are a full-blooded platonist at all, you will believe that it applies to every mathematical theory equally well. Thus, having a proof-theoretic reduction of one such theory to another will not give any special philosophical significance to the reduced theory, since the large-scale philosophical features of the reduced and the reducing theory are the same. To be sure, the reduced theory might be epistemologically distinguished. It might be easier for use to persuade ourselves that the reducing theory is consistent, and therefore a proof-theoretic reduction might give us strengthened confidence that the reduced theory is consistent as well. This is not to be ignored, but it seems to be mainly related to our confidence in theory, and it remains unclear how it relates to the large-scale philosophical issues we are concerned with now. A relative consistency proof for a full-blooded platonist might have the same importance as it has for anyone else. It might persuade us that the axioms of a certain theory indeed are consistent. Nice as it is, nothing of special philosophical significance seems to follow from that.

If, however, you believe in a form of stipulative instrumentalism then things might be different. Such instrumentalists will believe that entities satisfying certain descriptions are stipulated to exist for a certain purpose. And the purpose does not have to be to do mathematics or something that general. The purpose might be much more complicated. In particular, and contrary to the form of instrumentalism that we encountered above when we talked about conservative extensions, the purpose for which these stipulations are made does not have to be within mathematics proper. It can be, and presumably according to such an instrumentalist often will be, the case that the purpose for these stipulations is to deal with descriptions of

the non-mathematical world. In particular, a stipulative instrumentalist will not believe that every part of mathematics is the same. For which parts of mathematics stipulative instrumentalism is true and for which ones it isn't will depend on what people were doing when they introduced that part of mathematics and possibly other empirical issues. So, even if one is in principle willing to believe in instrumentalism for some parts of mathematics, it will still be a further, and largely empirical, question for which parts of mathematics instrumentalism is true.

I believe that instrumentalism is not true for at least one important part of mathematics: arithmetic. Contrary to what an instrumentalist would believe, I think that arithmetic truth is not truth given certain stipulations, or within a certain pretense, it is truth *simpliciter*. A certain form of platonism is true about arithmetic: arithmetic truth is objective truth, independent of any axiomatization, any stipulations or the like. In addition, every arithmetic statement is either objectively true or objectively false, regardless of whether it can ever be known. One could say that arithmetic is objectively complete. This is a form of platonism, but only platonism about objectivity. It is not *per se* a form of platonism about mathematical objects. In fact, I think that arithmetic truth is independent of the existence of any objects, including mathematical objects. The position I endorse about arithmetic is a form of platonism about objectivity while at the same time denying platonism about objects. It can legitimately be called a form of logicism, even though nothing hangs on it being called that. I only state the position here for the present discussion about relative consistency proofs, but I will not try to defend it here. I have defended it in Hofweber (1998, 1999), and I will also spell this out in detail on another occasion. What matters for us now is that if this were correct and if it were also correct that stipulative instrumentalism is true about a different part of mathematics then we would find a positive, though limited, role for relative consistency proofs in large-scale issues in the philosophy of mathematics. Under these two assumptions, that this view of arithmetic is correct, and that stipulative instrumentalism is correct about some other part of mathematics, we can see what large-scale philosophical relevance proof-theoretic reductions can have.

Suppose that arithmetic truth is objective truth, and suppose that truth in a different part of mathematics is not truth *simpliciter*, but truth given certain stipulations. If this is so then the latter might also give rise to a form of objectivity, namely it might be objectively so that a certain claim is true given the stipulations. But this is in certain important respects different. For one, truth given stipulations is different that truth *simpliciter*. In addition, the stipulations that are supposed to determine the properties of the things

that are stipulated to exist might only characterize them incompletely, and thus certain well-defined questions about the things that were stipulated to exist might not have an objective answer even in this sense of the word. The stipulations might not determine an answer, and thus there is not an objectively correct answer to this question, even given the stipulations.

Suppose that stipulative instrumentalism is true of a certain branch of mathematics B , and that T is a collection of axioms that correctly capture the stipulations that were made to introduce B and what it is about. To see whether a relative consistency proof of T in some theory can make a positive contribution to large-scale issues in the philosophy of mathematics, we will have to see what importance the consistency of T can have for this, and what importance a proof of the consistency in some theory can have.

The first is not hard to see, and was already discussed above. If T captures the principles that were used to stipulate what B is about then the consistency of T not only guarantees that our reasoning in B will not lead to contradictions, it most of all guarantees that B can be said to be about anything at all. If stipulative instrumentalism is true of a branch of mathematics then the consistency of the principles used in these stipulations is of great importance for this domain, and lies at the heart of this branch being well defined.

