Mathematics, Meaning, and Commitment

by

Mark Charles Reginald Smith

A thesis submitted to the Department of Philosophy

in conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

April 2009

ABSTRACT

This thesis offers a defence of realism in mathematics, principally through an *a priori* re-construal of the indispensability argument as revelatory of the commitments of mathematical thought. Following the articulation and defence of that argument—called the 'argument from constraint'—in the first chapter, I address semantic and metaphysical consequences and objections in turn. The second chapter broaches the value of Fregean semantics for arithmetical language, and the indispensability of semantic considerations in accounting for mathematical thought. The third chapter assesses and responds to objections from the fictionalist account of mathematical thought and meaning, and the final chapter articulates a programmatic proposal for mathematical epistemology.

ACKOWLEDGEMENTS

I owe many thanks to Adèle Mercier, Henry Laycock, and Joshua Mozersky for their criticism, insight and contributions over the course of this project. My thanks also to Ole Nielsen and James Robert Brown for their valuable comments.

I owe a debt of gratitude to David Bakhurst and Sergio Sismondo for their support, and to Alistair MacLeod, who gave me the opportunity for my first teaching experience.

To Maxine Wilson, Jackie Doherty, Marilyn Lavoie and Judy Vanhooser, all my thanks for pulling me out of the fire and sorting me out more often than I care to remember. I'm grateful for every minute of your help and conversation.

To my students over the years: you have taught me a great deal, and you continue to make philosophizing worthwhile.

I owe much more than I can express here to Amanda Gibeault. Suffice it to say that your presence, Amanda, infuses life with a deeper significance.

To my mother, Maria Smith-Bruynes: my heartfelt gratitude for your continuing love and support. The way has been long and you've seen me through every step and bend. I couldn't have done it without you.

I dedicate this work to the memory of my father,

Reginald Charles Smith (1935 – 2000).

I miss you.

Table of contents

CHAPTER 5: Structure and mathematical knowledge

1

Introduction

1. A proposed reorientation

This essay offers a defence of object-realism in mathematics, in a sense of 'realism' that will emerge gradually over the course of the first two chapters. The defence begins with an assessment of the indispensability argument, and a proposal for reorienting that argument in a novel direction.

The familiar form of the indispensability argument (IA) for the existence of specifically mathematical abstract objects invokes their indispensability for empirical science. I borrow the following fairly standard formulation from Colyvan (2001: 7):

> If apparent reference to some entity (or class of entities) ξ is indispensable to our best scientific theories, then we ought to believe in the existence of ξ.

The idea might be cast like this: apparent reference, combined with indispensability (in some sense to be made more precise) in the formulation of a body of theory which we know on independent grounds to be true, *secures* reference. Commitment to mathematical objects, by these lights, is the very same sort of thing as commitment to quarks and other unobservable entities. If 'quark' appears to refer, and despite our best eliminative efforts we cannot achieve the same scientific success without reference to quarks as we can with such reference, then 'quark' refers. In a similar vein, if 'the real line' appears to refer, and we cannot achieve the same degree of explanatory success in physics without the real line as we can with it, then the expression refers.

This is an *a posteriori* argument: one (or more, depending on the formulation) of its premises is empirical. The argument's fulcrum is the extent of the empirical support we can bring on behalf of this thesis: that mathematics cannot, without loss of explanatory content, be paraphrased out of empirical science. An IA of this kind is, then, an empirical hypothesis. In what follows I will not take a very firm stand on the empirical conjecture, except to use it as a starting point for reflection in a new direction. The arguments I offer here present a different version of IA, and a reorientation of its focus. This version is cast *a priori*, and argues for *metaphysical* rather than empirical indispensability.

By and large it's not a bad idea to hold that the burden of proof in metaphysics should fall on whoever is looking to inflate the ontology. Whoever wants to have something added to the inventory of what there is needs to make a case for inclusion. But if the burden falls to the inflators, it is worth wondering what basic ontology is somehow given in advance, what basic ontology we are thinking about inflating. This is where my version of the indispensability argument comes in: such an argument is not meant to crowd ontology unnecessarily, as is sometimes charged, but to supply convincing reasons for thinking that, in the case that will interest me here, there already are mathematical objects in the basic ontology that lies in the warp and weft of many of our true beliefs and successful practices. My claim will be that mathematical objects are indispensable in the explanation of features of true thought and features of the inferential structure in which specifically mathematical thought occupies a place. So while the arguments that follow proceed on metaphysical, semantic and epistemic fronts in turn, ultimately the goal is to situate these features of mathematics in the theory of thought.

2. Positions

A strong allergy to the imperceptible is a recurring feature of empiricism: sensory evidence connects us to the particular objects of knowledge, and empirically describable cognitive processes allow us to weave together the data of experience into a coherent whole. The weaving together is naturally partly theoretical, and our cognitive processes – especially of abstraction from instances to generalities – always outstrip the evidence they take as their grist, but on the whole this epistemic outlook is plausible and fruitful. In the modern period, Hobbes, Locke, and Hume defended the substantial insight that human knowledge is a natural process in a natural world, and that human integration with the rest of nature is achieved through the stream of causal interactions between sensory information and the natural psychology of processing. Given this outlook, the twin questions of the possibility of mathematical knowledge and the nature of mathematical thought and its objects pose a genuine problem, and it's one to which a number of solutions have been proposed, from Locke and Hume, through the Kantian synthesis and all the way to Kitcher (1983) and beyond. But there is also an epistemic tradition that sees mathematical knowledge not as a problem that empiricism needs to solve, but as a *refutation* of global empiricism. This is the tradition stemming loosely from Plato (though Plato was also its first critic: see his *Parmenides*), embodied in the early modern period by Descartes and the rationalists, and thriving today in a realist tradition that stems largely from Frege. It is a tradition that takes seriously the claim that mathematical thought and talk do not need to be re-analyzed and reformed in the light of a prior epistemology or of antecedent metaphysical commitments, but rather any comprehensive philosophy needs to take account of what kind of beast mathematics really is, and accommodate its presence as a datum to be accounted for. The general orientation of this essay is with the latter school.

At the broadest level on the metaphysical front, the current territory is staked out between realists and anti-realists, or between platonists and nominalists, but these divisions are too coarse-grained to be enlightening. Each camp divides in turn into many, and there are sometimes greater affinities across realist and anti-realist positions of certain kinds than there are within one or the other stance. It is, for instance, fair to say that the nominalism of Boolos has more in common with the neo-Fregean realism of the Scots School of Wright and Hale, than with Boolos's fellow nominalist Field. Positions cross-pollinate and hybridize as they unfold. Nevertheless, I offer arguments to the effect that we should plant both feet in the realist camp with respect to mathematics, or at least to a core portion of mathematics. The core portion at issue is the portion that philosophers of mathematics usually have in mind, and often tacitly. It is essentially composed of number theory ('arithmetic' very broadly construed so as to include analysis and algebraic number theory) with set theory as its foundation and the logic of sets as its language. A good deal is unapologetically left out: statistics, probability and geometry are generally ignored, and will be here as well. When I speak of 'mathematics' in what follows, I will use the term in the way that *philosophers* of mathematics use it, not in the more catholic sense.

3. Reading Frege

Since Frege's thought plays an important role in the positions I will articulate and defend, I should say a word about my reading of his work. I see a substantial unity and continuity in his thought from the *Begriffsschrift* of 1879 all through his major works until his very late period, when he appears to have abandoned the metaphysical core of his own views and espoused Kantianism in the foundations of mathematics. I take *The Foundations of Arithmetic* (1884) as the central statement of his philosophy, and I read most of his subsequent major writings as a variety of contributions to, and struggles with, the elements of the overarching programme that that book set out. Thus, my reading of Frege sees him as primarily a philosopher of mathematics, who was led to deep reflection in the philosophies of language and thought by his central commitments in the philosophy of mathematics. Those commitments are as follows:

- *Logicism*. This is the thesis that all of mathematics (in the sense of 'mathematics' sketched above) is derivable from, and achieves its epistemic security in virtue of, principles of logic combined with explicit definitions.
- *Mathematical platonism*. Numbers are self-subsistent objects, in a logical sense of 'object' to be explored.
- *Semantic realism*. Mathematics is a body of theory that has a subject-matter and a determinate content; it is not to be understood as a formal game having no reach outside itself. In general, truth values are determinate and mind-independent.
- *The integration of mathematical thought and thought in general*. The thesis here is that mathematical thought is not substantially different in kind from thought in general. A comprehensive theory of thought, which includes the theory of inference as a component, must integrate mathematics if it is satisfactory at all.

To these theses we might add anti-psychologism, but it follows from the second and third points combined. In this essay I will be concerned with all but the logicist thesis, about which I remain agnostic. The first three points are explicit in *Foundations*; the last is largely tacit, but it pervades the book and floats to the surface through the examples that

Frege raises and through his insistence that we must not lose sight of the semantic (and thus thought-involving) continuity between mathematical and natural language occurrences of numerical expressions, and the inferences that we make in both mathematical contexts and the ordinary course of everyday life. It is particularly fruitful to read the crucial subsequent works 'On Concept and Object' (1892), 'On Sense and Reference' (also 1892), 'Thought' (1918) and others besides through the lens of the 'integrationist' thesis, as I will illustrate in chapter 2 below. The *Basic Laws of Arithmetic* (vol. i 1893, vol. ii 1903) stands, of course, as the attempt to work out the logicist thesis in formal detail, though it also embodies the other three. Boolos (1999) and Hale & Wright (2001) have shown that logicism emerges as far more promising than Frege himself came to believe, but I will say very little about logicism in what follows.

This reading of Frege has the virtue of explaining why he addresses various problematic issues in the philosophy of language, such as the appropriate semantic treatment of empty names, fictional discourse, *oratio obliqua* and so on more as themes in passing than as central topics in their own right. These features do not crop up in mathematics or in language 'directed toward truth and truth alone' belonging to 'scientific exposition' (Frege 1918: 357). It may be that my reading is unfaithful to the historical evolution of Frege's thought, but I shall make a case for the merits of a Fregean view of this kind regardless. The merit I claim for it is independent of its historical accuracy.

4. Abstract objects and material adequacy

Logic, metaphysics, semantics, and thought are the respective themes of each of the theses set out above. There are many possible ways of casting what binds them, but in this section I want to select one that will come to the fore in the positive arguments that unfold in the first two chapters.

The logic of *Begriffsschrift* inaugurated the modern discipline through two principal innovations: the analysis of propositions into function and argument(s) rather than into subject and predicate in the manner of Aristotle, and quantificational theory. The former breakthrough allows for *n*-place functions and thus affords the technical apparatus required for a logic of relations, which is obviously crucial for a precise articulation of the structure of functional expressions in mathematics. The second feature, the theory of quantifiers, arises out of the first and makes possible the logical treatment of expressions of multiple generality, as in "every number (greater than 0) has a successor": $\forall x \exists y\ (Syx)$. This new expressive power in coding the structure of *n*-place quantified expressions is combined with a theory of valid inference which captures vastly more than what can be coded in the Aristotelian syllogistic or in propositional logic on its own. The notion of logical validity itself, though, is preserved across syllogistic and modern logic: a valid argument is one whose conclusion is true in every case in which all its premises are true. The explanation of validity lies purely in the form of the expressions deployed in an inference: logical validity is indifferent to the semantic assignment of cases, it is indifferent to interpretation. Things are not quite as straightforward in the case of mathematics, though; as we shall see in chapter 2, Frege takes mathematics both as language *and* as inferential structure entirely seriously, with an autonomous subject matter; and, in order to set out foundations on which the inferential structure rests, Frege's ambition is to supply explicit definitions of the fundamental number terms precisely because their explicit definition will, in turn, supply an articulation of the subject matter of arithmetical language. The explicit definition, then, is held to a standard of material adequacy that, as we shall see, cannot be met by implicit definitions. The material adequacy of the definition in its turn is a component part of the motivation for Frege's thesis that numbers are objects: their objecthood is built into the structure of mathematical thought and the application of that thought to the realm of experience. They are metaphysically indispensable. As will emerge, explicit definition is food for material validity in the process of mathematical inference.

5. Things to come

The following chapters proceed like this. Chapter 1 lays down the groundwork of the argument for metaphysical indispensability, whose core is the argument from constraint given in § 1.1. Elements of both continuity and contrast with the usual IA from empirical science are laid out. I go on to evaluate Hartry Field's much-discussed nominalist programme in the light of the constraint argument, and then assess what Quine has to teach us about both the ontology and the semantics of commitment. I will urge a view on which, while Quine is right about the semantics of commitment-detection, it does not follow that *metaphysical* commitments to abstract objects of the kind indispensable for mathematics are of the very same kind as the unobservables in physics. The upshot of this section is that, while Quine's work does indeed supply the right commitment-detecting strategy, his argument leaves us in a position of indeterminacy with respect to the particular range of objects to which we ought to be committed.

Chapter 2 then embarks on a Fregean semantic project whose ultimate aim is to motivate and restore requirements of determinacy, requirements which are filled by the components of the appropriate theory of meaningful mathematical thought. In this chapter I discuss the notions of explicit and implicit definition in some detail, examining the reasoning that underwrites Frege's explicit definition of number terms and his rejection of attempts at their contextual definition. These considerations ultimately feed an account of why objects are required at all. The first two chapters taken together supply the elements of the positive argument for mathematical realism.

It is nevertheless still the case that an adherent of mathematical fictionalism might endorse at least the semantic aspect of the Fregean argument, and yet reject the realist conclusion in favour of a re-construal of mathematical thought and talk as fictional, or otherwise non-commitment-making, talk. I therefore address various strands of fictionalism in the third chapter. Though the five forms I assess there are surely not exhaustive, they are the prevalent ones in the current literature, and I will offer a view of the common core of any fictionalism, a core which is subject to an overall objection which is, I think, decisive: that any fictionalism will be unable to capture what I'll call mixed-context inferences, because fictionalism demands a rigid kind of epistemic quarantine, a requirement of isolation from assent, for the mathematical premises in mixed inferences.

The closing chapter raises epistemological issues. My aim here is to express and defend a view on which the question of mathematical knowledge should really be split into two quite separate questions: the question of how the genesis of any mathematical beliefs at all is possible on the one hand, and the question of how those beliefs are to be justified on the other. Running these two questions together has made the problem of a suitable mathematical epistemology seem even less tractable than it is; while I do not propose a full-fledged theory of mathematical knowledge, this chapter is meant to carve out the conceptual territory in a rather new way in order to suggest a fruitful avenue of inquiry. I take a few steps down that avenue, especially in the direction of motivating the split between two kinds of epistemic processes. On the matter of the genetic process, which answers the question of how it is possible to have any mathematical beliefs at all, I espouse a minimalist structural epistemology in a sense that will emerge. The question of how to justify the elaborate array of mathematical beliefs that at least some of us do have, though, is a different question and needs to be answered in different terms. Those justifications are the sorts of materially good inferences that constitute the practice of mathematics as a discipline with its own justificatory canons. My aim in this chapter is principally to deflect a Benacerraf-like charge, which holds that one can either have mathematical truth while knowledge remains mysterious, or one can have mathematical knowledge but no satisfactory theory of truth for it. The main virtue I claim for my conceptual split between questions is that it makes that vexing puzzle less difficult to solve.

The overall purpose of this essay, then, is to make a contribution to the philosophy of thought, by examining mathematics as a case study through some of its metaphysical, semantic and epistemic components. Mathematical thought, and the talk that expresses it, is meaningful thought with an autonomous province of inquiry. What follows is an attempt to defend that thesis, and track its consequences.

2

Indisponsability and constraint

The scientist does not like to see his algebra get up, shake itself, and walk away.
-Lawrence Durrell, *Tunc.*

The strange thing about physics is that for the fundamental laws we still need mathematics.
-Richard Feynman, *The Character of Physical Law.*

2.0. Introduction

What rational grounds might there be for thinking that at least some portion of mathematics wears its subject matter on its sleeve, and is about what it appears to be about – numbers, sets and so on – rather than about some other range of objects, physical ones perhaps, or about nothing at all? A forceful argument for the realist face-value claim is one version or other of the indispensability argument, whose main thrust is that we cannot do without true mathematics and the inferences and predictions it licenses, in a variety of contexts.

The indispensability argument (IA) I'll examine gets its bite from more than the contingencies of how our scientific practice has been set up as an essentially mathematical discipline since the 17th century. An opponent of IA might say that it's on the face of it implausible that we could get substantial ontological results out of some formal – mathematical – language which could ultimately be written out of our scientific practice and leave everything else intact, though less economically expressed. But the force of IA lies not in its claim that we can't do without the particular mathematics we happen to have, but rather lies in this fundamental observation:

> (*) In order to make predictions and draw sound inference we must have not just a formal system, but an *interpreted* formal system. Formal systems may be designed in many ways, some of them mutually incompatible; but the interpretation we assign to our formal system of mathematics – that is, the meanings we assign to its non-logical terms – must be *the right ones* in order to explain how our predictions and inferential practices get their purchase on the world outside the mathematics.

There's a great deal packed into this (*) sentence. Much of this chapter and the next are designed to unpack it. This first chapter deals with the commitments of our explanatory theories and the way we detect, and then confirm, those commitments on the basis of our predictive and inferential abilities. This is a first step in defending the IA. The trouble, as we shall see in the section on Quine, is that while we are metaphysically committed to mathematical objects, there is at this stage no straightforward way of assigning a single correct interpretation to the formal systems we deploy. Frege's approach to that problem is the subject of the next chapter.

