FIELD

Field, Harty
Realism, Mathematics and Modality. Oxford: Basic Blackwell, 1989

1

Introduction: fictionalism, epistemology and modality

A number of themes run through the papers collected in this volume, but certainly the most dominant is the idea that we can develop a satisfactory fictionalist account of mathematics. I characterize this fictionalist position in section 1 of this introduction, and discuss some of the considerations which motivate it, and some which make it difficult to achieve, in sections 2–4. Section 7 addresses what has been perhaps the most common worry about the form of fictionalism I have defended. Sections 5 and 6 discuss a second theme prominent in several of the essays in the volume: that modality has some role, but only a very limited role, to play in ontological discussions.

PART ONE

1 Fictionalism

A mathematical realist, or platonist, (as I will use these terms) is a person who (a) believes in the existence of mathematical entities (numbers, functions, sets and so forth), and (b) believes them to be mind-independent and language-independent. So there are two different (at least, verbally different) kinds of anti-platonist position. One involves a disbelief in mathematical entities; the other takes the idealist position that mathematical entities exist but only as some sort of 'mental construction' or 'construction out of our linguistic practices'.[1] It is not

[1] An example of the linguistic variant of idealism: 'The existence of numbers is just constituted by the fact that there is a legitimate practice involving discourse with a certain structure, and that certain products of this discourse meet the standards of correctness that it sets' (Stalnaker 1988). Stalnaker describes his position as platonist, but this does not seem like Platonism as it is normally understood.

immediately obvious that the idealist version of anti-platonism ought to be regarded as importantly different from the simple denial of mathematical entities – just as it is not immediately obvious that Berkeleyan idealism should be regarded as importantly different from the denial of the existence of tables and chairs and the like. I will not pursue this issue here. I will make a few very brief remarks about the idealist form of anti-platonism in section 4 of this introduction, but for the most part I intend to ignore it.

If we ignore mathematical idealism, then we can say that anyone who adopts an attitude of literal belief toward mathematical theories taken at face value is a mathematical realist. For a mathematical theory, taken at face value, is a theory that is primarily about some postulated realm of mathematical entities: numbers, or functions, or sets or whatever (or some combination, like numbers and sets together).[2] You can't consistently believe the theory without believing in the entities it postulates.

This suggests that (if we continue to ignore the possibility of mathematical idealism), an anti-platonist should embrace fictionalism about mathematics – or at least, fictionalism about mathematics-taken-at-face-value. A fictionalist about mathematics-taken-at-face-value is someone who does not literally believe mathematical sentences, at least when they are taken at face value. (Or, if you prefer to 'semantically ascend', a fictionalist is someone who does not regard such sentences, taken at face value, as literally true.) The fictionalist *may* believe that there is some non-face-value construal of mathematical sentences under which they come out true; he or she may even believe that some such construal gives 'the real meaning of' the mathematical sentence, despite its departure from what the mathematical sentence appears to mean on the surface. My own view, though, is that the second of these additional claims is an uninteresting verbal one insofar as it goes beyond the first; and that the first, though of some interest, unnecessarily constricts the fictionalist. (I'll have a few brief remarks on the latter point at the end of this section.)

A fictionalist needn't (and shouldn't) deny that there is *some* sense in which '2+2=4' is true; but granting that it is true in some sense does

[2] The theory may mention non-mathematical entities too, for instance if it says that for every entity, even a non-mathematical one, there is a set with that entity as its sole member. But I would not count it as mathematical unless it postulated or purported to refer to some mathematical entities. This is a necessary but not a sufficient condition for being a mathematical theory: a theory of mathematical physics that uses mathematics to describe the physical world is not an example of what I am calling a mathematical theory, for it is not the kind of thing that any mathematician would claim to believe on 'mathematical evidence' alone.

not commit one to finding any interesting translation procedure that takes acceptable mathematical claims into true claims that don't postulate mathematical entities. Rather, the fictionalist can say that the sense in which '2+2=4' is true is pretty much the same as the sense in which 'Oliver Twist lived in London' is true: the latter is true only in the sense that it is true *according to a certain well-known story*, and the former is true only in that it is true *according to standard mathematics*. Similarly, the fictionalist *believes* that 2+2=4 only in the sense that he or she believes that standard mathematics *says that* (or, *has as a consequence that*) 2+2=4;[3] just as most of us believe that Oliver Twist lived in London only in the sense that we believe that the novel says that or has as a consequence that Oliver Twist lived in London. If one believes only this, it seems rather natural to say that one does not literally believe that Oliver Twist lived in London (after all, one doesn't believe that if one had gone to London in the nineteenth century one would have found Oliver there); similarly, the fictionalist who regards the comparison as reasonably apt will find it natural to say that he doesn't literally believe that 2+2=4. (He or she can still advise a young school-child to say that it's true on a true–false test, of course; similarly in the Oliver Twist case.)

I am strongly inclined to think that the fictionalist view of mathematics is correct, though I have to acknowledge that there are *prima facie* obstacles that the fictionalist needs to find a way to overcome. But for now let us defer arguments against fictionalism, and focus only on what the fictionalist view *is*.