But why is it important to prove the consistency of a theory in some other theory? Consistency statements are complicated combinatorial statements, and as such the same questions that apply to other mathematical statements apply to them. For one, there is an issue whether or not questions about consistency have an objective answer. For questions about consistency, just as for questions about real numbers, we have to ask ourselves, do they have an answer only given certain stipulations or do they have an answer objectively and independent of any stipulations? What a stipulative instrumentalist about a certain branch of mathematics would want to have is, of course, that it is objectively true that the principles governing that branch are consistent. If it only were true given certain other stipulations we would push the issue back to the consistency of these stipulations. Now, if we grant the picture of arithmetic that I outlined above then it will be the case that what is provable in arithmetic will be objectively true, not just true given certain stipulations. Thus, given the view of arithmetic outlined above, Peano Arithmetic will be objectively true, and so will be Peano Arithmetic + the statement asserting the consistency of Peano Arithmetic. Thus, if there is a relative consistency proof of T in Peano Arithmetic then it will be objectively true that T is consistent. Thus, a relative consistency proof in arithmetic guarantees that it is objectively true that the principles used in the stipulations are consistent.

And thus it guarantees that it is objectively true that what was stipulated is well-defined.

The question whether or not a certain instrumentalist stance towards this part of mathematics is the correct one will be independent of the question whether or not a proof-theoretic reduction of this theory to arithmetic is possible. What the correct stance is towards that theory will depend on what people do or did when they formulated that theory. It will be largely an empirical question. But if this empirical question would turn out in favor of instrumentalism, and given the form of logicism about arithmetic mentioned, then proof-theoretic reductions can make a positive contribution to the large-scale philosophical view of mathematics.

4. CONCLUSION

We have looked at what positive contribution successful proof-theoretic reductions can make to the large-scale philosophy of mathematics. We saw three promising candidates for this. One is to limit the application of the indispensability arguments when combined with a certain form of instrumentalism about the deductively inessential parts of a certain theory that allows for the expression of scientifically applicable mathematics. Another was connected to the notion of a foundational reduction. The third was to combine a certain logicist view about arithmetic with a certain other form of instrumentalism about a different part of mathematics. I found the third to be most promising, even though it would only work if certain other claims were true. One is that arithmetic is a special or distinguished part of mathematics, the other that the empirical facts about a certain part of mathematics point toward a form of stipulative instrumentalism. Both of these claims are, of course, quite controversial.

The position I outlined at the end resembles Hilbert's in several respects, even though there are many differences. The present view endorses the claim that certain parts of mathematics are philosophically special, it endorses a form of instrumentalism, it gives meaning to the claim that consistency implies existence, rightly understood, and it uses relative consistency proofs to do special philosophical work. The role of these technical results, however, is not as central as it would have been for Hilbert, but they nonetheless make a positive philosophical contribution to the large-scale issues in the philosophy of mathematics. The real philosophical work, however, will be somewhere else.

ACKNOWLEDGEMENTS

Thanks to Sol Feferman, Grisha Mints and Richard Zach for discussions and comments, and to an anonymous referee for *Erkenntnis* for helpful comments. Earlier versions of this paper were presented at the Workshop on Notions of Reduction in Mathematics and Philosophy, Munich 1997, and at the Stanford–Berkeley Conference, Stanford 1999.

NOTES

- ¹ See Feferman (1988) for a survey of a number of results in this area.
- ² Again, see Feferman (1988) for many more examples, and the details of this example.
- ³ See Feferman (1993) for more on the details of **W** and further references.
- ⁴ See Feferman (1960).
- ⁵ See Smoryński (1985) for more on this.
- ⁶ For more on this see Feferman (1993, 1998).
- ⁷ See Linsky and Zalta (1995) and Balaguer (1995) for an elaboration of such a view.
- ⁸ In fact, standard full-blooded platonists will also believe that two theories that have different axioms but are consistent with each other will be about different objects, for example ZF + the axiom of choice, and ZF without this further axiom.
- ⁹ Again, see Linsky and Zalta (1995) and Balaguer (1995).

REFERENCES

- Balaguer, M.: 1995, 'A Platonist Epistemology', *Synthese* **103**, 303–325.
- Feferman, S.: 1960, 'Arithmetization of Metamathematics in a General Setting', *Fundamenta Mathematica* **XLIX**, 35–92.
- Feferman, S.: 1988, 'Hilbert's Program Relativized', *Journal of Symbolic Logic* **53**, 364–384.
- Feferman, S.: 1993, 'Why a Little Bit Goes a Long Way: Logical Foundations of Scientifically Applicable Mathematics', *PSA 1992*, **ii**, 442–455.
- Feferman, S.: 1998, 'What Rests on What? The Proof-Theoretic Analysis of Mathematics', in *In the Light of Logic*, Oxford University Press.
- Hofweber, T.: 1999, *Ontology and Objectivity*, PhD Thesis, Stanford University.
- Hofweber, T.: 1998, 'Number Determiners, Numbers, and Arithmetic', unpublished manuscript.
- Linsky, B. and E. N. Zalta: 1995, 'Naturalized Platonism and Platonized Naturalism', *Journal of Philosophy* **92**, 525–555.
- Mancosu, P.: 1996, 'Hilbert and Bernays on Metamathematics', *From Brouwer to Hilbert*, Oxford University Press.

Smoryński, C.: 1985, 'Nonstandard Models and Related Developments', *Harvey Friedman's Research in the Foundations of Mathematics*, North-Holland, Amsterdam, pp. 179–229.

University of Michigan
Department of Philosophy
2215 Angell Hall
435 S. State Street
Ann Arbor, MI 48109-1003
U.S.A.