So, in this chapter I examine the general character of indispensability arguments, and go on to present the versions offered by Quine for the existence of specifically mathematical objects. The main thrust is that, while the indispensability argument (IA) offered by Quine (as well as Putnam, whom I will not discuss directly) is sound with respect to those mathematical objects put to work in empirical science, it is prima facie *limited* to those objects and silent with regard to branches of mathematics that are not empirically applied; moreover, its range only decides objects of commitment up to structural isomorphisms. One strand of argument in this chapter is that the indispensability argument functions as a commitment-detector, and that those commitments are genuine. But at the same time, the Quine-Putnam IA has *too* restricted a scope: while empirical applications are an initial mode of confirming the worth of our ontological commitments, there is no principled reason to razor off only those commitments which find empirical applications, and yet simultaneously to remain agnostic about the unapplied.

Section 1.1. articulates and discusses the IA-form premise by premise; there I also offer a novel argument for metaphysical indispensability; 1.2. is given over to Field's nominalism, especially as expressed in his *Science without Numbers* (Field 1980) and related subsequent works; 1.3. discusses and defends Quine on IA and ontological commitment. Section 1.4. argues that we have rational grounds not just to believe that the results of metaphysical commitment to abstract mathematical objects are *good* – that is, useful – but to believe also that those commitments are true.

2.1. Indispensability arguments, commitment and the role of 'real'

One of the principal arguments for platonism about mathematical objects is really a family of considerations grouped under the title 'indispensability arguments' (IAs). These IAs can be articulated in a number of ways, but the archetypal IA can be boiled down to this argument-form:

1. A theory that successfully explains phenomena in the physical world is most likely a true theory.
2. The commitments of a successful theory are most likely true commitments (at least up to isomorphism).
3. Mathematics has applications that successfully explain phenomena in the physical world.

4. Mathematics is committed to abstract objects like numbers, sets, groups, etc.
5. Therefore, mathematical commitments are commitments to real objects.

Generalizing this argument, we get the schema:

1. A theory T that successfully explains phenomena (physical or otherwise) is most likely true.
2. The commitments of such a T are most likely true commitments.
3. T_N has applications that successfully explain phenomena.
4. T_N is committed to entity-kinds **J**, **K** and **L** (where **J**, **K** and **L** are place-holders)
5. Therefore, **J**s, **K**s and **L**s are real.

Also, I will call a 'mixed context' any expression or inference in T which deploys both mathematical and non-mathematical vocabulary (e.g. 'The forces exerted on the tunnel that collapsed were greater than what its supporting mechanisms could withstand').

The generalized form of IAs (call it GIA) is intended to display the commitments to some range of entities, some of which may be abstract, of which mathematical entities are taken to be a subclass. GIA raises a number of difficult questions. Objectors inclined to nominalism will deny 2, and possibly 4 for any potentially abstract entity-kinds. I devote section 1.2 below to an argument against Field's nominalism in general (understood as the wholesale rejection of *any* abstracta). Others will object to the transition from 4 to 5, and perhaps also deny 1, but in the closing section of this chapter I will argue that we have rational grounds to make that transition in the case of metaphysical commitment to mathematical objects. First, though, a word on the specifically mathematical version of IAs.

Aside from the isomorphism-related problems that face the general version GIA, which I'll discuss in 1.3. below, IAs for mathematical entities seem to have some specific problems of their own, not the least of which is the nature of the indispensability in question. The arguments I offer in what follows present a different-than-usual version of IA, and a reorientation of its focus. This version is cast *a priori*, and argues for *metaphysical* rather than empirical or theoretical indispensability. To illustrate the difference, consider this: in the course of an argument on behalf of mathematical objects' indispensability in natural science, Alan Baker (2005) raises the very interesting example of North American 'periodical' cicadas and the resources that need to be brought to bear on the explanation of their life-cycle. Briefly, the case is this: in each of three species of cicada of the *Magicicada* genus, 'the nymphal stage remains in the soil for a lengthy period, then the adult cicada emerges after either 13 years or 17 years depending on the geographical area.' (Baker 2005: 229) That the life-cycle should fall into a prime-numbered-year sequence is curious, and wants explaining. Baker canvasses two rather different biological explanations of why this is an evolutionarily advantageous sequence, both of which, despite their differences, invoke the number-theoretic fact that prime periods minimize intersection compared to non-prime periods. (Exactly what they minimize intersection *with* is open to interpretation, but Baker's discussion of the phenomenon is uniform across both of the interpretive versions he cites: see the references in his article.) In any event, Baker writes:

> [T]here are genuine mathematical explanations of physical phenomena, and the explanation of the prime cycle lengths of periodical cicadas using number theory is one example of such. If this is right, then applying inference to the best

> explanation in the cicada example yields the conclusion that numbers exist. (Baker 2005: 236)

Nevertheless, he continues,

> Whatever cases of putative mathematical explanation the platonist might come up with, there will always be some leeway for nominalist objections since the role of mathematical posits is unlikely ever to *exactly* match the role of concrete unobservables, such as electrons. (*ibid.*)

I think it's helpful to cast Baker's argument like this: since explanation is achieved by invoking a property that is instantiated by nothing at all unless it's instantiated by an abstract object, the relevant abstract object is indispensable to the explanation. A further feature of the empirical IA is that it sees its criterion of success as the ability to locate cases in which the explanation of contingent physical facts ineliminably involves reference to abstract objects, whose existence is necessary if they exist at all. This opens up an opportunity for the nominalist to drive a chisel into the interstitial explanatory space between concrete and abstract particulars, and then hammer off the *abstracta*. Now, my claim will be that while examples like Baker's cicadas are compelling, we can further shore up the mathematical realists' case by reorienting the focus of the indispensability argument from an *a posteriori* inference to the best explanation—in which abstract objects emerge as posits on an ontological and epistemic par with physical unobservables—to an *a priori* inference that will allow us to derive the explanation of a properly physical *impossibility*: rather than looking to what mathematical objects rule in, I want to have a look at what they rule out. Mathematical objects don't need to match the role of concrete unobservables: they do a different kind of duty, but explanatory duty nonetheless. In what follows I will also do more to flesh out what's meant by 'real' than is usually met with in this sort of argument.

Consider, next, this query put to the defender of IA: what about those branches of higher mathematics for which there is no physical application? The reply to this will hinge on whether one thinks that the truth of commitments to mathematical entities has to be defended on a case-by-case basis, or whether one thinks instead that if IA is true for any class of mathematical entities, it might as well be true for the whole lot. To put it another way, one could argue in either of two ways:

> *a*) The IA is sound; a subset of mathematical entities exists; so that's good evidence for the existence of the whole bunch.
>
> *b*) The IA is sound; a subset of mathematical entities exists; but this is only evidence that we should retain the ones we demonstrably need for the explanatory purposes of our science and remain agnostic about the rest.

Option *b* treats putative mathematical entities *exactly* on a par with putative physical, though unobservable, entities like quarks: 'only commit to as many kinds as you absolutely cannot do without' is the motto. This is the option Quine preferred. The problem with *b* is that it is not clear what status it gives to a lot of actual mathematical practice: there is substantial continuity of assumptions, methods, patterns of inference and canons of justification between the mathematics that gets applied and the body of it that doesn't (or doesn't yet), and thus there is no principled reason to razor off one subset of commitments when the quantificational requirements of commitment extend through the whole body. Call this an argument from methodological continuity. The general form of the argument is this: if it is rational to believe some thesis T (along with its commitments) on the grounds of some method M, and M also establishes an extension T+ of T, then it is equally rational to believe T+ on those same grounds M. So version *a* is preferable. Version *a* claims that, while there is an analogy between IA and the general form GIA, the analogy is not exact. Mathematical entities are not exactly like physical postulates: the difference is that while *b* says that in IA and G the justification procedure is the same as the statement of ontological requirements, option *a* points out that they are *different*. Justification, on option *a*, is up to the mathematicians and the constraints they have found effective, constraints of proof and plausibility and so on; the empirical scientist is subject to different constraints of inference and theorizing, and it's up to the scientists to settle on them. What IAs show is not that science and mathematics justify their claims in the same way: they just show schematically how theories make commitments. Note that premise 1 in both IA and GIA invokes (likely) *true* theories. There is no claim that the sources of evidence for the truth of mathematical propositions on the one hand, and of empirical theories on the other, need to be the same. Indeed, they're not. Mathematical and scientific justifications are fundamentally different, for this reason: the resources invoked in methods of justification for the latter are essentially empirical and inductive, while the resources (the resources of M from the argument above) are essentially *a priori* and deductive. Criteria of mathematical and scientific ontological *commitment*, however, are not different. We need to be very sensitive to exactly what IAs are claiming, and the extent to which they are similar to GIA. Let's take the argument premise by premise.

IA (1): Any theory that successfully explains phenomena in the physical world is most likely a true theory.

What needs exploring here is the relation of explanatory success to truth. 'Success' and 'truth' are not necessarily coextensive. Suppose I'm a Roman builder of aqueducts. My apprenticeship has taught me how to build good aqueducts; they're so good they'll stay standing for two thousand years. Suppose further that I believe that when I build aqueducts according to the principles my teacher instilled in me, they stay standing because the engineering principles I was taught (e.g. to arrange the bricks in certain patterns) are maintained in place by the gods. Other, less durable, arrangements of bricks collapse because they are displeasing to the gods. Operating with this false theory I can successfully build mile after mile of durable aqueduct. The [gods + bricks = good aqueducts] theory is extremely successful, but false.

The physical phenomenon that wants explaining is that my structures stand. The elements that do the explanatory work are arrangements of bricks and the intervention of the gods. There's no reason for me to be dissatisfied with my explanatory theory (as long as I remain a theist) since it's borne out in practice time after time. The more I build, the more it's confirmed. Indeed, starting with this theory, I gradually take my engineering success to be evidence for the existence of the gods whose postulation I take as necessary in the *explanation* of that success. As George Eliot put it in *Middlemarch*, "incantations will kill a herd of sheep, if administered with a sufficient dose of arsenic."

An objector to IA (1) might then draw on this observation: you can't get evidence for ontological commitments out of the elements of a merely successful theory because any number of weird and outrageous commitments is compatible with success. We measure success by a capacity to predict what will happen next ('if the theory is right, the ball will drop at *this* rate…'), but success underdetermines the theoretical structure that purports to explain it. Consider also the well-known story of the professor and the flea. A professor trained a flea to jump at the sound of a bell. The professor ripped off one of the flea's legs, rang the bell, and the flea jumped. He tore off another leg, rang the bell, and the flea jumped again. The professor carried on this way until the flea had no legs left. When he rang the bell, the flea just stayed put. Thus it was proved: fleas hear through their legs.

So IA (1) seems shaky, taken on its own. Predictive success is the best sign of the truth of a theory, but the theory is always underdetermined because we can cook up scenarios in which the same phenomena would follow from radically different ontological commitments. And it's no use saying that in that case, the radically different commitments are equivalent if they play the explanatory role they're supposed to – to claim for instance that whatever saves the phenomena is explanatorily adequate – because there is a clear intensional, semantic difference between commitment to gods and witches on the one hand and to the objects required by counterfactually applicable laws of nature on the other. Moreover, the same evidence can support very different conclusions. This feeds into premise 2, with its cautionary parenthesis to flag the underdetermination problem:

IA (2): The commitments of a successful theory are most likely true commitments (at least up to isomorphism).

Hartry Field (1980) denies this premise in the case of the indispensability argument for mathematical entities, unless we take the commitments in question to be ones which are nominalistically acceptable; that is to say, commitment is restricted to space-time points and regions composed of these, whose causal interactions are what does the real explanatory work. His point is summarized in the phrase "mathematics needn't be true to be good". The idea is that even if we have a language whose vocabulary and patterns of valid inference allow us to draw successful conclusions about (or provide successful explanations of) observable physical phenomena, we are still not warranted in making the step from the apparent commitments of the language to the existence of the correspondents of those commitments.

Field's argument has intuitive force, but, as I shall argue, only up to the point at which its strands need to be teased apart. If I say – and believe – that a leprechaun bit me in the ankle, I do indeed make an existential commitment to a nonexistent object that plays a role, possibly even an indispensable role, in my theory of the world. But it's a role that snakes or hamsters or other ankle-high biters could also play, and play better. Commitment to snakes *versus* commitment to leprechauns is not to be decided by investigating the respective structural or logical roles of the terms 'leprechaun' or 'snake' in our theories, because they play the very same such structural role. The terms do not, however, make the same contribution to the truth-conditions of the sentences that involve commitment to them: where it is the case that 'a leprechaun bit me in the ankle' and 'a snake bit me in the ankle' differ only in extravagance of ontological commitment, the snake as value of the variable is to be preferred over the leprechaun on grounds that are thoroughly non-logical or non-structural. Truth is not at play in Field's objection to IA to the extent that truth is a formal relation between utterances (or propositions) and conditions of satisfaction for those utterances (or propositions), because many potential nonequivalent candidates for values of the existentially quantified variables are available, and we need independent reasons to choose one kind of value over another. Rather, truth lies in the explanatory adequacy of our choices.

Now, keep in mind the slogan that mathematics needn't be true to be good. I might have a good theory of the causes of my sensation that includes leprechauns, but I needn't be committed to leprechauns (i.e. they don't have to be real, 'leprechaun' need not be a referring expression, there might not *really be* leprechauns, etc.) for my theory of my sensation to be a good theory. Field's point is that a richly descriptive language *can* get us to the matters of empirical fact in a hurry, but if this expedience happens at the cost of extravagant ontological commitment to things as outlandish as numbers or witches or leprechauns, we should pause and look more carefully at the relationship between the descriptive (explanatory) language and ontology. According to Field's (1980) nominalistic reconstruction of some aspects of scientific theory (notably Newtonian gravitation), ontological commitment should not be taken lightly: if we can build good explanatory science with less ontological profligacy, at the cost of spending more time reconstructing the language, then we'll get a better picture of what the shortcut-less language really commits us to, and leprechauns and numbers will be written out of the story. The next section (§1.2.) is specifically devoted to a critique of Field's nominalism, but one further preliminary point. We must distinguish between conditions of justification and conditions of explanatory adequacy. Field distinguishes between conditions under which certain empirical theoretical claims are true and conditions under which those same theoretical claims are explanatorily adequate. My claim will be that, while the explanatory adequacy of the mathematical language in which we express laws of nature and the validity of inferences we make from them is prima facie reason to accept the existence of (say) real numbers, the primary purpose of indispensability arguments is to reveal that to which we are committed if we are to explain the validity of mathematical inferences themselves, in contexts both pure and mixed. That is, accounting for valid mathematical inference and the explanatory purchase it affords us requires abstract objects as the semantic values of the non-logical components of inference in *mathematical* cases, of which there are instances that allow the explanation of constraints (to be discussed) as a further advantage.

IA (3): Mathematics has applications that successfully explain phenomena in the physical world.

This is just a bridging premise and is uncontroversially true.

IA (4): Mathematics is committed to abstract objects like numbers, sets, groups etc.

One way of denying this premise is the fictionalist strategy that will be discussed in chapter 3. Another way of denying it is to place restrictions on its scope and make it subject to some epistemological or special semantic constraints (as in intuitionism and constructivism) that I discuss in chapter 4.

IA (5): Mathematical commitments are commitments to real objects.

What grounds might there be to motivate this claim? The concept 'real' is notoriously thorny and difficult to characterize in a way which is sufficiently general, clear and informative. The contrast between realists and antirealists about the commitments of a range of discourse is variously cast as *semantic* (with realists claiming that propositions have determinate truth values quite independently of any human mode of coming to discover those values for each proposition, and their opponents denying this) and as *ontological* (with realists claiming that there *really* are entities of whatever kind, and the antirealists denying this). But simply to *pick* semantic and ontological realism over the corresponding antirealisms at this stage would be to beg the question, since what is at issue is to discover independent motivation for picking one side over the other.

A more helpful, and non-question-begging, way of getting a foot in the door is to re-cast the issue in the manner of Reynolds (2006), who directs his attention to the semantic function of 'real', which is, he claims, to mark a transition from the level of representation to the level of the world, rather than as a transition from one level of representation to another. In other words, when we say or think that some object *a* is real, the purported subject-matter of the thought or the talk is not another level of representation, but is rather the world itself. The argument comes largely from Austin (1962), who observes first that 'a real *a*' is "substantive-hungry", i.e. that it wants a noun to modify, and second, that 'a real *a*' should be understood by way of *contrast* with such locutions as 'a fake *a*', 'an imitation *a*', 'a picture/representation of *a*' and so on.