One natural question about fictionalism is this: surely the fictionalist must grant that standard mathematics is in many ways a *good* 'story': but how could he or she give any content to this except by saying that its goodness consisted in its being true? Of course, there is a dismissive reply: that the Oliver Twist story is a rather good one too, but that *its* goodness doesn't consist in its truth. But this is hardly satisfying by itself: obviously, the way in which the 'story' told by standard

[3] For impure mathematical claims it's slightly more complicated: the fictionalist believes that 2 is the number of planets closer than the Earth to the Sun in the sense that he or she believes that this follows from standard mathematics *together with purely non-mathematical facts*: among the purely non-mathematical facts is the fact that $\exists x \exists y \{x$ and y are planets closer than the Earth to the Sun; $x \neq y$; and for any planet z closer than the Earth to the Sun, either $z=x$ or $z=y\}$. I take 'standard mathematics' to include a 'bridge law' that connects up such 'statements of non-mathematical facts' to their more mathematized counterparts. (I take 'the non-mathematical facts' here to involve no claim about mathematical entities, *not even that such entities don't exist*: otherwise, the possibility of an inconsistency between standard mathematics and the non-mathematical facts would threaten the above account.)

mathematics is good is very different from the way in which the Oliver Twist story is good. There are a number of differences that might be mentioned here, but perhaps the most important is that mathematics is good *as an instrument* that can be applied in domains outside mathematics, and nothing like this is true of the Oliver Twist story. (When one wants to stress this difference, one uses the word 'instrumentalist' instead of 'fictionalist' in the case of mathematics.)[4] Because of this difference, there is certainly room for initial suspicion that for standard mathematics the only reasonable account of goodness involves truth – or, perhaps, necessary truth.

I think, though, that the fictionalist can provide an alternative account: it was developed in Field 1980 and elaborated in several of the essays below. On this account, truth isn't required for goodness (so necessary truth isn't required either); what is required instead is something called conservativeness, which embodies some of the features of necessary truth without involving truth. I don't want to say that conservativeness is the *sole* virtue in a mathematical theory – there are secondary virtues, such as interestingness, elegance, the having of applications outside mathematics and so forth, and these secondary virtues are quite important. (A believer in mathematics recognizes these secondary virtues too, in addition to the primary virtue of truth or necessary truth.) But I contend that none of the virtues requires that the mathematical theory be true. It may be, of course, that conservativeness is too weak a virtue to account for the applicability of mathematics to the world – that we need truth in addition – but now I am only explaining what the fictionalist view is, not saying that it is correct.

I have heard two arguments for the unintelligibility of fictionalism about mathematics that deserve comment. Argument one is that it is unintelligible to deny the truth of mathematical assertions like '2+2=4', since it is simply a consequence of the meaning we have assigned to '2', '4', '+', etc., that this and similar assertions hold. It seems to me, though, that this can't be right. There *is* a *somewhat* plausible claim that underlies it: that our number-theoretic concepts are such that we wouldn't count anything as the number 2 unless there were also something we counted as the number 4 and unless we recognized an addition function mapping the thing we counted as 2 and itself into the

[4] Mathematical instrumentalism is sometimes differentiated from mathematical fictionalism: the instrumentalist is said to hold that mathematical claims lack truth value, while the fictionalist is said to hold that they are false. In my opinion this is a totally uninteresting difference; the important point to both positions is that the acceptability of a mathematical claim is in no way dependent on any truth value it may have. In the case of mathematics (and also in the case of physics), where the applicability is evident, I will use the words 'fictionalism' and 'instrumentalism' interchangeably.

thing we counted as 4. In other words, the claim 'If there are numbers then 2+2=4' has some claims to count as an analytic truth, indeed one so obvious that its denial is unintelligible. In a similar way, perhaps, nothing would count as Santa Claus unless it were human, lived at the North Pole, flew reindeer every Christmas Eve, etc. – or at least, did most of those things. But it can't be analytic (or, a purely conceptual truth) that there *is* a Santa Claus that lives at the North Pole and flies reindeer, and it can't be an analytic or purely conceptual truth that there *are* objects 1, 2, 3, 4 etc. obeying such laws as that 2+2=4. An investigation of conceptual linkages can reveal conditions that things must satisfy if they are to fall under our concepts; but it can't yield that there are things that satisfy those concepts (as Kant pointed out in his critique of the ontological argument for the existence of God).

Argument two is that it is simply unintelligible to say that one regards standard mathematics as good, and believes that 2+2=4 is true according to standard mathematics, but yet one doesn't *believe* 2+2=4 (or standard mathematics); what more could there be to *believing* standard mathematics than *regarding it as good*? The Oliver Twist and Santa Claus comparisons are inept, the argument continues. For in the case of the sentence 'Oliver Twist lived in London', one can give clear content to the idea that one only believes that this holds *in the story*: the fact that one doesn't literally believe the claim itself comes out in such facts as that one wouldn't expect to find Twist listed in the records of the London orphanages of the period. But, the argument goes, no analogous content can be given to merely believing that 2+2=4 holds *in the story*, rather than strictly believing that 2+2=4. One can't give content to the claim that one holds a fictionalist view by saying that one denies that an inspection of the universe will turn up any mathematical entities, for of course all but the most mystical platonist would agree that mathematical entities are not available for that kind of inspection.

In reply, it is first worth noting that even if argument two succeeds, it does not undermine what might be called 'weak fictionalism': the doctrine that mathematical claims are true only in the way that fictional claims are true. Rather, argument two is really an argument that any sensible form of platonism *agrees* that mathematical claims are true only in the way that fictional claims are true; what is wrong, according to the argument, is really only 'strong fictionalism', that is, weak fictionalism coupled with the doctrine that weak fictionalism and platonism are to be distinguished.