Thus, to say of a duck that it is real is to say that it is not a decoy or a representation of a duck; to call real cream 'real' is to say that it is not some other *ersatz* stand-in for whitening one's coffee. Reynolds puts it as follows:

> [T]he word 'real' usually indicates that the speaker… is not using the noun phrase to refer to something that, in the telling of some story [i.e. the reporting of the contents of a representation], merely represents the sort of thing that the noun phrase usually refers to. Thus, a diamond is said to be real when it is desired to indicate that it is not being called a diamond merely because it represents a diamond in some story… The word 'real' in "it's a real

> diamond" indicates that 'diamond' is not being used that way. (Reynolds 2006: 475)

And furthermore:

> '[R]eal' does not express a quality of things, … we don't need, and probably couldn't have, one test for whether any given thing is real. Ducks, cream, and progress are very different things obviously, and we have different ways of finding out whether something is a duck or cream or progress. But that is all we do when we find out whether something is a real duck, or real cream, or real progress. On my view, the contrast is not between lacking and having a mysteriously discernible and/or surprisingly variable quality of reality, but rather between only representing a duck, cream or progress in the telling of a story about those things, and being a duck, cream, or progress. (Reynolds 2006: 478-479)

The Quinean principle of semantic ascent has shaped our ways of talking about the real. The principle is that we should talk about how we talk, not about things. Semantic *ascent* is just that: to ascend by a logical level is to shift the subject-matter of our theorizing from what we once (perhaps naively) supposed was the world, so that the new subject-matter becomes our *talk* about that world. Indeterminacy, underdetermination, inscrutability and isomorphisms have shown us some of the complexity in the channels connecting us to language-independent reality, so we ought to be thinking about our *talk* first and foremost. The world we try to get some purchase on is available through the commitments we make in our quantified talk, and we pick our commitments through our best empirical strategies; but our commitments can be revised. Sometimes the revision is piecemeal, occasionally it's wholesale. The test is always the empirical, *explanatory*, traction our language allows us. The real is what most stubbornly resists elimination from our scientific talk, or from our best explanatory talk in general. Quine was distressed to discover that he was a platonist, at least about the lowest grade of set theory. But he made his peace with Plato, because you can't do physics without the real numbers and you can't account for the language of physics without quantifying over the real line. Some day, perhaps, the real line will be eliminated, like the Greek pantheon, on the grounds that we don't need it in our language any more. But for now at least, it is indispensable, as much a real part of life as bricks. Semantic ascent is a strategy for spotlighting the commitments we can't do without.

I say that this doctrine has framed much of the way in which we talk about reality, because it pinpoints the *locus* of ontological disagreements. The *locus* at issue is the extent of the resources we need to make our best current sets of sentences come out true. Do we need to have our quantificational commitments satisfied by 'real' entities, or can we appeal to other – and hopefully less distressing – resources, whatever they may be?

Austin, like Quine, was also thinking carefully about 'real' and its cognates. Like Quine, but for very different reasons, Austin adopted the method of talking first about talk, and letting objects fall into their places as needed. In virtue of what we discover through semantic ascent as a methodological strategy, we are locked into talk about what we at least *take to be* real. Sometimes those assumptions will not be borne out. As I've said in previous sections, by Quine's lights, at least, it is predictive success and explanatory work that will sort the real from the spurious or the dispensable. Reality is *not* just the history of what we say, as though Zeus were once real and is now no more, or as though electrons popped into existence from the void some time around 1880.

Rather, reality is approached obliquely through ways of talking, and those ways are better or worse as their commitments resist, or bend to, elimination from the best tale we have to tell about the world.

Returning now to IA (5), that mathematical commitments are commitments to real objects, what can we say to give the claim some legs? There would have to be an argument to persuade us that mathematical thought and talk take the world (albeit abstract parts of it) as their subject matter, and are not merely reports on the contents of another level of representation whose subject-matter is something else (as both the nominalist and fictionalist would like to urge). Now, I think there *is* such an argument, and I will call it the argument from constraint. While the usual interpretation of IA gets its mileage from its explanation of why mathematics allows us to *accomplish* certain feats, like building bridges and landing probes on Mars, the argument from constraint is an interpretation of IA that proceeds by explaining why we *cannot* accomplish certain things.

2.2. The argument from constraint

It goes like this. One of three famous ancient Greek construction problems in geometry is called the problem of squaring the circle (the two others are trisecting the angle and doubling the cube). The problem is this:

> **(SC)** Using only an unmarked straight-edge and compasses, and given a circle of area A, construct a square having that same area A.

Greek geometers struggled with this problem for a long time, as did many mathematicians until the 19^{th} century (mathematicians were similarly exercised by the other problems too). It turns out that the construction demanded in (SC) is, in fact, *impossible*, and we can prove it. (The other constructions are impossible too, for related reasons.)

Suppose a circle of radius *r*. The area of a circle A_C is given by the formula $A_C = \pi r^2$. Now, let $r = 1$, to give us an arbitrary unit length. The area A_S of a square of side *a* is, of course, a^2. This means that to construct $A_S = A_C$ we would need to construct a square with side $a = \sqrt{\pi}$.

This is what's not possible. π is a transcendental number: this means that it does *not* satisfy a polynomial equation (i.e. one of the form $a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 = 0$). Now, polynomial equations operate exclusively with the functions of addition, subtraction, multiplication and exponentiation: these are the operations that can be geometrically constructed, that is, carried out in principle with the tools set in (SC). But π does not satisfy any procedure that involves just these operations: it is transcendental (the proof was given by Lindemann in 1882). The upshot is that there is no constructive geometric 'performance' – so to speak – by which we could accomplish the task set in (SC). So a result in pure mathematics, the transcendence of π, gives us an *explanation* of why something is impossible, an explanation of a constraint on the possibilities that are open to us. It is an explanation given exclusively *a priori* which invokes a particular mathematical property, transcendence, belonging to a mathematical object, π.

At this stage the following objection might be raised.[1] If I host a party, and I run out of beer at a quarter past midnight, my not being able to serve Emily a beer at twenty past midnight is explained by my lack of beer, or by the dearth of beer in the fridge, or something along those lines. But this explanation is not an explanation that commits me – or rather, commits the theorist who is interested in examining my drink-serving

[1] I owe the objection to Joshua Mozersky.

constraints at the party – to such things as dearths or lacks, as though they were objects. The proper explanation of my inability to serve enough drinks will make no mention of such an *entity* as a dearth of beer. Why suppose, then, that the ontology of explanation for the impossibility of (SC) will be any different from the ontology of the explanation of why Emily can't have a beer?

The reply is this. There are two fundamentally different kinds of constraint invoked in these two scenarios, and two different corresponding kinds of explanation. My inability to serve enough drinks is a contingent constraint, and that constraint is explained by causal features of the situation. The relevant causal explanation will invoke an ontology of contingent relations between events. The impossibility of carrying out the (SC) construction, on the other hand, is not an *event* since it does not unfold in space-time, and is not open to causal explanation since the constraint expresses a *metaphysical* impossibility: there is no possible world in which the (SC) construction can be carried out, whereas there is obviously a possible world in which I have more beer than I have at my party in this one. We might call the mathematical explanation of the constraint an essential explanation: one that is necessary, and thus non-causal. Part of the argument from constraint's bite is that it points to a necessary property, instantiated in any metaphysically possible world whatsoever. This particular necessity points in turn not just to theoretical indispensability for natural science, but to metaphysical indispensability in understanding a portion of the structure of possible worlds, this one among them. The argument from constraint, then, is an argument from the requirements of essential explanation to metaphysical indispensability. In the next section I will address Hartry Field's nominalism, but a preliminary word is in order here: the argument from constraint is one whose commitments *cannot* be nominalized, because to nominalize a range of commitments is – in Field's way of thinking, at least – to bring that range into the province of the causal. But no such causal integration of the requirements of the explanation of (SC)'s impossibility is available, because the relata of the explanation are not contingent events.

Now, recalling the conclusion IA (5) that mathematical commitments are commitments to real objects, we can, by way of this argument from constraint, also give some content to what is meant by 'real' in this context. (Of course by 'real' here I don't mean 'part of the set **R** of real numbers'; I'm using 'real' in the more colloquial sense.) To say that at least one mathematical object is real is to say that *that very object* is the subject matter of, for instance, Lindemann's proof of its transcendence. The mathematics is not *about* a further layer of representation standing between the mathematical language and the world. Representations don't make certain actions impossible, any more than the Criminal Code makes murder a physical impossibility. I conclude that the argument from constraint supplies rational grounds for believing IA (5).

To illustrate this thesis about what's meant by 'real', consider the following nice passage from Dennett:

> If we devised a program that simulated a cow on a digital computer, our simulation, being a mere representation of a cow, would not, if "milked", produce milk, but at best a representation of milk. You can't drink that, no matter how good a representation it is, and no matter how thirsty you are.
>
> But now suppose we made a computer simulation of a mathematician, and suppose it worked well. Would we complain that what we had hoped for

> was proofs, but alas, all we got was mere *representations* of proofs? But representations of proofs *are* proofs, aren't they? (Dennett 1981: 94 – 95).

Dennett writes this in a completely different context – in fact this passage occurs as part of a discussion of the Chinese Room argument against Turing machines as authentic thinkers – but it points very nicely to what I mean in saying that the semantic function of 'real' is, in the contexts that interest me here, one of marking a difference between levels of **subject-matter**. Here are a few examples to illustrate:

1.2.1. That's a real Rembrandt, but *that* one's fake.

1.2.2. That's not real milk [said of the digital cow's output]

1.2.3. That's not a real place, you know [said pointing to the background in Leonardo's *Annunciation*]

1.2.4. *She's* a real lawyer, but he's a fraud

These are contrastive uses of 'real': in each case we are marking a distinction between actually satisfying a property, and just *seeming* to satisfy a property: being a lawyer, say, versus playing one on TV, or getting up and speaking in court dressed in the gown without actually being a member of the Bar. This particular use of 'real' is marking different levels of talk, ones that point to a level of seeming and another of being, so to speak. A picture might seem to most to be a Rembrandt, until Kenneth Clark marches in and pronounces it a fake: it's still a real picture, but it's a fake Rembrandt. The real picture represents a Rembrandt, but isn't one.

Now, my claim is that:

1.2.5. Transcendental numbers are real

deploys the word 'real' in just this sense: transcendental numbers are the subject-matter of the mathematical claims that are made about them – say, that they do not satisfy algebraic equations – and the subject-matter is not some additional level of representation in which we have a case of *seeming* to satisfy a property without *actually* satisfying it. It is, moreover, the case that π's instantiating the property claimed for it explains a constraint on the procedures of construction that are open to us. This is the explanatory work it does. For illustration, think of the distinctions between a human heart, a plastic model of a heart in the anatomy classroom, and a functioning artificial heart. The first is a real heart; the second is not. The third is not a human heart, but it's a *real* heart in the way it works: it pumps blood and allows its possessor to carry on living. The plastic model *represents* a heart: the working artificial one doesn't represent; it *works*.

Now, if we attend to the doctrine of semantic ascent and pay attention to our way of talking, we are led to a commitment to the things that play an explanatory role in our best theories in order to make them come out true. *De*scending semantically, from language back to its traction on the world, we are then led to conclude that this argument from constraint is an argument for mathematical realism in the requisite sense of 'real'. It's not the representation of a property that's stopping you carrying out the (SC) construction; it's the very property, actually instantiated, itself. And that's what we expect of what's real

.

2.3. Field's nominalism

I turn now to the following question: are there compelling grounds to endorse nominalistic strictures even in the light of the foregoing arguments, and thus deny IA (4) and (5) in favour of a nominalistically acceptable reconstruction of apparent commitments through strategies of paraphrase? Hartry Field writes that "ultimately the only serious argument [i.e. the IA] for the view that mathematics is true as well as conservative turns on the premise that nominalistic bodies of assertions are not always available" in the formulation of our best explanatory theories (Field 1989: 59). 'Conservative' in the sense deployed here means this: suppose a mathematical theory M and a body N of assertions, nominalistic ones, about the physical world. Take any nominalistic assertion A about the physical world. M, the mathematics, is conservative over N iff A is only entailed by N + M if A is already entailed by N alone. If N + M entail A, then N entails A. In essence, M is conservative over N if it doesn't entail anything not already entailed by N by itself. The interest of conservative mathematics lies in its ability to smooth out the inference process, but it doesn't enjoy any truth over and above the truths delivered by the nominalistic assertions alone. Conservativeness is a reducibility thesis. Field's celebrated slogan that mathematics need not be true to be good is given its content by this very thesis: conservative mathematics is good for inference, but is not true beyond its nominalistic underpinning; or, to express the point in rather different terms, mathematical subject matter is nothing that is not nominalistically acceptable. Mathematical truth does not overstep the bounds of some nominalistically expressed equivalent. Hence Field's claim that nominalism hinges on the availability in principle of reductions of mathematical assertions to nominalistically equivalent paraphrases (this is a thesis in-principle, because it might well be the case that such paraphrases, though logically possible, are too long ever to be expressed, or too unwieldy to be useful). If all the mathematics does is facilitate essentially nominalistic inference, then we have no reason to suppose that mathematics is a body of non-nominalistic truths and IA has no ontological purchase. If, on the other hand, mathematics is genuinely theoretically indispensable (irreducible, i.e. at least partially not open to paraphrase), then IA goes through, by Field's own admission. The project of

Field's *Science without Numbers* (1980) is to defend the former alternative, and offer just such a nominalistic reduction of a portion of classical Newtonian gravitational field theory in flat space-time to show its conservativeness over nominalistic assertions. In general, the question is one of the possibility of reduction – and hence elimination – through nominalizing paraphrase[2], since, if successful, IA is stonewalled. What, then, of the theoretical dispensability of mathematics? Field's words again: "Given any application of mathematics, then, it is natural to ask whether the utility of mathematics in that application is due to its conservativeness or its truth" (Field 1989: 62).

Now, the ontology of Field's thesis goes like this. The subject matter of the mathematics deployed in physics is, at rock-bottom, an array of space-time points and the regions they compose (this gives us the $\mathbf{R}^4$ structure of space-time, effectively a nominalist's real line in four dimensions of measurable properties), as well as the causal properties of those points and regions. Field's physics thus *quantifies over* not just the array of points and regions – each of which he is happy to call a physical object – but also over the (causal) properties of the array. He therefore has to help himself to the resources of second-order logic in order to account for the inferences made with the predicates that, on this picture, are conceived as genuine referring expressions.

A number of people, including Malament (1982) and Resnik (1985), have justifiably complained that this construal isn't recognizable as nominalism any more. Indeed, it lets abstract objects right back into the picture, except that rather than being quantified over in the first-order language, they reappear in the quantificational requirements of the second-order language. They may have changed their address, but

[2] Note that nominalistic reductive paraphrase is not intended to be a *meaning*-preserving strategy. A nominalistic equivalent will not be a semantic equivalent, since on Field's view the expressions in M are largely false, while the expressions of N are true.

that's a far cry from elimination. If all that Field has done, metaphysically speaking, is file a change of address form for the *abstracta*, then it is not clear that this is a measure of ontological economy. That being said, Field could no doubt counter this objection ('you've got abstract objects too') by stressing that the quantificational requirements of the second-order language are confined to those predicates that stand for strictly *causal* properties of proper subsets of points of the array and, being so confined, they are indeed nominalistically acceptable. This seems to be the thrust of Field's remark that "to take the notion of a field seriously, space-time points or regions are full-fledged causal agents" (Field 1980: 114 fn.23). But Field furthermore writes that "regions don't need to be connected, or measurable, or anything like that: very 'unnatural' collections of points count as regions" (Field 1980: 114 fn.26). This, however, sounds exactly like a set-forming choice function over the sets of points in the array, and to count these as regions is then to reintroduce the very quantification over mathematical objects, sets in this instance, that was supposed to be reducible through nominalistic paraphrase. The resources of mathematics are then squarely resettled in the formulation of physical theory. If one has no sets, then one has no probability distributions, for instance, and thus no kinetic theory of gases or quantum mechanics.

Can Field now try the rejoinder that, while indispensable, the set theory is nevertheless not *true*? Field endorses a version of fictionalism about set theory (and mathematics in general), and I reserve a discussion of the specifics of fictionalism for fuller treatment in chapter 3, since it is an active programme of current research and deserves detailed assessment in its own right. Here I will confine my attention to the relationship between truth and conservativeness in Field's sense of the latter term. Is it possible to maintain that set theory, while required for empirical science, is itself conservative over a nominalistically acceptable range of commitments?

This strikes me as a hypothesis about what is technically feasible, which has yet to be borne out. It is one thing to nominalize a portion of Newtonian field theory in flat space-time. The possibility of engineering the same reduction for the universe of Einsteinian curvature is quite another matter, and doing the same for quantum-level reality is yet another. Is there any reason-in-principle for being optimistic about its prospects? An antecedent commitment to the nominalistic outlook is surely not sufficient warrant for thinking so, and the presence of set theory in the language Field is willing to countenance dims the prospect of uniform nominalistic reduction across the whole of well-entrenched empirical science.

Admittedly the foregoing paragraph does not furnish a direct argument. But the direct argument is this: to say that the subject-matter of the mathematics deployed in a portion of physics is really an array of space-time points and regions mereologically assembled out of such points (essentially by a choice function) is to say neither more nor less than what set theory says already, *except* that Field adds the clause that such points and their sums are nominalistically acceptable because they are capable of entering into causal relations. The grounds for the extra clause, however, stretch the concept of causation to its snapping point. Recall that regions (which include mereological sums of points) need not "be connected, or measurable, or anything like that"; but this over-generates what is acceptable as a causal region. Nothing in this view prevents us from selecting outrageous sums as causal agents in Field's sense, like the sum of the magnetic North Pole, the eleventh sentence of *Pale Fire* and the stubble I shaved off my face last week. The point is this: if Field-conservative mathematics requires these arbitrary sums *as* *potential causal agents* in order to provide the requisite nominalistic underpinning, then we should reject the underpinning on grounds of arbitrary over-generation and therefore reject the hypothesis of conservativeness that led to it. Nominalistic paraphrase is attractive if it can supply us with all of our mathematical requirements and simultaneously pare down the range of our theoretical commitments. But Field's space-time points and their summation into arbitrary, though causal, regions leave us with an intolerable inflation of the ontology of the causal.

2.4. Quine, commitment and structural isomorphism

Quine writes:

> Discourse in general, mathematical and otherwise, involves continual reference to abstract entities of this sort – classes or properties. One may prefer to regard abstractions as fictions or manners of speaking; one may hope to find a method whereby all ostensible reference to abstract entities can be explained as mere shorthand for a more basic idiom involving reference only to concrete objects (in some sense or other). Such a nominalistic program presents extreme difficulty, if much of standard mathematics and natural science is to be really analyzed and reduced rather than merely repudiated; however, it is not known to be impossible. (Quine 1947: 121)

Indispensability is equivalent to ineliminability from an explanatory role in theorizing. The passage cited above indicates that eliminability, or in the case of mathematical objects reduction of a theory committed to them to a theory that has no such commitments, presents technical difficulties that Quine, at least, thought were likely insuperable in practice, though Field's (1980) project is precisely to take up the challenge of such a grand feat of elimination. In this section I explore Quine's version of the indispensability of mathematical objects and some of its consequences.

Quine again:

> To show that a theory assumes a given object, or objects of a given class, we have to show that the theory would be false if that object did not exist, or if that class were empty; hence that the theory requires that object, or members of that class, in order to be true. (Quine 1969: 2)

There is a constraint on theories T in the argument-schema from section 1.1 above that governs their allowability as features of the premises in the inference: if any T is to feature in the argument, it cannot be reducible to any T* that has fewer or more economical commitments. Since the inference is from commitments and explanatory success to existence, the commitments that feature in the premises must be the fewest possible needed to ground explanatory success. There is thus an irreducibility constraint on T, and it falls to the defender of indispensability arguments to justify the irreducibility of T's commitments.

The puzzle is therefore to discover which existentially-bound variables are ineliminable from a successful theory, which commitments are such that their elimination would render the theory unsuccessful. Such is the Quinean decision-procedure for working out the ontological commitments of a theory, but it's a procedure that offers no solution to the leprechaun problem. It is a structural procedure that has no particular content in advance of a particular body of talk that is then put through the mechanism of what Austin called the 'survival test' (Austin 1956: 185): even if we suppose that the retrograde confirmation of the 'referentiality' (so to speak) of the genuine singular referring expressions is the right way to cash out the explanatory worth of ontological commitments, still 'leprechaun' and 'hamster' play the same role of ankle-high biter in the structure of the theory and they are extensionally non-equivalent, so arguments independent of considerations about the structure of theories need to be adduced to decide the issue between structurally equivalent theories that differ not only in intension – 'leprechaun' and 'hamster' mean different things – but in extension as well. Structural equivalence obtains when all the place-holders of two or more theories hold the same places; extensional equivalence obtains when, for any two or more values used to saturate the place-holder, those values have all the same properties[3]; intensional equivalence obtains when two or more terms used as names for the values that saturate the place-holders have the same meaning.

Returning to indispensability, and equipped with these distinctions, we're in a position from which we can see:

> [1.3.1.] an ontological point: features *internal* to a theory are not able to adjudicate between rival ontological commitments; and
>
> [1.3.2.] an epistemic point: explanatory success is not the external feature we can appeal to in order to tell the difference either.

On point [1.3.1.], the retrograde confirmation of the role of whatever value for the place-holders we choose will tell us nothing about the values themselves – that is, the things in the world we want to be talking about – but only about their *role* and whether it is a role worth playing. Nothing in the internal structure of the rival theories will allow us to tell the leprechauns from the hamsters, because despite being both extensionally and intensionally non-equivalent, they play the same *structural* role in saving the phenomena. Short of allowing a disjunctive commitment to an either-leprechaun-or-hamster, Quine

[3] Like Quine, I'm assuming objectual quantification here.

leaves us with the demand that we either find a criterion that is external to the theory and which will split the leprechauns from the hamsters, or that we bite the bullet of point [1.3.2], admit that structural equivalence is the best we can hope for, and modify IA(2) accordingly. In that case, IA(2) would read: 'the *structurally equivalent* commitments of a successful theory are most likely the right range of true commitments' and we would give up on pinning the commitments down any further, which is precisely the problem of isomorphic theories with intensionally different commitments.

Both avenues raise their own problems. The first, finding a criterion of ontological commitment from an Archimedean point that lies outside the confines of the theory, is, for Quine, impossible: this impossibility derives from his (and Duhem's) familiar arguments for semantic and scientific holism, which I will not canvas here. The second option, in which structural equivalence is the best we can hope for from the epistemic point of view, is more or less the one that Benacerraf (1965) leaves us with – at least for the specifically mathematical case. Indeed Quine's arguments (or the arguments I have offered in Quine's name) lead us to a structural realist conclusion in the philosophy of mathematics, for (*a*) existential commitment to mathematical objects is indispensable for the explanatory purposes of our best natural science, and (*b*) there is no way of telling the difference between structurally equivalent values used to fill the extensional place-holders. If Quine is right, these joint theses tell us that there are mathematical objects to which we are metaphysically and logically committed, *and* in the face of rival theories of them we can't know which ones they are. This is what Quine's version of the indispensability argument amounts to: platonism with an indeterminist twist.

I don't think that this is the right platonism, nor is it the right kind of indispensability argument. Quine's (and to an extent, Putnam's (1971)) arguments for indispensability are much too strongly focused on indispensability for natural science, and there are some fundamental problems with arguing for mathematical platonism in this way. I will explore what I think are better arguments in the next chapter, on Fregean platonism. But first, a preliminary word on where I think Quine goes wrong.

Natural science, despite its vicissitudes and well-known disorder in both internal coherence and methods of discovery, is the best touchstone for *empirical* truths that we have and are likely to have. This is so much the case that a great many philosophers have seen the accessibility of a body of empirical truths as the paradigm case of truth, and have argued that ultimately all truths must have an empirical ground. As a consequence, if there are mathematical truths (and there are), they must be explicable on the basis of some empirical bedrock. Since mathematics is such a valuable, indeed indispensable, language for *expressing* empirical truths, the temptation is strong to look for mathematical truth in the ways in which mathematical language might be referentially tethered to the deliverances of sensory experience. But mathematics isn't just language; it is language combined with rules of inference, language that comes with a built-in way of getting from the known to the hitherto unknown. Feynman (himself no friend of philosophy, even as he philosophizes) hit on something important when he wrote:

> [M]athematics is not just another language. Mathematics is a language plus reasoning. It is like a language plus logic. Mathematics is a tool for reasoning… [I]f you do not appreciate the mathematics, you cannot see, among the great variety of facts, the logic that permits you to go from one to the other. (Feynman 1967: 40-41)

Feynman naturally tethers mathematics, replete with its logic, to the empirical world, to the move from one lot of empirical facts to another lot of empirical facts as that movement is given in the mathematical expression of empirical laws such as $F = G(mm'/r^2)$, which in turn are fodder for valid inferences.

But even the quite startling realization that mathematics, language and system of rules that it is, is more than adequate for our empirical, descriptive and predictive purposes[4], is no argument for platonism about mathematical objects and the indispensability of ontological commitment to such objects. Frege's arguments for platonism, which are usefully cast as jointly constituting an indispensability argument, are stronger and more insightful to the extent that they take mathematical language and inferences *as mathematics* rather than as primarily the expressive tools of empirical science. I elaborate on Frege's arguments in the next chapter. But in the closing section of this chapter I want to examine the connection between the commitments of good theories (i.e. explanatory and predictive ones), and *true commitments*. They are not the same, but the good can give us a bridge to the true.

2.5. The good and the true

That there are no leprechauns is *not* a matter of logic or of mathematics. The same goes for flying horses. We've never seen them and we don't need them, and it would be

[4] It is startling, for instance, that non-commutative algebra, developed by Sir William Hamilton from 1843 and published in its most complete form in his *Lectures on Quaternions* in 1853, turned out to be exactly what was needed for Schrödinger to make explicit the equations of motion in his expression of the dynamic vectors of subatomic particles at the quantum level, descriptive equations in which Hamilton's tensors and scalar products, among other features, proved to be crucial. Schrödinger discovered these equations in the mid-1920s, fifty years after Hamilton's death. Small wonder that Wigner writes: 'the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and there is no rational explanation for it' (Wigner 1967: 233), and: 'the miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve' (*ibid.*: 237).

irrational to believe in them, though not strictly speaking illogical. We simply don't have good epistemic grounds for inserting mythical creatures into our inventory of the real, though they are not logically ruled out. Conversely, logical consistency will not tell us what, intensionally speaking, we have good grounds to believe. Someone suffering from persistent, wildly false delusions may well have a logically coherent belief set and draw nothing but valid inferences from it, yet they are essentially mistaken about the real.[5] Now, the metaphysical structural realism that I derive from Quine's commitment-detecting strategy leaves us with an open epistemic question of how to select those objects, from among isomorphic sets of intentionally non-equivalent candidates, to which we shall be committed. In the case of macroscopic empirical ranges of choice, the guiding principles are fairly clear. Choose the snake over the leprechaun. Unobservable entities are a more problematic case, but I do not want to dip into the ocean of that vast literature here because, as I have urged against Quine, mathematical objects are *not* like the unobservables of physics, as we learn from the argument from constraint: mathematical objects are necessary, and they play a role in non-causal, essential explanation. But this question still wants an answer:

> **Query:** The commitments are useful ('good' in that sense), but why think that they're also *true* besides?

The fictionalists' reply to this query will be addressed in chapter 3, so I leave that option in abeyance until then. Another option is to espouse pragmatism in the style of William

[5] I heard a sad and instructive anecdote in this context. A professor had a doctoral student who suffered from paranoid delusions. The student's behaviour became increasingly bizarre; in particular, he thought that his professor was making airplanes follow him around. The professor insisted that, if they were to continue working together, the student must go to a psychiatrist. He did, and came back to his professor. The professor asked "how was the session?" The student replied, "very good, we had a productive talk". Professor: "did you tell her about the airplanes?" Student: "of course not, she would have thought I was crazy." Professor [dismayed]: "well… how do you think that makes *me* feel?" Student: "but you already know: *you're* the one controlling the planes!" Clearly the unfortunate student's reasoning is coherent. (See also the next footnote.)

James when faced with these worries, and reply that 'true' (when predicated of theories, at any rate) is equivalent to 'good'. It's a sort of testimonial to usefulness. But here is an argument to give truth rather stronger legs.

Letting 'RG' stand for 'rational grounds' and 'MP' for 'mathematical proposition', we know:

(RG 1) Some MPs are predictive and explanatory.

It is also plausible that:

(RG 2) Predictive and explanatory power constitute rational grounds for belief.

A bridging premise:

(RG 3) We should believe MPs for which we have rational grounds.

Furthermore:

(RG 4) Canons of inference and standards of justification for MPs are continuous across the applied and the pure.

RG 1 and 2 together lead us to conclude that it is rational to assent to applied MPs; RG 3 and 4 together entail that if it's rational to believe any MPs, it's equally rational to believe all MPs which are established by the relevant canons of justification. And it *is* rational to believe some MPs; by the principle that it's also rational to believe what our well-grounded beliefs entail,[6] we have:

(RG 5) We have rational grounds for believing MPs pure *and* applied.

By standard Tarski-biconditionals for truths (insert your favourite MP for which you have rational grounds, e.g. '1 + 1 = 2'):

[6] Even this is not totally immune from objection, though. Consider the paradox of the preface: you write a book in which you assert $p_1 \wedge \ldots \wedge p_n$ and, being honest, you believe each of these because you have rational grounds for each. But, being modest, you know that you're likely to have made a mistake somewhere, which you confess in the preface. This entails that you also have rational grounds to believe ~ $(p_1 \wedge \ldots \wedge p_n)$. Priest (1998) takes this to be a sign that rational belief is not closed under logical consequence. He thinks that we can and do rationally believe contradictions.

(RG 6) MP $\leftrightarrow$ 'MP' is true

This, I think, is a tidy way of getting from useful mathematics to mathematical truth, or from the good to the true; and mathematical truth brings commitment.

Now, this *still* leaves the selection problem open (i.e. 'precisely which mathematical objects are there?') But it does flesh out structural realism to a realist thesis about the truth-values of MPs, and thence to realism about the objects to which we are committed by those MPs via RG 6. In much of the remainder of this essay, I will examine these issues: the selection problem in chapter 2, objections to my 'rational grounds' reply to the query in chapter 3, and epistemological worries in the final chapter.

3

Frege: definition, indispensability, and objects

Nous sommes dans l'inconcevable, mais nous avons des repères éblouissants
René Char

3.0. Introduction

The story so far goes like this. The IA supplies, first, a *method* of commitment detection through the strategy of semantic ascent, and second, *grounds* for thinking that mathematical thought and talk comit us to specifically those mathematical (especially set-theoretic) objects which feature indispensably in the explanation of the traction applied mathematics delivers. But, building on Quine, the argument from methodological continuity shows that there is no reason in principle to split apart the commitments of applied and pure mathematics; commitment-detection is thus an appropriate method for examining the subject-matter of pure mathematics as well. Furthermore, the argument from constraint provides additional, and *a priori*, support for the genuineness of commitments to such purely mathematical objects as transcendental numbers. Finally, the concluding argument from the last chapter shows that we have rational grounds for the move from the explanatorily effective to the true: that is, grounds for *belief* rather than some weaker epistemic attitude of agnosticism.

But, derived also from my account and extension of Quine, we are left in a position of *metaphysical* commitment to an epistemically *underdetermined* range of mathematical objects. This problem – call it 'the selection problem' – is brought into sharp focus when we ponder the most fundamental of arithmetical notions: the concept of natural number. This chapter is devoted to the selection problem with respect to that concept, and proceeds by way of an account of Frege's thought on the matter. I offer some arguments inspired by, and drawn from, Frege's works. I think that Frege would have endorsed them and I offer some textual evidence for that claim, but I don't pretend that Frege explicitly offered them himself. Rather, this chapter digs out some themes and arguments in Frege's writings for my own purposes, and ascribes them – I think reasonably – to Frege.

This argument for platonism about arithmetical objects is a version of the indispensability argument, with a twist. IAs have traditionally been viewed as *a posteriori*, in the sense of being founded upon the empirical success of a host of physical theories which indispensably involve mathematics, as Quine urged. But a Fregean version of IA can – and should – be read as an *a priori*, semantic argument.

It might seem that this undermines the whole point of indispensability arguments. After all, IAs are forms of argument that move from mathematical language to that language's reach outside itself into the bits and pieces of the physical world; IAs depend on the remarkable fact that mathematical entities manage to get mixed up with bits of matter-energy in space-time in such a way that they play an explanatory role. The objector says that this is reminiscent of the pineal gland: it's explanatorily idle. In this chapter I offer arguments to give Fregean platonism some real content, and to show how these explanations can be managed for mixed contexts.

Section 2.1 examines explicit and implicit (contextual) definitions and offers a preliminary critique of contextual definition. The latter is precisely what Hilbert and Peano strove for: contextual definitions uninfected with particular commitments to a lot of problematic objects. A definition, as we shall see, is a means of *selecting* a subject-matter: one kind of definition or another is thus a mode of approach to the selection problem. *Implicit* definitions would, if workable for arithmetic, undercut the selection problem itself. I will explore Frege's reasons for thinking that the issue cannot be tidily done away with, though. The reasons for this have to do with the inevitable semantic character of inference.

So the remaining sections of this chapter explore the reasons for which Frege rejected Hilbert's and Peano's views, and held onto the determinacy of mathematical – or more specifically *arithmetical* – objects. In the final section I urge that on the basis of these Fregean arguments there is good reason to extend the class of objects of mathematical commitment beyond the class of the empirically applicable, regardless of applicability – or even potential applicability – to physical phenomena. The reason the extension is motivated is that the assumptions, methods, and inference-patterns are, as I have mentioned, continuous through applied and pure mathematics and there's no principled, mathematical or semantic reason to pry them apart.

3.1. Of definitions, implicit and explicit

"When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relations subsisting between the elementary ideas of that science. The axioms set up are at the same time the definitions of these elementary ideas." (Hilbert 1902: 437) So wrote Hilbert, expressing the fundamental principle of implicit definitions for formal systems. What is the implicit/explicit contrast, and what rides on it?

At its most basic, a definition is a bundle of information sufficient for transmitting the meaning of a term or concept to a person who otherwise has no previous cognitive handle on that term or concept. But this is too vague. What is it to transmit a meaning? There is disagreement here, and in the case of definitions of terms and concepts deployed in formal systems like logic and mathematics, the disagreement turns on whether definitions should be implicit or explicit. The passage from Hilbert illustrates his conception of implicit definition. Let's take an example of a very stripped-down (and rather uninteresting) formal system to begin to illustrate.

The following formal grammar consists of only seven symbols and four rules, which I'll call axioms, and which are merely syntactic rules for well-formed formulae (wffs):

Symbols:

R av ar g d S Pi

Formation rules (axioms):

(Use/mention is left intuitive)

1. av and ar may be followed by g, d or R, but only once and by no more than one of those symbols.
2. Strings may begin only with av, ar, S or Pi
3. R closes a string, and only R closes a string
4. av, ar, S and Pi may occur anywhere in a string except as the final symbol.