But this point aside, I do not think that this argument that weak fictionalism and platonism coincide is correct. At least, it does not seem to me correct as applied to the version of fictionalism that I incline towards, the kind which I have defended in many of the papers in this volume. For there are clear substantive differences between fictionalism of that sort and platonism: in particular, the fictionalist differs from the platonist in regarding different theoretical questions as important. More fully, fictionalism of the sort I recommend has both positive content and negative content. Its positive content is that it commits one to abjuring all appeal to mathematical entities in explanations when the chips are down: it must be possible, for instance, to develop theoretical physics without any appeal to mathematical entities. The claim that such an elimination of mathematical entities from explanations is possible is a very substantial claim; there are serious difficulties in defending it, difficulties which have led most philosophers to think that this sort of fictionalism is unworkable. (I think that there are also benefits of showing how to eliminate mathematical entities from explanations, independent of issues about fictionalism – see the end of section 3 of this introduction, and section 5 of essay 6.) Whatever the merits of the eliminability claim, that claim is the positive content of the form of fictionalism I recommend. The negative content of fictionalism is that it avoids having to answer some questions that seem to need answering on a platonist view. This is the main *motivation* for fictionalism, and I will have more to say about it shortly. It seems to me a legitimate matter of dispute whether, at the present stage of knowledge, it is a good exchange to shift the important questions in the way that (this sort of) fictionalism recommends; my present point is only that for better or worse fictionalism does license such a shift, and this shows that contrary to argument two it is an intelligible position genuinely distinct from platonism.[5]

Fictionalism about mathematics is one alternative to platonism. Now, much as phenomenalists have claimed that rejecting matter didn't prevent them from literally believing that they were sitting at a table, some anti-platonists say that rejecting platonism doesn't prevent them from literally believing that 2+2=4 or that there exists a number between 15 and 17; those claims are literally true, the anti-platonist can hold, as long as they are construed in the appropriate fashion. (And 'the appropriate fashion' of construing them needn't be an idealist one: it could instead

[5] There may be some dispute as to whether fictionalism really does differ from platonism in 'negative content': one line of defence of platonism has been to argue that its commitments are less than the anti-platonist contends (though I have never seen this argued very convincingly). But at least by virtue of its positive content, a fictionalism of the sort I defend is substantively different from platonism. Against a form of fictionalism like van Fraassen's (1975, 1980) that would avoid the hard work of showing how mathematical entities are eliminable from explanations, the argument of the previous paragraph has a better chance of success.

be, for instance, a modal construal that involved no reference to any objects other than physical ones, as suggested in Putnam 1967b or Hodes 1984b or Chihara 1984.) In advocating fictionalism, I do not really mean to be opposing such views (though I do mean to refrain from endorsing them): for such views *are* fictionalist about mathematics *taken at face value*, and I have not committed myself to a fictionalism more radical than that.

I should add as an aside, though, that I think that there are difficulties in trying to force anti-platonism into the 'non-standard construal of mathematical theories' mode. One problem arises from the need to give an account of the applications of mathematics to the physical world. In the case of number theory, perhaps, the problems are not so great: for there is one specially central kind of application of number theory to the world, one that works by using numerals to 'encode' numerical quantifiers (see Hodes 1984b); and provided one is willing to utilize a rich enough logic,[6] one can give a general translation procedure for number-theoretic claims that gives them a natural role to play in these applications. But in other cases, such as the theory of differentiable functions of a real variable, there seems to be no single canonical application: one can apply such mathematics to physical space, to degrees of belief, and to much else. I think that this makes much more difficult the problem of finding a translation procedure for the mathematical theory that gives rise to a natural account of its applications. I will not pursue this, though. (I have exhibited an approach to applications elsewhere that does not rely on any translation procedure of the mathematics: it is in Field 1980, but there are some sketchy remarks about it in several of the essays below. Sections 6 and 7 of essay 7 discuss special difficulties for using modal translations of mathematics in applications.) It also seems to me doubtful that forcing one's anti-platonism into the mode of a nonstandard construal of mathematics has benefits that compensate for the difficulties: on this point, see section 8 of essay 7. But there is no need to press this point: the nonstandard

[6] Hodes uses impredicative second order logic and a certain sort of not-purely-logical modality. I am dubious about both; in particular, his view that the former is available to the anti-platonist rests on a sharp distinction between 'objects' on the one hand and 'properties' (or 'Fregean concepts', as Hodes prefers) on the other, and on the definition of platonism as the view that there are mathematical *objects*. To me this definition of platonism seems perverse, for the epistemological and other difficulties that mathematical objects like sets have been thought to raise seem to arise with equal force for the mathematical 'non-objects' (properties or Fregean concepts) that are in the range of the second order variables, and there are other problems there as well (such as the mysteries surrounding Fregean 'unsaturatedness'). But perhaps there are less ontologically loaded logics that would have served Hodes' purposes as well.

construal approach differs from a more thorough fictionalism only in a point of strategy; the anti-platonism is what's important.

2 Initial Plausibility

Fictionalism is often portrayed by platonists as a radical position, quite at odds with the views of the average non-philosopher. I rather doubt that this is so. I don't think it at all obvious that the average person who calculates, or the average physicist who quantifies over mathematical entities while theorizing, or even the average mathematician, literally believes that there are mathematical entities. The average non-philosopher, I suspect, has not thought enough about what platonism involves and what fictionalism involves to have anything like a consistent view of the matter.

But whatever the views of the average non-philosopher, there are considerations that appear to favour the platonist, and which the fictionalist will have to deal with. These considerations have to do with the fact that mathematics is not just an autonomous discipline. Rather, mathematics has many applications outside mathematics – in scientific explanation, in the description of our observations, in metalogic, and in many other areas. It is the fact that mathematics appears indispensable in applications (indispensable without incurring high costs, that is) that provides the main source of arguments for platonism.

In several essays below I have gone further: I have said that the *only* serious arguments for platonism depend on the fact that mathematics is applied outside of mathematics. These remarks (which I stand by) have led quite a few people to accuse me either (i) of ignoring the possibility that mathematics is known by investigating conceptual interconnections, or (ii) of ignoring the reasons that mathematicians actually give for their mathematical beliefs. As for (i), I have already noted in the previous section ('argument one') that reflection on conceptual interconnections cannot yield knowledge of the existence of mathematical entities: it can at best yield knowledge that *if* there are entities that we can correctly call mathematical, *then* they obey the usual mathematical laws. (ii) deserves a more extended discussion, but I will now argue that it too is incorrect.