Since we now have a vocabulary and a syntax, however impoverished, we have a full implicit definition of our symbols, by Hilbert's lights. (Notice that this is not a deductive system: there are no inference rules; this is merely a bare-bones illustration.) We can construct as many wffs of the system as we like. The following is a wff:

S – av – g – S – Pi – S – Pi – ar – R

These are not:

S – av – g – g – Pi – S – Pi – ar – R (this breaks axiom 1)

av – g – Pi – S – S (this breaks axiom 3)

What we have is an if-then system for constructing strings. To start a string, write 'av', 'ar', 'S' or 'Pi'. If you next write 'g', then the next symbol cannot be 'g', 'd' or 'R'. And so on. With these rules, we can both construct wffs and decide, for any string of the symbols, whether it is a wff of the system. So we have both construction and decision procedures for wffs. This completes the implicit definition of the seven primitive symbols.

It is, perhaps, a bit startling that this could count as a definition. One feels cheated, because one feels cheated of *meaning*. But on the Hilbert conception a more complex version of exactly this is what is required for defining the primitive terms of mathematical vocabulary. If we take the geometrical terms 'point', 'line' and 'plane' as examples of primitives, we are actually misled – according to Hilbert – by the fact that these words also occur in natural language: we are misled into thinking that they must have a meaning independent of the axiomatic rules which govern their syntactic behaviour within a geometrical system. As Hilbert famously put it, we ought for formal purposes to be able to replace these terms with 'table', 'chair' and 'beer mug' without loss of formal content. There is no mathematical meaning except that which is given by the governing axioms. It is senseless to ask "what does 'point' *really* mean?"; if you know and obey the axioms, you have the meaning implicitly and there is nothing further to inquire after. It is senseless to ask of the system above, "what does 'Pi' *really* mean?" as well. The reply will simply be: " here's how it behaves by the grammar: look at the rules".

My implicit definitions for the system above say nothing about what the system is *for*, or what – if anything – its symbols stand for independently of the rules that govern their concatenation. This, in Frege's view (and Russell's as well), is intolerable as a definition. Why?

Russell, in the Introduction to the second edition (1937) of his *Principles of Mathematics*, puts the objection snappily: "[t]he formalists [i.e. proponents of Hilbert-style systems] are like a watchmaker who is so absorbed in making his watches look pretty that he has forgotten their purpose of telling the time, and has therefore omitted to insert any works." (Russell 1903/1937, ix) When it comes to defining the natural numbers in the implicit style, the strategy makes numbers' connection with the actual world of countable objects unintelligible (*ibid.*).

In other words, defining numbers implicitly by stating axioms makes counting and other mixed context applications seem thoroughly mysterious. An explicit definition, on the other hand, by giving the meaning of terms (in this case, number terms) *independently* of the axiomatic system in which they figure, is able to lay bare the reach that mathematics has outside itself. Where Hilbert favours an uninterpreted formal system, Frege and Russell demand that the formal system be interpreted explicitly, i.e. that its fundamental terms be assigned meanings.

Returning to my little system, here's an explicit definition of its symbols, or an interpretation of the system. It is a symbolic codification of rules for dancing: it's a way of writing down a choreography (a bad one, I'm sure, but that doesn't matter). Here's what the symbols mean:

Symbol	**Meaning**	**Corresponding movement**
R	révérence	bow or curtsey
av	en-avant	move forward
ar	en-arrière	move backward
g	plié gauche	*plié* leftwards
d	plié droite	*plié* rightwards
S	saut	jump
Pi	pirouette	spin on tip-toe

The rules are designed to allow only choreographies that are physically realistic for dancers, and a choreographic string ends with a bow or curtsey, to mark its conclusion.

The table given above amounts to a semantic interpretation of what is otherwise just a sequence of syntactic procedures, of rules for manipulating symbols. The table gives rules for dancing – I reiterate that they're very silly, but you get the point – not just for writing. The symbols mean something that has reach outside the way the strings are built. Mathematics, for both Frege and Russell, needs an explicit *semantics*, an interpretation, or an assignment of cases. Implicit definition is insufficient for accounting for mixed contexts.

Why, then, might implicit definitions for mathematical terms appear attractive in the first place? There are at least two reasons. The first has to do with the development of non-Euclidean geometries in the nineteenth century. These are, on the face of it, incompatible with Euclidean geometry, which had been thought of as the absolute, unchangeable (indeed sometimes downright sacred) *a priori* science of space. They are "incompatible" at least in the sense that some propositions true of Euclidean space are not true of non-Euclidean space. In light of the emergence of coherent systems of non-Euclidean geometry which contradict some fundamental Euclidean principles, implicit definition of geometric terms appears attractive because it makes no commitment to there being one single "correct" geometry to be chosen over the others, one single geometry that has a monopoly on geometric truth to the exclusion of all others. Since implicit definitions do not assign system-independent interpretations to the non-logical components of the geometric axioms, there is no reason to suppose that there is any conflict at all between rival systems: each has its terms and functions with respect to its own set of fundamental principles, and there is no further requirement that we be able to adjudicate *which* geometry is correct *outside* its own self-contained system. In this sense there can well be mutually incompatible systems of geometry which are all equally *objective*: the objectivity of the derived theorems lies in their very derivability in some system.

The second, related, motivating consideration is that implicit definition seems to secure *a priori* knowledge. A set of axioms is laid down (because it is a self-evident set, or conventional) and its deductive properties are investigated. So long as the resulting system is internally consistent, i.e. it is not possible to derive a contradiction, and so long as the results are interesting and rich in new avenues of research, new discoveries can be made on the basis of those axioms about the properties of, and relations among, the objects of the geometry. If the axioms are self-evident or conventional, they are in any case indifferent to the evidence of the senses; the various deductions carried out on those axioms are similarly indifferent. Implicit definition offers a tidy route to the *a priori*. Frege, as we shall see below, saw these attractive features and explored them in the *Grundlagen* (1884) in an experimental sort of spirit, to see if implicit definition could be made to work for the natural number concept. As it turned out, in his view, it could not.

We can view rival theories of definition as a contrast between definition-as-detection on the one hand (which is what Frege will ultimately endorse), and definition-as-organization on the other (the preferred route for Hilbert and Peano). The "detection" view sees definition as something like a conceptual metal-detector: sweep it over the patch of ground you are investigating, and it will bleep when you encounter an object of interest. It singles out only some relevant objects of interest from the complexities of the environment. The "organization" picture, on the other hand, sees definition more as a matter of choosing an organizing principle, such as a unit system for measuring temperature (degrees Kelvin, Fahrenheit, Celsius…) or a method of classifying books in a library (Dewey Decimal, Library of Congress…). The first *selects* particular objects and thereby excludes others; the second does not.[7] Deciding between pictures is a matter of sorting out just how much selection needs to go on in order to characterize the subject-matter of mathematics and its inferences. Explicit definitions *select*, and thereby commit; implicit ones do not, or at least do not need to. I will trace the fate of contextual definition in Frege's thought in the next section.

3.2. Contextual definition in the *Grundlagen*.

Frege (1884) proposes two alternative contextual (implicit) definitions of numerical singular terms, and rejects them both. Frege's two experiments with this strategy are carried out first by means of numerically definite quantifiers, and next by means of an analysis of 'numerical equivalence', or what has come to be known as 'Hume's Principle' (so baptized by George Boolos). I consider each in turn.

Frege is concerned to show that numerical terms are genuine referring expressions, and furthermore that numbers are indeed objects. In other words, Frege's claim is that numerical expressions are both commitment-making *and* committed to a determinate range of objects. Naturally many substantives *used* as referring expressions are committed to things that just don't exist: nouns occur in fiction too. These substantives are faking it; but Frege wants to show that there's no make-believe in the use of numerical terms in carrying out valid mathematical inferences.

---

[7] It's worth noting that neither of these matches what 'definition' means in ordinary English. This is technical parlance at some remove from the pretheoretical use of 'definition'. As Quine put it, in his entry on 'definition' in *Quiddities*, "one tends to speak of definition in the object-directed sense when puzzled less about the use of the expression than about the nature of its objects" (p. 44).

The first attempt at a contextual definition comes in § 55 and is given by means of numerically definite quantifiers:

> [T]he number 0 belongs to a concept, if the proposition that *a* does not fall under the concept is true universally, whatever *a* may be.
>
> Similarly we could say: the number 1 belongs to a concept *F*, if the proposition that *a* does not fall under *F* is not true universally, whatever *a* may be, and if from the propositions "*a* falls under *F*" and "*b* falls under *F*" it follows universally that *a* and *b* are the same.
>
> It remains still to give a general definition of the step from any given number to the next. Let us try the following formulation: the number (*n* + 1) belongs to a concept F, if there is an object *a* falling under F and such that the number *n* belongs to the concept "falling under *F*, but not *a*". (Frege 1884, § 55)

In current notation, using '$\exists_n x$' for 'there is a number *n* of *x*s such that…', Frege's proposal is to define 0, 1 and *n* + 1 contextually as follows:

(0) $\exists_0 x\, Fx \leftrightarrow \forall x\, (\sim Fx)$

(1) $\exists_1 x\, Fx \leftrightarrow \sim\forall x\, (\sim Fx) \wedge \forall x\, \forall y\, (Fx \wedge Fy \rightarrow x = y)$

***(n* + 1)** $\exists_{n+1} x\, Fx \leftrightarrow \exists_n x\, (Fx \wedge \exists_1 y\, (Fy \wedge x \neq y))$

These definitions of 0, 1 and *n* + 1 provide the conditions governing the use of propositions of the form 'there are 0 Fs', 'there is just 1 F' and 'there are *n* + 1 Fs', and on this basis any *n* can be contextually defined. But Frege immediately recognizes that this first attempt at a definition of number is a failure, because the definition overdetermines the range of possible uses of 'number'; it lets too much into its scope, and by overdetermining the possible contexts of use, the definition thereby *under*determines any particular number. Thus Frege:

> [W]e can never – to take a crude example – decide by means of our definitions whether any concept has the number JULIUS CAESAR belonging to it, or whether the same familiar conqueror of Gaul is a number or is not. Moreover we cannot by the aid of our suggested definitions prove that, if the number *a* belongs to the concept F and the number *b* belongs to the same concept, then necessarily *a* = *b*. Thus we should be unable to justify the expression "*the* number which belongs to the concept F", and therefore should find it impossible in general to prove a numerical identity, since we should be quite unable to achieve a determinate number. (Frege 1884, § 56)

The problem is that this version of a contextual definition does not delimit acceptable uses in fine enough grain, leaving the door open for 'Julius Caesar' to muscle in as a numerical expression, which it obviously isn't. A corollary of this problem is that the definitions given above underdetermine *the* number belonging to a concept, underdetermine 'the number of Fs', because they are not sufficient for establishing numbers as objects[8]. Why not? Frege writes:

> [W]e have already settled that number words are to be understood as standing for self-subsistent objects [in large part because of the syntax of singular terms]. And that is enough to give us a class of propositions which must have a sense[9], namely those which express our recognition of a number as the same again. (Frege 1884, § 62)

---

[8] Frege's notion of objecthood is discussed in 2.3. below.

[9] Note that 'sense' is not, at the time of the *Grundlagen* (1884), to be taken as equivalent to the technical term '*Sinn*' introduced in 'Über Sinn und Bedeutung' in 1892.

Because of the Julius Caesar problem, Frege sees that without identity conditions that apply to all and only substantival occurrences of numerical expressions, there's no way of excluding categorially mixed identity statements like '2 = Julius Caesar'. The concept of number is underdetermined by these first contextual definitions because the identity criteria are too widely applicable, in fact applicable to *any* singular term rather than just to the numerical ones. This makes it impossible to say of the numbers that they are numbers rather than ferns, bricks or conquerors of Gaul. The Julius Caesar problem is one of categorical boundaries, and the first attempt at contextual definition cannot by itself solve the problem. So Frege rejects the first attempt, and goes on to suggest another, this one framed by an analysis of 'equinumerosity' (*Gleichzahligkeit*).

Equinumerosity – or numerical equivalence – is an attempt at framing an identity criterion for mathematical singular terms. Introducing N as a term-forming operator, and '$\mathrm{N}x\mathrm{F}x$' to be read as 'the number of *x*s that are F' and 'F ~ G' as 'there is a one-to-one function mapping every F-thing onto a G-thing and every G-thing onto an F-thing', the second contextual definition runs:

(E) $\mathrm{N}x\mathrm{F}x = \mathrm{N}x\mathrm{G}x \equiv \mathrm{F} \sim \mathrm{G}$

and claims that the number of Fs is equal to the number of Gs if and only if there is a function such that F ~ G. (Frege 1884: §§ 62-63) The strategy is designed to provide 'a means of arriving at a determinate number and of recognizing it again as the same, [so that] we can assign it a number word as its proper name'. (Frege 1884: § 62) And indeed Boolos (1998: 276ff.) has shown that (E), if interpreted as taking the infinite set of natural numbers as a model, and added to second-order logic, in fact yields the Peano axioms of arithmetic. (This is what Boolos has dubbed 'Frege's Theorem'.) But what's wrong with (E)? Once again the underdetermination problem rears its head and Frege rejects this second attempt at contextual definition not for formal (axiomatic logical or mathematical) reasons, but because of problems connected with the semantics of identity statements generally.

Michael Potter puts the objection like this:

> The numerical equivalence lays down a condition for the truth of what is apparently an identity sentence, 'N*x*F*x* = N*x*G*x*'. What it does not do is to settle the meaning more generally of sentences of the form 'N*x*F*x* = *q*', where '*q*' is a singular term, except for the one case where *q* is given the form 'N*x*G*x*'. (Potter 2000: 75-76)

Even with an identity condition that applies to all and only substantival occurrences of *strictly numerical* singular terms (leaving 'Julius Caesar' behind and letting *q* = '9', say), there is no statement of the identity conditions governing the use of numerical singular terms unless they are given the form of 'the number of G-things', i.e. of numerical definite descriptions, i.e. the form in which a numerically definite quantifier is applied. And we cannot translate every occurrence of a numerical singular term into a numerically definite description simply because '9' is not everywhere substitutable for 'the number of planets', as belief contexts and modal contexts show. So much, it appears, for contextual definition in *Grundlagen*. As Frege puts it:

> It is not only among numbers that the relationship of identity is found. From which it seems to follow that we ought not to define it specially for the case of numbers.
>
> […]
>
> Our aim is to construct the content of a judgement which can be taken as an identity such that each side of it is a number. We are therefore proposing not to

> define identity specially for this case, but to use the concept of identity, taken as already known, as a means for arriving at that which is to be regarded as identical. (Frege 1884: § 63)

Frege seems to have thought that the underdetermination problems that face his contextual definitions are lethal to the project, and so he goes on in *Grundlagen* to supply an explicit definition of number in terms of the extensions of concepts. The next section addresses this definition, but I will return to the questions raised here in sections 2.3. (on objecthood and the context principle) and 2.4. (on sense, reference and concepts).

3.3. Apriority, explicit definition and Frege's extensions.

'By their deeds shall ye know them', say both the Bible and the behaviourists. It's not a bad slogan for partisans of contextual definition, the idea being that appropriate contextual observations of the "behaviour" (so to speak) of an object – whether it's a number or a fish or a planet – will add something to our knowledge of the object that an explicit definition of necessary and sufficient conditions for being of that object-kind just won't. One of Frege's principal aims in *Grundlagen* is to rid arithmetic of any sensory or otherwise empirical intuition and supply *a priori* conditions for tracing the boundaries of the number concept, with the ultimate goal of giving a full explanation of the process of mathematical proof. On the road to this goal, a stab or two at a contextual definition is a good idea, because a contextual definition of number has a chance at being both *a priori* and ampliative, i.e. both given independently of empirical intuition and genuinely informative. Note that this is *not* equivalent to Kant's conception of the synthetic *a priori*, because synthetic *a priori* claims *are* dependent upon empirical intuition. In Frege's words:

> [T]hese distinctions between a priori and a posteriori, synthetic and analytic, concern, as I see it, not the content of the judgement but the justification for making the judgement. Where there is no such justification, the possibility of drawing the distinction vanishes. (Frege 1884: § 3)

There's no reason to say that a claim justified *a priori* cannot be informative, *pace* Kant. 'On the basis of [Kant's] definition, the division of judgements into analytic and synthetic is not exhaustive.' (Frege 1884: § 88) Informativeness is a question of content, but apriority is a question of the kinds of epistemic resources we appeal to in *justifying* the claims we make. The *a priori* is not a stack of sentences on the left-hand column of the great compendium of analytic and synthetic claims[10], but is rather a name for a procedure of justification whose methods do not involve empirical intuition of any kind. So no question can arise over whether or not contextual definitions like (E) are *a priori*, even though they are informative, except to the extent that we wonder about the procedures involved in justifying (E) and its ilk. Frege's claim is just that the definition of number can be given without reference to any form of empirical intuition, and thus that the concept of number is *a priori* because our understanding of it is *a priori*, be the definition explicit or contextual. The contextual definitions failed (failed, at least, to Frege's mind) for reasons discoverable *a priori*: they failed because both of them overdetermined the extension of 'number' and so failed to define the concept of number generally, in such a way that it would include all and only numbers.

What, then, of the explicit definition? Well, in arguing against Kantian intuition, Frege writes of any kind of definition (contextual or explicit):

---

[10] This is the great compendium that the more hardboiled logical positivists hoped one day to have, with logical and mathematical claims on one side, informative empirical claims on the other, and every claim not in the book banished to moulder in the Theology Department.

> [T]he more fruitful type of definition is a matter of drawing boundary lines that were not previously given at all. What we shall be able to infer from it, cannot be inspected in advance; here, we are not simply taking out of the box again what we have just put into it. (Frege 1884: § 88)

Frege's point is one about the way in which claims about the boundaries of the number concept can be justified. There's much to be said on this issue (I'll say very little of it), and it's an issue that lies at the historical and conceptual root of the divergence between constructivists, platonists and fictionalists. But I will mostly ignore the historical aspects[11]. From the ontological point of view, the contextuality or explicitness of a definition of 'number' doesn't matter *except to the extent of its commitments*, commitments which are to be judged, like all ontological commitments, in relation to their fruitfulness. But in relation to what *purpose* are we to judge fruitfulness? Well, in the case of the Fregean attempts at a definition of the number concept, the purpose is an explanation of the justificatory process involved in arithmetical proof, and that's a process that – says Frege, anyway – crucially involves commitment to the existence of mathematical objects where 'commitment' is defined by the criteria offered in my discussion of the IA in the previous chapter. The explicit definition of 'number' given in *Grundlagen* is an attempt at an ampliative *a priori* account of the concept of number, and it's famous because it's dead wrong.

In more or less modern notation, Frege's explicit definition is this:

(D) $\mathrm{N}x\mathrm{F}x =_{\mathrm{df}} |X\text{: } X \sim \mathrm{F}|$

which, in English, stipulates that

[11] With the rider that there *really is* an elephant in the room. I'm just not paying attention to it.

> (D') The number of xs that are F is equal to the cardinality of some set X, where X is defined by the following property: there is a function mapping every member of X onto every F-thing, and mapping every F-thing onto a member of X.

(E) follows from (D), but (D) is fatally flawed: we can derive Russell's Paradox from it. Suppose that, on the right-hand side of the definitional equation, 'X' stands for 'the set X whose members are all and only those sets which are not members of themselves'. The point of (D), or Basic Law V in Frege (1893), is to give an explicit identity condition holding between concepts that determine all and only the same objects. But once again, as for the contextual definitions, the concept is underdetermined because it overgenerates: it's an *unrestricted* axiom. Moreover, as a concequence of its overgenerating, it leads to a contradiction. What's more, Basic Law V cannot be rescued simply by restricting the value-range of X to the numbers, precisely because the equivalence is meant to *define* that value-range. Identity-conditions are procedures for abstraction, and the hope was that the procedure would yield a *categorical* theory (in the logical sense of 'categorical' in which a categorical theory determines its objects uniquely). Nor can we rescue it by indicating that the resulting paradox is merely an indication that there can be no such impredicative set as the set whose members are not members of themselves: if a system of axioms leads to a contradiction, that is a *reductio* that infects the system as a whole. Anything at all follows from a contradiction (at any rate if, as in classical logic, we accept that *ex falso quodlibet* is provable), so since a contradiction is provable from the unrestricted axiom of abstraction, the system falls. In the end, all three of Frege's attempts at a definition overgenerate and underdetermine the numbers.

So far I've said little about Frege's arithmetical platonism, and indeed the failure of a definition of the concept of number is not in itself an indictment of platonism about mathematical objects. One principle of abstraction down the tubes needn't *necessarily* drag any others with it. What needs to be made clearer is the Fregean notion of 'objecthood', and the sense in which numbers are objects, the sense in which number terms are genuinely referring singular terms. This is the issue I address in the next section.

3.4. Objecthood and context.

Frege enjoins us 'never to ask for the meaning of a word in isolation, but only in the context of a proposition.' (Frege 1884: x) The context principle is crucially related to the notion of objecthood and so in this section I discuss the two together. Recall that above I gave a depiction of Frege's core semantic theory as '$\exists f\,(f\!: \ulcorner \mathrm{F}x \urcorner \rightarrow$ T-VAL)' : note that here 'F*x*' is a representation of a whole sentence expressing a thought: in order for it to be a meaningful expression at all it needs to satisfy the function *f* to a truth-value, and single words or subsentential expressions in general are thus by themselves meaningless because subsentential expressions by themselves do not have truth-values.

Kit Fine describes the context principle as follows:

> Let us suppose … that we are given a language *L* whose terms and sentences are already understood and that we add new terms to *L* to form the (as yet only partially interpreted) language *L*+. It is our aim in applying CP [the context principle] to *L*+ to secure reference, or at least potential reference, for the new terms of *L*+; and we do this by adding a set of 'bridge principles' *B*, which relate the truth (and falsehood) of sentences in *L*+ to the truth and falsehood of

sentences in *L*. We might call this set of bridge principles a *contextual definition* or *definition by CP* for the new terms. (Fine 2002: 60. Italics in original.)

The bridge principles *B* will constitute a semantic theory for *L*+, in this case the language of arithmetic, and will allow a truth-conditional specification of the value-ranges of the terms of *L*+. In introducing arithmetic we are effectively introducing a new collection of terms, predicates and operations whose syntax is specified but for which we have as yet no semantics. The context principle is a constraint on the elaboration of any such semantic theory: even if we were to accept the possibility of an explicit definition of number in *L*+, we would still be unable to specify its – that is, the arithmetical language's – semantics independently of the kinds of contexts in which *L*+'s terms t_1, t_2, t_3 and so on occur. The context principle is designed to constrain the interpretations we can give to such terms, *and* to ensure that the semantics we give for *L*+ will be the right one by assigning an interpretation to all and only those terms occurring in contexts in which they are meaningful and have a shot at being true. 'Eight' or 'number' have no determinate meaning by themselves independently of the compositional contribution they can make to sentences in which they occur. We cannot understand these terms independently of an explanatory context that specifies the semantics by in turn specifying a set of uses that have truth-conditions. This is why Frege's explicit definition is framed in such a way as to yield a determinate truth value for the statements of the identity conditions of numbers-as-objects.

But once the context is specified, we are in a position to generate a semantics, and this is the stage that has ontological import. Thus Frege, retrospectively describing his strategy:

> We … laid down the fundamental principle that we must never try to define the meaning of a word in isolation, but only as it is used in the context of a proposition: only by adhering to this can we, as I believe, avoid a physical view of number without slipping into a psychological view of it. Now for every type of object there is one type of proposition which must have a sense, namely the recognition-statement, which in the case of numbers is called an identity. Statements of number too are, we saw, to be considered as identities. (Frege 1884: § 107)

The 'recognition-statement' is a statement of the conditions under which we can recognize an object as being the same object on another occasion, i.e. an identity statement of the kind discussed in the two preceding sections. Frege's platonism is expressed in the conjunction of two claims:

> **C1**. in arithmetical propositions number-terms occur substantivally as object-terms.
>
> **C2**. there are true arithmetical propositions, and in order to supply them with a truth conditional semantics (which will also account for the falsehood of the false ones) we must postulate abstract objects, i.e. numbers.

In this section I have mostly discussed how to arrive at numbers' semantic objecthood; more needs to be said about their objectivity. This is the topic of the next section.

3.5. Sense, reference and arithmetical objectivity.

Frege's familiar distinction between sense and reference (Frege 1892) is designed in part to solve problems of natural language semantics connected with the truth conditions of utterances in intensional contexts (of belief and *oratio obliqua*) and modal contexts. But the distinction can serve as a powerful tool for analyzing purely extensional arithmetical contexts as well. The point of the distinction is to find a method for assigning truth conditions to propositions considered as composite, structured wholes expressing determinate thoughts. The sense (*Sinn*) of an expression is its 'mode of presentation', i.e. our means of cognitive access to its referent (*Bedeutung*), especially by way of a package of information sufficient for unique determination of the referent. Sense and reference together yield a meaningful proposition. Moreover, the distinction allows for the shareability of 'judgeable contents', the communicability of propositional content, by pointing out that it is when senses are shared between speakers – when two or more speakers cognitively pick out the referent in the same way under the same description – that communication genuinely occurs.

The distinction does (at least) two jobs: it offers a method for assigning truth conditions to expressions, and it explains a phenomenon familiar to every speaker, the phenomenon of linguistic communication. But it's a distinction that has ontological import: if the phenomenon is to be adequately explained, we need the right sort of semantic picture. Frege offers this semantic picture, and it's one which demands that objective, public, shareable senses be worked into the ontology, in addition to the referents the senses identify. Hence the objectivity of propositions (or thoughts or contents) on Frege's semantic picture: Frege's semantic platonism – platonism about propositions – derives from the consideration that objective propositions are indispensable in the explanation of the phenomenon of linguistic communication. They cannot be eliminated in favour of a physical counterpart because no physical object does an explanatory job equivalent to the work propositions do; and they cannot be eliminated in favour of mental entities because mental entities are inherently private and subjective and therefore not shareable in the way that linguistic communication demands. The third option, the Third Realm, is all that's left.

This very quick account of Fregean natural language semantics is meant just to lead into the discussion of the role that sense and reference play in the semantics of specifically arithmetical language. The Julius Caesar problem discussed in previous sections is a semantic problem: the question is whether there can be a way of assigning truth conditions to arithmetical sentences which will at the same time exclude the possibility of absurd mixed identities between categories of sortals such as 'Julius Caesar is the number 2'. The problem can't be solved merely by jury-rigging the arithmetical domain of definition, as we've seen, and moreover any solution will have to account for the fact that while Julius Caesar is not a number, he can still be counted: there's one of him. We can make sensible inferences such as 'Julius and Brutus are in the room; so there are at least two people in the room', and thus we need an account that is sensitive both to the semantics of pure arithmetical propositions (such as '7 + 5 = 12') and to the semantics of the applications of those propositions in non-arithmetical, or in mixed (e.g. 'The number of Jupiter's moons is four'), contexts.

The account that Frege ultimately arrives at is that number is a property of second-order concepts taking first-order concepts as objects, a.k.a. first-order predicates (or properties defining a set). He lights on this account as the only satisfactory way of giving a theory of how inferences in mixed contexts are possible. Thus Frege:

> I characterize as a concept that which has number, and in so doing indicate that the totality [encompassed by a set], which is here our primary concern, is held together by characteristics [i.e. a common property or properties], not spatial proximity, which latter can obtain only in special cases as a by-product of these

> characteristics, but which generally speaking is unimportant. Thus even the number zero, which otherwise is completely without bearer, becomes intelligible… Only in this way is it possible to fulfill the first requirement of basing all modes of inference that appear to be peculiar to arithmetic on the general laws of logic. (Frege 1885: 144)

And again:

> We do ask … e.g. 'How many moons has Mars?' or 'What is the number of Martian moons?' and from the answer 'The number of Martian moons is two' we learn something worth asking about. So we see that both in the question and in the answer we get a word, or a complex designation, for a concept… (Frege 1894: 81)

Number, thus, 'belongs to' a higher-order concept that subsumes a set as falling under its extension, and the set – defined by a first-order property, a property common to objects falling in the set – in turn is an *object* in its own right, denoted by a singular term.[12] On this analysis, number is both extensional and objective. The sense of a numerical singular term is the mode of identification of the extension of the set-forming predicate ('____ is green', for instance), and the referent is the extension thus determined.

The way of confirming that this is the right analysis is to check it against some of the problems an analysis must solve. The problem at issue here is of accounting for mixed-context inferences such as the one about Julius and Brutus mentioned above, and the analysis fares well in the face of the challenge. Thinking of number as the extension of a concept allows for this kind of inference, without regard for whatever particulars of

[12] There is an ambiguity in Frege's treatment of numbers as properties on the one hand, and objects on the other. The latter treatment is motivated and defended in the final section of this chapter, below.

whatever kind actually fall under the first-order predicate that defines the object-set. The problem of the unrestrictedness of the property-forming operation in Frege's naïve set theory remains, of course, and perhaps it can be solved by a Russellian hierarchy of logical types; but that's another story, connected with the success or failure of logicism as an epistemic programme for the foundations of mathematics, rather than with the adequacy of Frege's arithmetical platonism *qua* ontological thesis about the subject matter of mathematics.

3.6. Structure, and the definition question again: Frege and Peano.

The Peano axioms of arithmetic derivable from (E) provide a contextual definition of the elements in any arithmetical progression whatsoever, as follows:

P1. Zero is a number.

P2. Every number has a successor.

P3. Zero is not the successor of any number.

P4. No two numbers have the same successor.

P5. If 0 has the property P, and P is true of the successor of every number for which P is true, then all natural numbers have the property P.

With only two primitives ('zero' and 'successor') and the principle of mathematical induction (axiom 5), Peano is able to characterize the abstract structure of any discrete well-ordered sequence that has an initial term, is closed under the successor relation and satisfies mathematical induction. Out of this one can generate the whole of arithmetic; so why, from the Fregean point of view, is this not satisfactory? The answer to this will make it clearer why Frege rejected his own and anybody else's attempts at strictly contextual definition and opted instead for the explicit definition of 'number'.

Peano's axioms offer an implicit definition of 'number' because they characterize the numbers exclusively by giving the structure of the relations between whatever we might be inclined to call 'numbers', without assigning any *particular* model that actually satisfies the axioms. This is to say that the contextual definition is free of semantics: no particular values are assigned to 'zero', 'successor of zero', etc. A model in which 'zero' is assigned the value 2, and 'successor' is interpreted as meaning 'the result of adding 3', satisfies the Peano axioms because it assigns an initial term, defines the successor relation, and mathematical induction applies to it. Here, we've offered the model:

M = {2, 5, 8, 11, 14, 17, 20, 23, …, *n*, successor of *n*}

and the model satisfies the axioms. It also so happens that

M' = {0, 1, 2, 3, … *ad infinitum*}

satisfies the axioms too, but the sequence of natural numbers in model M' does not occupy any privileged position among the infinitely many possible models, possible semantic interpretations, that we can assign to the axioms. This indeed is the great advantage of the Peano contextual definition: it characterizes abstract structure without assigning a determinate semantics, and so it tidily answers some difficult questions. How is it that arithmetic can cope both with denumerably infinitely large numbers and the relations between them, as well as serve our purposes in ordinary language when we count heartbeats and twigs? The reply is that these are just instances of sequences that share a common structure, abstractly given in the axioms, and they are both legitimate collections of substitution instances for the uninterpreted primitives of the axioms. But why, then, should Frege be dissatisfied with this account?

Frege's objection comes from his insistance on the principle that mathematical thoughts expressed in judgements be *determinate* thoughts, that the objects invoked in the thoughts be characterized according to precise identity conditions, and that the signs for concepts be similarly uniquely characterized. Thus Frege:

> Let us suppose that we have two definitions and that both of them give a meaning to the same sign. Then there are only two conceivable cases: either both of them give the same meaning to the sign, or not. In the first case, we have again two possibilities: either both definitions bestow the same sense on the sign and say exactly the same, or not. In the first case, one of the two is superfluous; in the other, it would have to be proved that they assign the same meaning to the sign even though they give it different senses. Thus one of the two would have to be allowed to stand as a definition while the other would have to be changed into a theorem and proved. The reader is cheated of this proof if what should be a theorem is presented as a definition. Finally, if the definitions give different meanings to the same sign and not just different senses, then they contradict each other, and one of the two must give way. (Frege, Letter to Peano, 29 September 1896, in Frege (1980: 114))

A theorem, or indeed a true mathematical proposition generally, must express a determinate thought. Frege is here objecting to a purely formal characterization of the truth-conditions for arithmetical propositions such as is given by the abstract structure of Peano arithmetic. Another passage from Frege:

> It can still be asked [of Peano's system] what happens if something which is not a numeral at all – say the sign '⊙', for the sun – is inserted… The sentence '$\odot > 2$' is false, because the sun is not a number, and only numbers can be greater than 2. Accordingly the sentence
>
> '$(\odot > 2) \supset (\odot^2 > 2)$'

> would be true, regardless of whether its right hand side were true or false – and it ought to be one or the other. However, according to the usual explanation of the combination of signs 'x^2', the right hand side has no reference. Thus there must be presupposed here an explanation of 'x^2' such that a reference is always forthcoming no matter what sign is inserted for 'x', provided only that this sign itself has a reference – i.e., provided only that it designates an object. This makes evident the necessity of my claim that functions should be explained in such a way that they receive a value for every argument. (Frege 1897/1969: 10)

In other words, the reference of number terms needs to be secured in advance of the system of relations into which they can enter: hence the requirement of an explicit definition and the inadequacy of the implicit ones. Since purely formal, structural interpretations of mathematical languages (*à la* Peano or Hilbert, say) fail to secure reference, they thereby fail to express determinate thoughts and will therefore either fail to have a truth-value at all (which would make them meaningless) or be false (which would make them useless).

The account given of Frege in this section is an account of some negative views: reasons for rejecting a purely formal characterization of mathematical language and for underlining the importance of the semantic for the proper interpretation of the mathematical. In the next section I turn to the more positive aspects of Fregean indispensability.

3.7. Determinacy

To recap a bit: Frege rejects contextual definitions because they do not settle the domain of definition for identity claims such as $NxFx = q$, where 'q' is restricted to naming only one of the numbers: not walnuts, ears or emperors. Thus contextual definitions don't give a method for assigning a privileged interpretation to one model over another, and Julius Caesar works his way into the picture as a candidate for identity with the number 2. In fact this is precisely the feature that the structuralist points to as a great advantage; focus on the relational structure of a system of entities rather than the entities themselves, and you thereby solve a number of puzzles: on the one hand the applicability of mathematics is no longer a mystery since entities in the physical world are arranged according to structures mirrored in mathematical language and can serve as one among many lots of values for the variables invoked in mathematical abstractions, and on the other, the epistemic problem is transformed into one of working out how human beings manage to recognize and codify patterns in the world of experience: patterns are there for the detecting and we don't need to invoke a mystical relation between human minds and causally inert abstract objects.