There are two types of arguments that mathematicians tend to give for their mathematical 'beliefs' – I use the quotes since as remarked there is a good deal of question in my own mind as to whether the typical mathematician literally believes the sentences of standard mathematics. First, the mathematician may argue for a claim by proving it (perhaps sketchily). But what is called rigorous mathematical proof is really just logical derivation of the claim being proved from other mathematical claims (and what is called *sketchy* proof is *sketchy* logical derivation from other mathematical claims). The fictionalist and the platonist agree that logical derivations are to be trusted. The epistemological question about mathematics is not about these logical derivations, but about the mathematical claims used as premises of the derivation: unless they are believed (not just accepted as part of an interesting story), the derivation obviously offers no ground for literal belief in the claim derived. The second kind of argument that mathematicians typically offer for accepting a mathematical claim is that the claim has attractive consequences. Again the fictionalist agrees that it has those consequences, and is likely to agree that the fact that it has those consequences is good reason to accept it for many purposes; the fictionalist *might* indeed go further, and agree that if the attractive consequences are to be literally believed, then the argument gives some sort of hypothetico-deductive reason for literally believing the claim in question. But even this (perhaps dubious) concession does not give grounds for believing the claim in question unless those attractive mathematical consequences are themselves literally believed; and a fictionalist mathematician will refrain from literally believing them.

But, it will be said, I am ignoring the fact that mathematicians (and ordinary people) find many mathematical claims *initially plausible*; plausible *independently of argument*. Well, perhaps they do and perhaps they don't. I certainly grant that mathematicians and ordinary people find many mathematical claims *natural to accept* independently of argument, where 'accept' means simply 'incorporate into one's mathematical story'. There are obvious reasons why some mathematical claims are very natural to accept in this sense, and these reasons don't presuppose platonism. The reasons for finding certain mathematical claims natural (even if not literally believable) will vary somewhat from one mathematical claim to another. For instance,

(a) {Human females} $\cup$ {human non-females} = {humans}

is natural to accept largely because of its intimate association with the logical truth '$\forall x$(x is human if and only if either x is human and female or x is human and not female).' Similarly,

(b) $1+1=2$

is natural to accept largely because of its intimate association with logical truths like 'If there is exactly one apple on the table and exactly one green thing on the table and no apple on the table is green then there are exactly two things on the table which are either apples or green'; here 'there is exactly one' and 'there are exactly two' are numerical quantifiers, definable in terms of ordinary quantifiers plus identity. (The fact that (b) is intimately associated with certain sentences involving numerical quantifiers is, of course, what lies behind most standard applications of the natural numbers to reasonings about the physical world. The fictionalist view can easily make sense of these applications: see essay 2 and the postscript thereto.) The situation is similar with

(c) Between any two real numbers there is another real number;

this is natural to accept largely because of its intimate association with an analogous claim about points on a line in physical space. (Again, this intimate association is involved in some though not all of the most familiar applications of the real numbers.)

(d) For any physical objects x and y, there is a set containing x and y as its only members

draws some of its naturalness from the claim that there is an *aggregate* of x and y; admittedly, the claim about aggregates gives direct naturalness to (d) only in the case when x and y don't overlap, but the extension to the case where they do overlap is also natural because of its simplicity. And in each of cases (a)–(d), part of the naturalness of the claim doubtless comes from our mathematical education: taking (b) as an example, we are taught from early childhood both that $1+1=2$ and that we can pass freely between numerical quantifications and corresponding claims about numbers.

There are of course mathematical claims that unlike (a)–(d) are not intimately connected with non-mathematical claims, and which seem natural even when we first hear them, so that education is not directly a factor: an example is

(e) There are inaccessible cardinals.

But a fictionalist will find these natural too: they are natural ways to extend the 'story' of Zermelo–Fraenkel set theory. (We find some extensions of ordinary pieces of fiction more natural than others, so why not in mathematics?) It seems to me quite tendentious to take the uncontroversial fact that some mathematical claims are, for diverse reasons, natural to accept, and redescribe it in terms of their being 'initially plausible'.

But suppose that most mathematicians do find these claims initially plausible, rather than just natural. What follows? Well, one can try to make it look like a lot follows by drawing an analogy between initial plausibility judgements in mathematics and perceptual judgements about the physical world. There is some plausibility to the analogy: after all, part of what makes a claim like 'This is red' perceptual is that it is believed *independent of argument* in the appropriate circumstances, so

it too could be said to be 'initially plausible' in those circumstances. Given the perceptual analogy, should we not say that mathematical claims like (a)–(e) are in as good epistemological shape as perceptual judgements about ordinary objects?

There are two reasons to doubt this. One, which I shall *not* discuss now, is that there seems to be a crucial difference between the cases: the difference arises from the fact that there are unproblematic connections (typically causal connections) between what we perceive and our perceptual judgements, whereas there are no such unproblematic connections in the case of plausibility judgements in mathematics. (Some brief remarks on the significance of this difference can be found in section 4 of this introduction, and in section 2 of essay 7.) But the other reason for doubting the moral suggested at the end of the previous paragraph is that it rests on a very naive view of the epistemological significance of perception.

Whatever the defects of coherence theories of knowledge, one thing that seems right about them is that claims of 'direct perception' never have unchallengeable status. More specifically, not just individual perceptual judgements but whole practices of making perceptual judgements of a certain sort can be questioned, *when an alternative proposal for a perceptual practice is available.*

Here's an example (deriving from Feyerabend 1975). It may well be that the epistemic practices of astronomically ignorant people licensed the assertion 'The sun is rising' in the observational circumstances typical of the early morning, and that this assertion was always understood literally, as meaning that the sun was in absolute upward motion. But it is hard to deny that when the astronomical theory implicit in this observational practice was undermined, it became rational to modify the practice (either by modifying the words used to report observations; or by understanding these words differently; or by using the same words understood the same way, but no longer literally believing the observational reports and believing instead only that they are a useful though literally false way of conveying the truth about the change in angle between the horizon and the sun). *It seems to be part of our general methodology to consider alterations of our perceptual practices under certain circumstances: in those circumstances, we compare the old perceptual practice with a proposed new one, and make the alteration if it seems to lead to better results.*

A platonist might respond as follows: 'All this analogy suggests for the mathematical case is that we shouldn't hold on to certain types of mathematical plausibility judgements after they have come into conflict with other mathematical plausibility judgements or with perceptual judgements. But mathematical judgements never do come into direct conflict with perceptual judgements, and if we only revise mathematical plausibility judgements when they conflict with other mathematical plausibility judgements we will never be led to a fictionalist position.' But a platonist who so responds has failed to appreciate the force of the example. What happens in cases like the above is *not* that the old observational practice as a whole undermines a particular part of that practice (the part I've cited), by lending support to an astronomical theory that contradicts that part of the observational practice. For the fact is that the new astronomical theory could never be thought of as well-supported as long as one relied on the old perceptual practice. (The old observational practice leads so immediately to an astronomical theory in which the sun is in absolute upward motion in the morning that it is hard to see how a simple amalgamation of initial credibilities from the totality of our observations can lead to a new theory that radically conflicts with this.) What happens in the example is rather that the mere *suggestion* of an alternative perceptual practice together with some sketch of what life would be like if we accepted that alternative practice is enough to cast enough doubt on the original practice so that an alternative practice deserves a fair hearing. If the new practice looks better we shift to it; we need not (and in this example cannot) argue to it on the basis of the old perceptual practice.

In this example, in fact, it is plausibly maintained that the new practice can be argued to be better than the old practice *independently of any specific problems with the old practice*. For the new practice (involving only judgements about angle from the horizon, rather than about absolute motion) is less committal than the old practice and yet serves all our purposes just as well, so that it has advantages of economy over the old practice. If one accepts this, then the bearing on platonism is apparent: for the fictionalist's judgements about naturalness are less committal than the platonist's plausibility judgements, so one should shift to the fictionalist's practice unless one can argue that there are respects in which the platonist's practice serves our purposes better. *And the only way to argue that would be to appeal to considerations independent of initial plausibility, such as indispensability arguments.*[7]

[7] One can't very well avoid this conclusion by simply saying that the platonist practice has advantages over the fictionalist practice, in that it leads us to truths about mathematical entities; that would be like saying that the perceptual practice that licenses beliefs about absolute motion has advantages over the new one, in that it leads us to the facts about absolute motion. The point is that the alleged 'facts' about absolute motion need to be shown to be important; similarly for the alleged 'facts' about mathematical objects. Maybe they can be shown to be important, for instance by indispensability arguments: here I am only arguing that that is where the interesting issues lie.

The platonist might want to protest that what undermined the old perceptual practice in this example isn't just the existence of a new perceptual practice that is less committal, but also a critique of the old perceptual practice. That is, part of what was involved in the development of the new perceptual practice was the development of and/or emphasis on the concept of relative motion; *this then led to the realization that it is only relative motion that can be reliably detected by direct visual observation*, and it is that critique of the old observational practice, not *just* the development of the alternative practice and it's being less committal, that led to the undermining of the old practice. I have to agree that this response to the example makes a certain amount of sense. But it should be noted that a platonist who responds in this way implicitly admits that the platonist practice of making initial plausibility judgements about mathematical entities could be undermined if it were subjected to an analogous critique: that is, the platonist practice would be undermined if it could be argued that even on the supposition that there are mathematical facts, there is no way in which our plausibility judgements could be expected to reliably reflect them. Arguments of roughly this form *have* been suggested by various people (see section 4 of this introduction, and section 2 of essay 7). So our discussion leads, if not to the strong conclusion of the last paragraph, then at least to this weaker conclusion: even if we grant that initial plausibility judgements in mathematics are analogous to perceptual judgements, still arguments of the sort just alluded to might undermine them, and lead to their replacement by the less committal judgements of naturalness that even the fictionalist accepts.

There is another, simpler, route to the conclusion I have drawn – indeed, to the stronger conclusion of two paragraphs back, rather than the more concessive one of the last paragraph. This argument too will be made by analogy. Consider an ontological dispute about, not mathematical entities, but another somewhat controversial sort of entity: regions of space. Let us use the term 'substantivalist' for a person who believes that talk of regions of space is true taken at face value, that is, without reconstrual in terms of physical objects. Now, if one is a substantivalist, it will make perfectly good sense to say that one makes perceptual judgements about regions of space – for instance, it makes sense to say that one perceives that the region of space in front of one is empty of visible objects. So claims about space can have *perceptual* initial plausibility, and that is the kind of initial plausibility whose epistemological value is presumably least controversial. *Even so*, the claim that one can have perceptual knowledge of physical space does nothing to undermine a challenge to substantivalism (and to the practice of making substantivalist perceptual judgements) on the basis of ontological economy.[8]

The reason, as before, is that observational reports are always theory-laden; and this example (like the previous one) makes clear that if the theory with which they are laden is challenged by someone with an alternative theory, one's practice of making observational reports in the way that one does needs defence against that challenger. Faced with a relationalist who thinks that there is no need to postulate physical space (as anything other than a manner of speaking about ordinary objects), a believer in physical space needs to reply on the relationalist's own terms: to claim that one has perceptual knowledge that there are regions of physical space devoid of visible objects would in *that* context carry no epistemological weight. In essay 6 below I attempt to meet the relationalist on his or her own terms, with two types of indispensability arguments. My present point is simply that the platonist ought to respond to the fictionalist in the same manner, and that until he or she does so there is no reason to persist in regarding mathematical claims as initially plausible (rather than simply as being natural to accept for diverse reasons of the sorts I have described).