But Frege wants *the natural number sequence*, not just any sequence that shares its structure, or a portion of it. He wants an explicit delimitation of the identity conditions for numbers because number is a property of *concepts*, and not of objects as the structuralist interpretation would have it. The structural interpretation makes it look as though number were a property of objects occupying certain positions in structures (Peano structure, for instance). In general, the structural view is that '*n*' is elliptical for '*n*th place in sequence S' and that number terms do duty for properties of objects, objects that stand in certain ordered relations to each other. But this is a mistake: number belongs to concepts, more specifically to their extensions, and not to objects, for every object is one.

Still, why the insistence on determinacy for the natural number sequence? Dummett's exegesis of Frege on this point is interesting but, I think, mistaken:

> It is possible [on Frege's view] to make an inference only from a thought (only from a true thought, that is, from a fact, according to Frege): it would be senseless to speak of inferring to the truth of some conclusion from something that neither was a thought nor expressed one. (Dummett 1991: 256)

Chihara (2004: 249f) raises an objection to this account (as a position in its own right, not necessarily as a bit of Frege exegesis). The objection points to legitimate inferences made on the basis of *false* premises: if Holmes knows that either A or B is the murderer, and Dr Watson tells him that it cannot possibly have been A because of a medical condition of A's, Holmes deduces that B did the deed. But let's say that Watson is mistaken about the medical facts, and A was the killer after all. On Dummett's construal of Frege, it looks as though Frege is bound to say that Holmes has not made an inference at all, which is patently absurd: Holmes has drawn a valid inference, though it happens to be unsound. This seems to me to be reason enough for thinking that Dummett must be wrong about Frege in the parenthetical remark: surely Frege saw the difference between validity and soundness! We cannot possibly restrict the range of valid inference to the range of sound inference. But Chihara makes heavy weather of a parenthetical remark of Dummett's which could easily be excised, leaving us with the proposition that for Frege, sound inferences can only be made if the inference is a sequence of determinate thoughts. The reason for this is that only determinate thoughts have a determinate truth-value and, valid inference being the truth-preserving beast that it is, only a sequence of determinately true thoughts can yield a true conclusion[13].

But I want to make a further point in defence of the need for the determinacy of arithmetical propositions: that is, the possibility of the semantic applicability of numerical terms and the validity of (some) inferences in 'mixed' contexts, that is, contexts which bring numerical terms and operations into natural language use, such as 'if you believe that you have two children, you must also believe that you have at least one child' or 'since there are two ferns and a geranium on the sill, there are three plants'.

There are both pure and mixed contexts in which arithmetical terms – or mathematical terms generally – occur: the mixed are as instanced by the examples above, and the pure are those contexts in which propositions involve only the occurrence of arithmetical terms and operations (e.g. '2 + 2 = 4') to the exclusion of all others; and it is worth pointing out that most of mathematics is in fact expressed in *mixed* contexts. Pure mathematics is very little calculation, and very much metamathematical reasoning about the relationships between kinds of mathematical objects and their properties. When mathematicians investigate the hierarchy of sets or the properties of Abelian groups, they do so in mixed contexts, vastly more complex than the inference to three plants but of basically the same kind. Hence the requirement of determinacy: if truth-preserving inference is to be made in mixed mathematical contexts, the meanings of the terms used in the process of the inferences must be constant and settled in advance of the structure of the inferences themselves. A purely strutural characterization of the natural number system such as the one given in Peano Arithmetic would make mixed contexts mysterious to the extent that it is silent about the semantics of mixed contexts, and

[13] I reserve consideration of truth in fiction and the possible role of fictions in general for full treatment in Chapter 3 below.

therefore unable to deliver the determinate thoughts that have a truth value. We are therefore required to settle the mixed identities of the form N*x*F*x* = *q*, and to do it by means of explicit definition for the sake of determinacy. That definition holds that number belongs to the extensions of concepts, and that any particular expression of number thus characterized is a putative referring expression; moreover, it is a *genuinely* referring expression because the inferences involving it – both pure and mixed – must be either true or false. Mathematical objects are thus indispensable in accounting for validity. In the next section I offer some more considerations for motivating this last claim.

3.8. Why objects?

Why should we be committed to objects for the purposes of arithmetic, rather than to properties or concepts or something else that is not itself an object in Frege's logical sense of 'object'? Why think of number terms as singular terms, and not as any other kind of term? The motivation for a Fregean reply to this question comes from the theory of thought, and from reflection on what it means to entertain an arithmetical thought and make arithmetical judgements.

Thoughts are structured wholes that stand in a complex system of inferential relations to each other. Following Frege, "I mean by 'a thought' something for which the question of truth can arise at all" (Frege 1918: 353), thereby disallowing incomplete or unsaturated expressions, which have no truth value one way or the other. Furthermore, Frege writes: "How does a thought act? By being grasped and taken to be true" (*ibid*., 371). I propose this reading of Frege's meaning: that a thought is a structured whole, having a truth value, and available for inference: *grasping* and *taking to be true* enable a thought to 'act' by placing it in the appropriate set of inferential relations to other thoughts: a thought 'acts' through its inferential relations. What, then, is the required structure for something to be of the kind that can have an inferential role, and be available for truth-preserving inference? What components qualify something for inclusion in a deductive system?

The components are first of all syntactic, or rather the syntax of portions of language is the way of encoding the structure of a complete thought. The principal syntactic feature is the distinction between functions and arguments already laid out in *Begriffsschrift*, where we read as follows:

> Let us suppose that the circumstance that hydrogen is lighter than carbon dioxide is expressed in our formula language. Then in place of the [logical] symbol for hydrogen we can insert the symbol for oxygen or that for nitrogen. This changes the sense [meaning] in such a way that 'oxygen' or 'nitrogen' enters into the relations in which 'hydrogen' stood before. If an expression is thought of as variable in this way, it splits up into a constant component, which represents the totality of relations, and a symbol which can be thought of as replaceable by others and which denotes the object that stands in these relations. The former component I call a function, the latter its argument. The distinction has nothing to do with the conceptual content, but only with our way of grasping it. (Frege 1879: I, §9, Beaney's (1997: 65-66) translation.)

This is a rich little passage. The central thesis is that different ways of carving up logical form, different readings of the logical syntax of the complete expression of a thought, make different inferential features of that complete thought salient. When we turn our attention to the particular case of number terms, we observe that they can occur both adjectivally and substantivally, available for interpretation as standing for properties on the one hand, and for objects on the other. Such is the contrast between 'Jupiter has 4 moons' (an adjectival reading) and 'the number of Jupiter's moons is 4' (a substantival, or singular-term, occurrence). The first case has the form of a numerically definite quantification: $\exists_4 x$ (moon of Jupiter (x)); the second has the form $\exists n$ (number of Jupiter's moons (n)). I have already examined Frege's motivation for rejecting the first kind of interpretation (§ 2.2), which is that numerically definite quantifiers overgenerate their domains and thus underdetermine their objects (the 'Julius Caesar problem'). The goal of a logical language that will make arithmetical inferences transparent is partly to achieve a way of capturing the relevant domain which is in some sense *understood in advance*, though not with full precision. If the language allows into the domain such objects as Julius Caesar which ought clearly to be disallowed, then the language has failed to achieve its goal and it needs revision. Such is the further motivation for Frege's conception of definition-as-detection, as specification of the domain of discourse in question. Now I want to say more on behalf of the second reading, on which numerical expressions should have their occurrences in logical syntax read as singular-term occurrences, especially in connection with the theory of arithmetical thought.

A complete thought is that which is available to stand in inferential relations. A complete arithmetical thought saturates a function with one or more arguments, and yields a truth-value. The inferential relations made salient by the function-plus-arguments analysis of the logical syntax of arithmetical language are such that, holding the functional components constant and altering the arguments, we may alter the truth-value of the whole thought so expressed. Furthermore, our interest in discerning the logical form of arithmetical inference lies in our ability to make transparent the structure of inferences in which a 'constant component' or functional expression – such as '___ is prime' – is multiply instantiated by a range of objects denoted by singular terms used to saturate the function and thereby yield a truth value. Singular terms are required in order to make salient what the subject matter of inquiry really is, in the case of number theory. The theory of thought is the theory of how thoughts stand in inferential relations to each other, and in the case of arithmetic the inferential relations are structured in such a way that the constant components reoccur in the course of inference, while the objects to which the constant component is ascribed change. This is how we are able, for instance, to make salient the relevant patterns of inference involved in applications of the principle of mathematical induction over the integers.

It is helpful to have an example in mind. Begin with the principle of mathematical induction, articulated as follows. Suppose that the property P holds of some integer n, or, in more Fregean parlance, that the function $P(\ )$ with n as argument yields the value T. Furthermore, for every integer $k \geq n$, if P holds of all integers t such that $n \leq t \leq k$, then P holds of $(k + 1)$. In that case, $P(m)$ holds for any $m \geq n$; that is, the property cascades, or is heritable, through the whole sequence of integers from the initial term.

Now let's take one part of the fundamental theorem of arithmetic, that every integer $n \geq 2$ is either a product of primes, or itself a prime. Since we can take primes as single-factor products of primes ($n^1 = n$), the theorem is equivalent to the statement that every $n \geq 2$ is a product of prime numbers. (The other part is the statement of *unique* prime factorization, which I will ignore.) The proposition of prime factorization can be established from the principle of induction over the positive integers, the universe to which we restrict the discussion. Letting P be the property of prime factorability, the claim is that $\forall n \geq 2, P(n)$. The trick is to show that, for integers $k \geq 2$, the base case $P(k)$ and the inductive step $P(k + 1)$ both hold. P clearly holds for $k = 2$, our base case, since 2 is prime. Now, the inductive step is as follows: for every $k \geq 2$, if $P(t)$ holds for all t such that $2 \leq t \leq k$, then P holds of $(k + 1)$, which we establish by this reasoning: if $k + 1$ is prime, there's nothing to show; if $k + 1$ is composite, that means that it has factors a and b. Then $2 \leq a \leq k$ and $2 \leq b \leq k$, and by the induction hypothesis, a and b are each the products of primes. It follows that $k + 1 = ab$, a product of primes.

Every integer strictly greater than 1 is a product of primes, then. This universally quantified proposition has the inferential consequence that the property may be instantiated on every object in the domain: pick an integer, and it will have a prime factorization. What the argument from the previous paragraph does is make salient the heritability of a particular property possessed by infinitely many objects. This feature of mathematical thought gets syntactically coded as a multiply-instantiable function taking, in argument place, potentially infinitely many singular terms. These are at least *treated as if* they were referring expressions. Note that the foregoing considerations do nothing to *secure* reference for the arithmetical singular terms; that is another argument, or battery of arguments, already deployed. What it does is rather to motivate our treatment of the language of arithmetic *as if* numbers were objects, rather than concepts or properties, in order to make transparent the kinds of inferential relations that stand among arithmetical thoughts.

Nevertheless, the fictionalist option is still open. Conceivably, the fictionalist can accept the Fregean substantival reading, and happily treat the substantives *as if* they referred, without insisting upon actual reference. I devote the next chapter to assessing that proposal in various forms.

4

Fiction, quarantine, and the norm of truth

La littérature mène à tout, à condition d'en sortir.
-- Vercors

4.0. Introduction: the common core

A familiar procedure in both logic and mathematics is to make assumptions and formulate hypotheses for exploratory purposes, but to put those assumptions in quarantine: they are isolated from assent, ontological commitment, truth ascriptions and so on. So it goes when a supposition is made for indirect proof, for instance: to show a proposition P, assume that P is false, and then show that P's falsehood leads to an absurdity. The proposition P, assuming bivalence, is thereby proved. But there is a view in the philosophy of mathematics that holds that this epistemic quarantine, where hypotheses are isolated from belief, should extend to the whole body of what we normally take to be mathematical truths. Shakespeare captures the strategy rather nicely in the following passage from *Hamlet*:

> Your bait of falsehood takes this carp of truth;
> And thus do we of wisdom and of reach,
> With windlasses and with assays of bias,
> By indirections find directions out.
> (*Hamlet*, Act II, sc.I)

But there is a view in the philosophy of mathematics that holds that this epistemic quarantine, where "windlasses and assays of bias" are isolated from belief, should extend to the whole body of what we normally take to be mathematical truths. Now, any mathematical fictionalism, to the extent that it is properly termed a fictionalism at all, shares a common core claim with all the other fictionalisms. This core claim is that mathematical propositions are exactly like the propositions expressed in overtly fictional works such *Lolita* and *King Lear*, all across the ontological, semantic and epistemic dimensions of analysis. In the ontological dimension, mathematical propositions no more require actual reference to abstract objects than fictional propositions require actual reference to King Lear. There is no Lear, and there is no number π either. Along the semantic dimension, mathematical propositions are true only in the sense in which fictional propositions are true, in respect of their particular fictional frame: 'π is transcendental' is true, if it's true at all, only in the way in which 'Lear descends into madness' is true, whatever that way may be. In the epistemic dimension, mathematical propositions are not to be believed except in the sense in which we believe that Lear descends into madness. Mathematical propositions are sequestered in epistemic quarantine: if we were to release them, say the fictionalists, we would be infected in turn with an inappropriate ontology and semantics. Quarantine as an epistemic stance is the thesis that mathematical propositions are not to be included among the cognitive inventory of what is authentically believed. It is not merely an agnostic stance: it is a rejection.

This picture, to be fleshed out in various incarnations in what follows, is at bottom an empirical hypothesis whose primary claim is semantic: the idea is that mathematical thought and talk can be, and ought to be, interpreted as relativized to a framework which stakes out the boundaries of the quarantined talk, and that with this hypothesis in place we accrue a number of advantages (to be discussed). What goes on within the boundaries does not drain out the plainly literal province of speech and thought. My claim against this common core of fictionalist hypotheses is that it is crucially, and fatally, unable to capture either the semantic or the epistemic dimensions of mixed context inferences, because of its commitment to quarantine.

4.1. Staking out the territory

There are substantial advantages to a mathematical fictionalist picture, the most attractive, in the eyes of many, being that it disengages mathematics from uncomfortable ontological commitments. After all, the basic shape of the argument for realism about mathematical objects is that realism is required in order to make sense of mathematical truth; but if we are not committed to the *literal* truth of mathematical propositions because they are in quarantine, the requirement is undercut. In addition, a theory of truth, or of quasi-truth, is preserved in fictionalism by relativizing the truth-predicate to the context of a relevant body of fiction, and *a priori* knowledge is maintained without making reference to analyticity. And so the fictionalist seems to be able to offer non-trivial mathematical truth, knowledge independent of the course of sensory experience, and freedom from ontological profligacy.

A number of fictionalist avenues have been proposed. (1) The first is to interpret mathematical propositions as, so to speak, *nakedly* fictional: they are just plain false. This is, or at least once was, Hartry Field's suggestion, examined in § 1.2. above. (2) The second is to claim that the class of mathematical truths should be construed as truths relative to a particular theoretical language, but then to say that such truths have no reach outside the confines of that language. So argued Carnap, and Stephen Yablo has lately taken up that proposal. (3) A third view, also offered by Yablo among others, is to read mathematical propositions as truths-in-fiction, and again these 'quasi-truths' have no purchase on the world outside the fiction in question. (4) Yet another avenue is to construe apparent commitments to mathematical objects as what I will call 'didactic fictions', as convenient but actually non-commitment-making heuristic devices. (5) Finally, there is the suggestion (arising from work of Kendall Walton's in another context) that mathematical utterances can be interpreted as make-believe, and therefore as divorced from standard truth conditional analyses.

What these proposals have in common is that each suggests a way of putting our mathematical thought and talk in quarantine, by construing such thought and talk as fictional in one way or another, and therefore immune from assent. In what follows, I will argue that the common core of any fictionalist proposal, the 'epistemic quarantine' view, is *not* plausible, by arguing that the various attempts at non-literal interpretations of mathematical thought and talk fail.

4.2. Naked fictionalism

Fictionalism about mathematical entities is the thesis that despite their usefulness as expressive tools in the language of empirical science, and despite their empirical applications and (at least apparent) predictive power, mathematical existence claims should not be interpreted as true claims. Mathematical existence claims, on the fictionalist view, are close – though not quite analogous – to the idealized expressions of physical laws that invoke such nonexistent but theoretically useful 'entities' as frictionless planes and point masses. Physics students are told in the first week that *of course* there are no masses in no dimensions, but that it's best to treat masses as point masses for the purposes of, say, astrophysics, because for the sake of simplicity and beauty the laws governing gravitating bodies are best expressed when we treat those bodies as nondimensional, as having mass but no size.[14] The fictionalist claim is that just as we treat point masses as the subject matter of astrophysics, so we should treat numbers, functions, sets, groups and so on as the subject matter of mathematics: neither exists, though both are handy for explanatory and predictive purposes. The analogy is not exact because the mathematical explanation of physical features of the world is not analogous to the mathematical explanation of mathematical features of the world (even if we were to beg the question and assume that there *are* mathematical features of the world), and because there is no clear sense in which the mathematical platonist relies on some putative predictive power of mathematics for support of her platonism. But even given these caveats, the fictionalist objection cuts deep because it says that even if we suppose that commitment to mathematical entities is an indispensable element of our best theory of the physical world, still it need not follow either that such entities necessarily play an explanatory role in our theorizing or that mathematical existence claims are to be treated as true any more than existence claims about point masses are. Field (1980) poses the hard-nosed nominalist challenge:

> [W]hat good argument is there for regarding standard mathematics as a body of truths? The fact that standard mathematics is derived from an apparently consistent body of axioms isn't enough; the question is, why regard the axioms as *truths*, rather than as fictions that for a variety of reasons mathematicians have become interested in? (Field 1980: vii-viii)

[14] So it goes for theoretical physics. Obviously, engineering problems – like moon landings, trips to Mars and bridge building – had better not treat masses this way.