3 Indispensability Arguments and Inference to the Best Explanation

As I said early in the previous section, there *are prima facie* difficulties in maintaining a fictionalist position about mathematics: at least, there are for anyone who takes indispensability arguments seriously, as I do. An indispensability argument is an argument that we should believe a certain claim (for instance, a claim asserting the existence of a certain kind of entity) because doing so is indispensable for certain purposes (which the argument then details). In this section I will focus on one special kind of indispensability argument: one involving indispensability *for explanations*. (There is some discussion of other kinds of indispensability argument in several of the essays in this volume, especially essays 3 and 7.) To rely on this special kind of indispensability argument is to rely on a principle of 'inference to the best explanation'. Some such principle seems to underlie much of our knowledge of the physical world.

[8] I don't say that the claim that space-time regions are perceivable is of no epistemological importance: my claim is rather that it has importance only in a different kind of epistemological context than that which is here in question. In this different context, the existence of unproblematic connections between perceptual judgements and perceptual facts, not just the initial plausibility of perceptual judgements, are important. I will return to this in section 4.

More fully, suppose (a) that we have certain beliefs, beliefs about 'the phenomena', which we are unwilling to give up; (b) that this class of 'phenomena' that we believe in is large and complex; (c) that we have a pretty good explanation of these phenomena (in the sense of, a relatively simple non-*ad hoc* body of principles from which they follow); and (d) one of the assumptions that appears in this explanation is claim S, and we are pretty sure that no explanation of the phenomena that does without claim S is possible. The idea of 'inference to the best explanation' is that under these circumstances we have a strong reason to believe claim S. (If this seems vacuous, I should add that we do not accept as explanations claims of the form 'The phenomena are as they would be if explanation E were correct': as-if claims which ride piggyback on genuine explanations are not themselves to be construed as explanations (at least, non-*ad hoc* ones), for the purpose of understanding this principle. Given this, the principle, though of course vaguely formulated, is non-vacuous: it precludes us from accepting a large and complex set of phenomena as brute when there is a decent explanation of these phenomena in the offing.)

It seems to me that most of us accept the principle of 'inference to the best explanation', in the sense that this principle (or something pretty close to it) governs our ordinary inductive methodology.[9] Clearly something like inference to the best explanation is at work in us in giving rise to observational beliefs, that is, beliefs that could be independently checked by observation: we come to believe that a pipe behind the wall is leaking, since this best explains the stains on the wall-paper, the warped floor-boards, etc. The same principle seems to be involved when we arrive at beliefs about unobservable entities, or non-observational beliefs about observables. The principle of inference to the best explanation makes no discrimination among these three cases: if a belief plays an ineliminable role in explanations of our observations, then other things being equal we should believe it, regardless of whether that belief is itself observational, and regardless of whether the entities it is about are observable. That I think is the methodology we (nearly) all employ, and I think it would be unwise to change it. The fact that the principle does not discriminate over whether the explanation is observational (or whether it postulates unobservable entities) stands up well to reflection: intuitively, the observational nature of the explanation should make no difference in an inference to the best explanation. After all, in *any* case where we rely on inference to the best explanation, our belief goes beyond what we have observed; the fact that one belief *could* be fairly directly tested by observation while the other *couldn't*

[9] For a contrary view, see van Fraassen (1980).

seems to have no relevance to their evidential status *when such an independent test has not been made*. (When the independent test *has* been made – when the leak behind the wall has been directly observed – then we need no longer rely on inference to the best explanation. When we do rely on inference to the best explanation, our beliefs go beyond the observations we have made, and my point is that the difference with respect to *possible* observations that *haven't* been made is irrelevant to our *actual* evidential situation.)

To be sure, it would be possible to introduce a restricted form of inference to the best explanation that licenses only observational beliefs; but, firstly, this would be entirely *ad hoc* and unmotivated,[10] and, secondly, such a restriction would cripple our beliefs about observables. Many observational beliefs (e.g., the beliefs that the Los Alamos scientists had about what would happen when they made the first atomic test) seem to depend on beliefs about unobservables; they are not obtainable directly from other claims about observables, without detour through the unobservables. (For instance, they aren't obtainable from past observations by any straightforward sort of enumerative induction, since they concern observable situations very different from any hitherto encountered.) I don't deny that one could formulate a principle that would allow belief in the observable (though hitherto unobserved) consequences of inferences to not observationally checkable explanations, without believing in those explanations themselves, but such a position seems to me even more *ad hoc* than the simple restriction to observationally checkable explanations. (Also, the fact that it would deprive us of believable explanations of the things we believe – deprive us of a simple and unified body of beliefs from which many of the rest of our beliefs follow – seems to me to make it intrinsically unattractive.)

But if our belief in electrons and neutrinos is justified by something like inference to the best explanation, isn't our belief in numbers and functions and other mathematical entities equally justified by the same methodology? After all, the theories that we use in explaining various facts about the physical world not only involve a commitment to electrons and neutrinos, they involve a commitment to numbers and functions and the like. (For instance, they say things like 'there is a bilinear differentiable function, the electromagnetic field function, that

[10] If one is going to make an unmotivated restriction that we use inference to the best explanation only to arrive at observational beliefs (beliefs directly checkable in principle by human beings, wherever and whenever located), why not the further unmotivated restriction that we use it only to arrive at beliefs directly checkable *by me*, or directly checkable in principle *by beings located in this galaxy during the check*, or directly checkable *in an experiment cheap enough for the government to fund*?

assigns a number to each triple consisting of a space-time point and two vectors located at that point, and it obeys Maxwell's equations and the Lorentz force law.') I think that this sort of argument for the existence of mathematical entities (the Quine–Putnam argument, I'll call it) is an extremely powerful one, at least *prima facie*. It should be noted that the Quine–Putnam argument is not merely that just as there are good explanations in which the postulation of unobservables is essential, so too are there good explanations in which the postulation of mathematical entities is essential, so that if inference to the best explanation licenses one it licenses the other. The argument is stronger, in that it says that the very same explanations in which the postulation of unobservables is essential are explanations in which the postulation of mathematical entities is essential: mathematics enters essentially into our theory of (say) electrons. There seems to be no possibility of accepting electrons on the basis of inference to the best explanation, but not accepting mathematical entities on that basis, by saying that the explanations involving the latter are weaker than the explanations involving the former: for the very same explanations are involved in both cases. This fact makes arguments for the indispensability of mathematical entities *in explanations of the physical world* seem in some ways more compelling to a scientific realist than other indispensability arguments. (Not that other indispensability arguments may not be compelling too.)

Still, two points need to be made about justifying belief in mathematics by its apparent indispensability in science (or elsewhere). Together, they make the ultimate force of the Quine–Putnam argument hard to ascertain.

The first point is that any such 'indispensability to science' justification of mathematics can be undercut if we can show that there are equally good theories and explanations that don't involve commitment to numbers and functions and the like. I believe that such an undercutting of the justification is possible (but that no analogous undercutting is possible in the case of our justification for believing in electrons and neutrinos and the like). I originally made a case for this in my book *Science without Numbers*.

At present of course we do not know in detail how to eliminate mathematical entities from every scientific explanation we accept; consequently, I think that our inductive methodology does at present give us some justification for believing in mathematical entities. But this brings me to my second point, which is that justification is not an all or nothing affair. The belief in mathematical entities raises some problems which I and many others believe to be fairly serious. (I will briefly discuss two of those problems in the next section.) These puzzles provide reasons against the belief in mathematical entities, and to put it very crudely, what we must do is weigh the reasons for and the reasons against in deciding what to believe. Less crudely, what we must do is make a bet on how best to achieve a satisfactory overall view of the place of mathematics in the world. I do not declare that it is misguided to try to solve the puzzles that many people have found in platonism; on the contrary, much interesting work is being done in that direction, and I hope that a large segment of the philosophical community continues to pursue that line of research. Let a hundred flowers bloom. But my tentative bet is that we would do better to try to show that the explanatory role of mathematical entities is not what it superficially appears to be; and the most convincing way to do that would be to show that there are some fairly general strategies that can be employed to purge theories of all reference to mathematical entities.[11] In any case, I think it is an idea that needs to be pursued much further than the philosophical community has pursued it so far.

I think, indeed, that showing how to eliminate mathematical entities from explanations would be attractive for reasons not wholly dependent on anti-platonism. For *even on the assumption that mathematical entities exist*, there is a *prima facie* oddity in thinking that they enter crucially into explanations of what is going on in the non-platonic realm of matter. It seems to me that the most satisfying explanations are usually 'intrinsic' ones that don't invoke entities that are causally irrelevant to what is being explained.[12] 'Extrinsic' explanations are acceptable (as when we explain the behaviour of a non-human or non-English speaker by reference to English sentences that he or she believes or desires), but it is natural to think that for any good extrinsic explanation there is an intrinsic explanation that underlies it. This principle seems plausible independently of anti-platonist scruples (it wouldn't help to refer to inscriptions of English sentences instead of sentence types of English in our explanation of the non-human's or foreigner's behaviour); but it requires that there are mathematical entity free explanations underlying the platonistic ones, since mathematical entities (according to the views

[11] The result of such a purging would be a theory that a physicist would probably regard as merely a rewritten version of the original theory; just as a physicist would probably regard Newtonian physics formulated without talk of absolute rest as merely a rewritten version of Newtonian physics with absolute rest. But for a philosopher, there is an advantage in formulating Newtonian physics without absolute rest, and so too for formulating it without mathematical entities. What counts as two formulations of the same theory is a context-relative matter.

[12] This is over-simplified: a more careful formulation is needed to handle non-local quantum phenomena; also, independent of quantum considerations, a suitable kind of spatio-temporal connection may be enough for 'intrinsicness' even when the causal condition fails. (The alternative condition in terms of spatio-temporal connection might be enough to handle the quantum phenomena too.)

of those who believe in them) don't enter into causal interactions with the material world (in that for instance there is no interchange of energy-momentum between them and the material world).[13] I regard the acceptance of an extrinsic explanation as ultimate as at least slightly odd.

Whether or not one takes this *prima facie* oddity seriously as a motivation for showing that mathematical entities are in principle eliminable from physical theory, it does point up an important fact: the role of mathematical entities, in our explanations of the physical world, is very different from the role of physical entities in the same explanations. For the most part, the role of physical entities in those explanations is causal: they are assumed to be causal agents with a causal role in producing the phenomena to be explained. Since mathematical entities are assumed to be acausal, their explanatory role (or roles) must be somehow different.

It may be thought that the difference in the explanatory role of mathematical entities and physical entities is enough to motivate a restriction of inference to the best explanation to the latter. The position would be (1) that we should literally believe in the existence of electrons and their properties as postulated in our physical theories, since there are good explanations in which they are assumed causally relevant, and

[13] Here I exempt the views of Popper (1972), who seems to hold that mathematical entities can causally affect our immaterial minds, and thereby indirectly affect our bodies by way of our pineal glands. (Gödel is sometimes taken to hold this too, on the basis of some much-quoted remarks in his 1944 and 1947.)

Note also that weakening the notion of intrinsicness in accordance with the previous footnote is unlikely to help, since, among other things, mathematical entities are usually taken to have no spatio-temporal location. Again there is an exception: Maddy (1980) holds that at least some mathematical entities have spatial location. ({Reagan}, {{Reagan}, Reagan}, and so forth, all have the same location, namely that of Reagan himself; similarly, {Reagan, Carter}, {Reagan, {Carter}}, etc. all have the same location as does the aggregate of Reagan and Carter. The number 3 is a property of all three-membered sets; presumably each region of space is the location of a three-membered set (say, the set of its left third, its middle third and its right third), so that 3 is a property that is instantiated everywhere. I don't know if the number 3.782 and the exponential function have spatial location on her view, or if she views them as properties of things that have locations; but I suspect that if they do have or apply to things that have location, they are located or instantiated everywhere.) Maddy's view, unlike Popper's, does not strike me as at all silly. Rather, it is a perfectly sensible convention about how to talk about mathematical entities. I doubt, though, that it is enough to resolve worries about the explanatory legitimacy of the use of mathematical entities in explanations, especially the more complex ones typical of advanced physics. I should add that it was not her intention to deal with *these* worries; she was concerned rather with the epistemological issues that we will deal with in the second half of the next section. I suspect that the doubts I have raised about how helpful it is to assign sets spatial location apply in that context too, but there the issues are in some ways more complicated.

there is no obvious prospect of eliminating them from these explanations; but (2) that we shouldn't literally believe in mathematical entities, since there are no good explanations *in which they are assumed to be causally relevant*, despite the fact that there is no way of giving explanations that avoids postulating them in an *acausal* role. The problem is that if one takes this line, then the properties of electrons that one literally believes in can't include any properties that require mathematical entities for their expression; so if mathematics is not eliminable or close to eliminable, there are going to be very serious limitations on stating explanations in terms of electrons without going beyond what one believes. Perhaps one can maintain a belief in electrons and a belief in those of their properties which are describable without mathematics, on *something like* inference to the best explanation grounds, without a belief in the explanations one gives; but it seems to me a very delicate position to maintain.[14] Consequently I am inclined to think that unless a very substantial amount of explanation involving electrons can be given in a mathematical entity free fashion, the prospects for maintaining realism about electrons without maintaining platonism are dim.[15]

4 Problems with Platonism

I do not intend to discuss in detail the various problems that many have felt to arise in the platonist position. It is, though, worth commenting briefly on two of them.

A. *'What Numbers Could Not Be'*

A noteworthy feature of mathematics is that there is a tremendous amount of arbitrariness as to the identification of different types of mathematical objects. The most famous example of this is the one highlighted by Paul Benacerraf (1965): if one wants to identify natural numbers with sets, it seems rather arbitrary which sets one picks. But of course this example is just the tip of the iceberg: just as there is no uniquely natural set-theoretic explication of natural numbers, there is no uniquely natural set-theoretic explication of real numbers (for instance, one can use Dedekind cuts or equivalence classes of Cauchy

[14] See Cartwright (1983) for an attempt to maintain it.

[15] I like to set my goals high – I like to set myself the task of entirely doing without mathematical entities in physical theory. It may prove in the end that this task cannot be carried out, but that substantial parts of it can be carried out: that substantial bodies of explanation can be given in a mathematical entity free way, even if not full physical theory. If that were to turn out the case, I think that the position contemplated in this paragraph would need to be seriously considered.

sequences); or of ordered pairs; or of the tensor product of two vector spaces; or of the tangent vectors at a point on a manifold; and so on *ad infinitum*. It seems absurd to suggest that in each such case there is a fact of the matter as to which sets the mathematical objects in question are. Indeed, the problem is deeper than this: for there is also an arbitrariness in thinking that it is sets that are to be regarded as the basic objects. As Benacerraf mentions, it is possible to take ordinal numbers as basic and define sets in terms of them. Also, we could take functions as basic, and define sets as functions of a certain sort (say, identity functions); or we could take relations as our basic objects, and define sets as relations of adicity 1. The choice between regarding relations as sets of a certain sort and regarding sets as relations of a certain sort seems no more factual than the choice between regarding natural numbers as Zermelo sets and regarding them as von Neumann sets; or between regarding real numbers as Dedekind cuts or as equivalence classes of Cauchy sequences; or between regarding the tensor product of the vector spaces V_1 and V_2 as a space whose elements are equivalence classes of formal sums of pairs of vectors, and regarding it as a space whose elements are the linear functionals on the bilinear forms on the direct sum of V_1 and V_2.

But what exactly is the significance of the arbitrariness in these choices? This is of course quite controversial. In the case of the identifications of natural numbers with sets, some philosophers (for instance, Steiner 1975) have taken the position that the arbitrariness of one identification over the others shows that all such identifications are false: numbers and sets are simply distinct sorts of entities. Perhaps this position has some plausibility in this case (and in the case of the identifications of real numbers with sets, and in a few other special cases); but surely as a general position it defies credibility. Are we to say, for instance, that a topological space is distinct from any set-theoretic construct, since it could be taken to be either a pair of a set and the set of its open subsets, or of the set and its set of closed subsets, or of the set and its closure operator? And are we to say the same for each other mathematical structure, even the arcane and specialized ones? Perhaps saying this is *slightly* more natural than saying that topological spaces *definitely are* set-theoretic constructs of some particular sort. But I think that by far the most natural conclusion is that topological spaces, and numbers, and ordered pairs and functions are neither definitely sets nor definitely not sets: there is no fact of the matter about whether they are sets or not, in addition to there being no fact of the matter as to which sets they are if they are sets.

Now, supposing that the arbitrariness in identifying one type of mathematical object with another is as pervasive as I have been saying