Consistency is clearly not the stamp of truth, nor is it grounds for ontological commitment: that much is true by anyone's lights.[15] Even if we suppose that Conan Doyle's stories of Sherlock Holmes are completely consistent, that's no reason for saying that there was such a person as Sherlock Holmes.

But Field's claim is that between truth and usefulness lies the great gulf between the true and the 'factually defective' (cf. Field 1994 & 2001b). Presumably this is not quite the same as the gulf between the true and the false since the 'factually defective' is that for whose truth we have no evidence. What is it, then, that would count as evidence for any *a priori* claim at all? In discussing specific claims to *a priori* knowledge – of which mathematical knowledge is a species – Field writes as follows:

> In virtue of what is it reasonable to use modus ponens on no evidence? The difficulty of providing an answer to this question is one of the main reasons that apriority has seemed so mysterious… [T]he proper question is, why value a methodology that allows the use of modus ponens on no evidence? Well, one needs some methodology, so the question can only be why favor this methodology over alternatives, and the answer will depend on what alternative methodologies are possible… The question then reduces to showing what is wrong with particular methodologies of each type. (Field 2001b: 371-372)

I take it that Field is asking what, given a particular methodology, falls *outside the reach* of that methodology: a method will be evaluated according to its successes and failures in

[15] Perhaps it would be better to say that *local* consistency is not the stamp of truth. It is certainly true that making up a consistent story involving unicorns, Sherlock Holmes and Prester John involves no commitment to any of the cast of characters. A local coherence theory of truth is no theory at all. Global consistency, on the other hand, is a more complicated issue. It may be that once all the facts are in – say at some Peircean end of inquiry – there will only be one coherent history and inventory of the world, and that that is the one that we will call 'true'. It is tempting to say that interpretations, multifarious as they are, can only extend as far as the facts will allow, but that claim depends on an interpreter-independent notion of 'fact' which in turn depends on a notion of truth. It seems any attempt at getting rid of the idea of correspondence truth, globally interpreted, is doomed.

the face of relevant phenomena, granted. The method at issue is to look at arguments for the truth of mathematical existence claims, and evidence for the arguments' truth or falsity is to be found (according at least to the Putnam-Quine indispensability argument) in their empirical applications. Mathematical existence claims make, by Field's own lights, testable claims, and we test them by evaluating their explanatory worth against their rivals'. My reply to Field is therefore that what is wrong with the fictionalist methodology is that it does not deliver mathematical truth even by its own lights because the fictionalist methodology cannot deliver an explanation of the empirical applications of mathematics, nor can fictionalism account for constraints of the kind I raised in discussing the impossibility of squaring the circle (in chapter 1). The following arguments will show, I think, that we cannot leave substantive questions of mathematical ontology behind when fictionalism crops up. The fictionalist denial of ontological commitment is defeated in two ways. First, (at least mathematical) existence claims are not made ontologically inert by the fact that there can be structurally equivalent yet incompatible theories of a phenomenon. Second, they are not made ontologically inert by the denial of correspondence truth, because fictionalist denials themselves suppose a background against which the denials may be adjudged, and that background is one in which existence claims are intelligible only in relation to whether or not they are true and available for inference.

4.3. Relativity to *L*-internal structure

But one might argue that existence claims are *not* intelligible. One might argue with Carnap (1955) that existence claims are at best only intelligible within a linguistic 'framework', and unintelligible from the point of view of any other linguistic framework.

One might so argue because a relativism of vocabulary is sometimes tempting, for example for epistemological reasons, especially if one has a demarcation criterion in mind between what is meaningful and what isn't. Carnap's view is roughly that from within the 'framework' of classical mathematics, it makes sense to ask whether there are one or more primes between 7 and 13, because in the framework there is an intelligible way of telling the right answers from the wrong ones, i.e. of saying that 'yes, and the number in question is 11' is a right answer while 'yes, and the numbers in question are 11 and 12' or 'no, there aren't' are wrong answers.

The question of the intelligibility of ontological questions depends for the most part on whether one thinks that there is a criterial distinction to be made between frameworks, or whether on the other hand there is no criterial distinction between frameworks, theories and languages. Cast another way, it depends on whether there are limits to the literal mode of declarative theorizing that can be carved off from the non-literal, metaphorical bits. If there are such limits, as Carnap (1950, 1956) and Yablo (1998) have argued, then it might be the case that ontological questions can only legitimately arise from within the bounds of the literal, and that questions raised from the outside are 'devoid of cognitive content' (Carnap 1950:212). The claim is equivalent to saying that utterances (theoretical sentences or propositions) only have truth-conditions when expressed in the vocabulary, and according to the well-defined rules, of some language *L* within which ontological commitments can be made and are truth-apt. Any expression lying outside the (stipulated?) sphere of *L* does not make ontological commitments or truth-apt claims because the truth conditional domain for those expressions is not defined. Thus, for Carnap, if we take *L* to be composed of the vocabulary, formation rules and derivation rules of number theory, it is intelligible to ask about properties of the distribution of primes – the question has cognitive content because by examining the rules of *L* we can arrive at an answer – but ontological questions raised from outside *L* have no content because they are not defined on any domain and are therefore just not answerable. Questions of whether there *really are* primes, if they are not questions about what *L* says there are, are senseless, or at best metaphorical.

My reply to this worry is twofold. First, Carnap's objection does not bar ontological inquiry into the status of the ontological commitments made within *L*, unless we accept the epistemic constraints (of verificationism and constructibility,[16] say) that are implicit premises in Carnap's picture and which would need a lot of defence in their own right. It is still intelligible to argue between, for example, *L*-platonism and *L*-nominalism if the dispute is over whether or not the claims made within *L* demand platonic truthmakers, and we reject verificationism as the epistemic constraint on arbitration. But even then, this platonist reply is rather feeble if Carnap is right that cognitive content is only to be had within *L*-specific theoretical domains, and so in the second part of my reply it falls to me to argue that we cannot make sense of a criterial divide between sub-domains of literal, truth-apt claims on the one hand, and supersets of claims that are either unintelligible or metaphorical on the other.

Let's turn to Quine:

> If there is no proper distinction between analytic and synthetic, then no basis at all remains for the contrast which Carnap urges between ontological statements and empirical statements of existence. Ontological questions then end up on a par with questions of natural science. (Quine 1951b: 134)

[16] Note that Carnap's version of constructivism is not of the usual Brouwer-Weyl-Dummett kind; Friedman (2001) illustrates Carnap's greater tolerance.

The idea here is that in Quine's reading of Carnap, the *L*-external questions are analytic and the *L*-internal questions are synthetic: this makes internal questions meaningful insofar as they are answerable and proposed answers are evaluable for truth, and it strips external ones of cognitive content since they are just uninformative logical truths. But since, by the arguments in Quine (1951a), the distinction between analytic and synthetic propositions is untenable, the internal/external distinction collapses: if one accepts Quine's argument in 'Two Dogmas', the sentences that we are prepared to call 'analytic' don't enjoy any special epistemic status beyond being at the furthest possible remove from directly observational sentences. Questions of the ontological status of putative entities are then *all* synthetic, and resolved by the usual Quinean method of detecting the ineliminable commitments of our best theories which make those theories come out true.

Yablo argues that the analogy between Carnap's external/internal divide and the analytic/synthetic divide that Quine wants to dissolve is a weak one, on the grounds that 'existence-claims of the kind Carnap would call analytic show no particular tendency to be external' (Yablo 1998: 235). It seems that Yablo is right. But rather than showing that Quine's criticism is mistaken, this still leaves Quine's criticism intact. Some existence claims, those made *L*-internally, are indeed legitimate and answerable by Carnap's lights (e.g. 'is there a prime between 7 and 13?') and it is indeed an analytic question whether *L*-internal mathematical properties of certain kinds are or are not instantiated, insofar as it is answerable *a priori* given the rules of *L*. But Quine's point is not to say that Carnap can't cope with *L*-internal existence claims; rather, it is that existence claims that are *L*-*ex*ternal according to Carnap *do* have cognitive content because the bounds of the synthetic are pushed right into the centre of the web of belief. There's nothing for '*L*-external questions' to be external to.

But, questions of interpretation aside, Yablo looks to make the case that something like Carnap's limits of intelligibility can be rescued if we take the contrast to be between theoretical utterances meant to be taken literally, and those that are metaphorical. Thus Yablo:

> [L]et us identify frameworks outright with practices of such and such a type, where it is independently obvious that to engage in these practices is not thereby to accept any particular doctrine.[17]
>
> Now, what is our usual word for an enterprise where sentences are put at the service of something other than their usual truth conditions, by people who may or may not believe them, in a disciplined but defeasible way? It seems to me that our usual word is 'make-believe game' or 'pretend game'. Make-believe games are the paradigm activities in which we 'assent' to sentences with little or no regard for their actual truth values. (Yablo 1998: 243)

In this light one could imagine the following bit of conversation:

> WISEGUY: So, are you saying that there are point masses?
>
> PHYSICIST: No, we're just using those as tools for simplicity's sake.
>
> WISEGUY: But doesn't your best theory quantify over them?
>
> PHYSICIST: Sure, but it's metaphorical quantification.

We're back to Field's slogan that a set of theoretical sentences need not be true to be good. But as Field himself acknowledges: 'what makes scientific methods better [than

[17] That is, let us admit that we can usefully talk of idealizations such as point masses for scientific purposes without being committed to there being any such thing, and call that kind of talk a part of the 'framework of physics'.

alternatives] … is that *they do* lead to more truth and less falsehood than those other methods' (Field 2001b: 383). In other words, a background of assumptions about truth is a necessary condition of judgements about whether or not certain methods, certain commitments, are better than others. Even if we were to say that ontological commitments are in some sense metaphorical, we evaluate the aptness of metaphors at least partly according to the insight we gain through them *with respect to their subject matter*. A metaphor is better or worse insofar as it is an apt description of its object, and so even if we were to suppose that ontological commitments are metaphorical, that does not lead us into fictionalism; quite the opposite: it reenforces the claim that there is something to which we are committed and which is well or poorly approximated by a metaphor used to pin it down. And supposing further that there is no set of necessary and sufficient conditions that will allow us to shave the metaphorical away from the literal – and it seems very unlikely that there are such conditions – the norm of truth remains undamaged. A theory of meaning, at least under one very plausible description, should aim to establish the right kind of pairing between a linguistic item and a non-linguistic item. It may be the case that for every linguistic item there is a number of possible candidates for the corresponding non-linguistic item (this is Quine's point about the indeterminacy of reference and is also the thrust of Russell (1923)), and it is certainly the case that for each non-linguistic item there are alternative descriptions expressible in one or more languages. After all, any state of affairs that can be described in one language can be described in another. Ultimately, the core constraint on a theory of meaning must be that it answers the question "how does expression E manage to mean what it does?", and this constraint is fulfilled when word-world pairs are settled in the right way, whatever that might be. In the house of linguistic items are many mansions, some literal, some metaphorical, some elliptical, and a legion of others. Philosophy itself is riddled with metaphor (for instance when a philosopher speaks of "full-blooded" or "thin" theories), as any discipline must be to the extent that it is expressed in a language at all. But the fact that metaphor has great expressive power should not obscure the fact that metaphors express something that's evaluated relationally: evaluated with respect to subject-matter. The subject-matter is not linguistic; delight in a finely tuned metaphor arises from the expressive power the metaphor has with respect to the bit of the world it captures.[18] Metaphors that sound well but shed no new light on a bit of the world are empty. Moreover, it may well be that the use of metaphor is a fundamental element of human cognitive engagement with the world, as Lakoff & Johnson (1980) argue, but this use is precisely *engagement* with a world whose features are liable to being well or poorly represented in a variety of ways, one of which is metaphorical. Thus, to claim that some bit of language is metaphorical is not to free it from an ontological tether. Indeed the possibility of metaphor itself could be used to argue for ontological commitments since metaphor relies on better or worse approximations to its object. In any event, the argument that metaphorical uses of language are free of ontological import is thin at best, since metaphor is one among the many uses of language that can be evaluated with respect to truth, with respect to a constitutive and necessary evaluability in the face of non-linguistic items.

[18] It's interesting to think about musical "language" in this context. Some music is downright sarcastic (Shostakovich was a master of musical sarcasm), some is funny (there are many funny passages in Mozart), some is cruel (e.g. in Berg's *Wozzek* or Reimann's *Lear*). The sarcasm, cruelty and so on of the music rely on representing bits of the world through metaphors in musical "language", musical approximations of their subject-matter. For instance, if I have a criticism to make of some of Britten's operas, it is that the music is sometimes *too* literal (e.g. in *The Turn of the Screw*), but at the same time I recognize that the sheer literalism of the music serves a crucial expressive purpose in the musical "argument". Musical metaphor is (at least) *treated* as evaluable in much the same way as the strictly linguistic. We cannot account for this in terms of structural similarity between musical metaphors and their objects; Britten, for instance, used A-natural on a horn to express innocence and childish wide-eyed wonder, but there's nothing structurally similar between that note of music and the "corresponding" mental state.

To reconstruct the fictionalists' claims about mathematics in more detail, and flesh out the critical points above, we need to make some further distinctions.

4.4. Standard overt fictions

First, there is fiction outright, such as we find in novels, fairytales, and so on, for which it is the case that some propositions are true-in-the-fiction, some are false-in-the-fiction, and many no doubt are indeterminate. Thus,

(3.4.1.) Humbert Humbert drove Lolita across the United States

is true-in-the-*Lolita*-fiction, and

(3.4.2.) Humbert Humbert is content and at peace

is false-in-the-*Lolita*-fiction, and no doubt

(3.4.3.) Humbert Humbert wore blue trousers on the first Thursday after meeting Clare Quilty

is forever indeterminate-in-the-*Lolita*-fiction. We can utter true and false things about the explicit contents of the fictional world of an outright fiction, and must remain forever agnostic about the truth value of anything not made explicit in the fiction or deductively following from it. In the cases above there is no question of ontological commitment to Humbert Humbert, his mental states, Lolita, Quilty, trousers or Thursdays. All of these claims are tacitly subject to the intensional operator 'In-the-*Lolita*-fiction, ___ '.

As I have said, the great attraction of fictionalism as a general thesis in mathematical ontology is that it allows for a certain kind of realism about the truth values of mathematical claims while dodging the apparent commitments of those claims by interpreting them on an analogy with fictional discourse. In particular, if we take '_____ is true-in-a-fiction' as an intensional operator akin to 'X believes that _____' (as

Lamarque & Olsen (1994) suggest), we know that certain inferences, particularly existential generalizations, are thereby blocked. The inference ticket in intensional contexts of this kind won't allow a ride from a claim that's true-in-the-fiction to any extra-fictional reality. This offers ontological economy, and also responds neatly to Benacerraf's dilemma since it both allows for mathematical truth and for the opportunity of acquiring and extending mathematical knowledge. On the epistemic front, the idea is that we make mathematics up and are then free to 'discover' new features of our created world through inferences of the appropriate kinds made on the make-believe objects and the relations between them.

There are, however, at least two problems with this kind of elementary fictionalist view of mathematics. The first is the problem of applications to the physical world, which is the launching pad of indispensability arguments in the first place. It would be thoroughly absurd to claim that aircraft are supported in the air by passengers' and crews' collective faith in make-believe entities. The second point on which the analogy breaks down is the question of coherence: internal contradiction is not (or at least not *necessarily*) damaging to a fiction, while it is lethal to a body of mathematical theory. If we read *The Count of Monte-Cristo* closely enough – so I'm told – we discover that in different parts of the fiction the Count is reported to be in different cities at the very same time. In other words it is true-in-the-fiction that P and not-P. Is this damaging to the book as a work of fiction? It doesn't seem so to me, partly because the contradictions are very hard to notice at all, and partly because even if one does notice them, the strict ordering of a content reported by a storyteller as though it were factual is not what makes a fiction interesting. It need make no difference to our appreciation of a fiction that it is internally incoherent. (Of course, sometimes it does: if a murder mystery offers a solution which, from the information given in the story, is impossible, we rightly feel cheated; but the point is just that this *need* not be the case.) To take another example, when the first British edition of Melville's *Moby-Dick* was published, it was printed without the Epilogue that explains how Ishmael survived the wreck of the 'Pequod'. Without the Epilogue it makes no sense that Ishmael could be the first-personal narrator of the tale. But *Moby-Dick* is no less a masterpiece without its Epilogue; in fact the Epilogue strains belief anyway. It smacks of the "and I alone am escaped to tell the tale" sort of dodge and rather spoils the climactic drama of the book. In the mathematical case, on the other hand, and accepting the principle of *ex falso quodlibet*, anything at all follows from a contradiction and therefore in a body of mathematical theory that asserts contradictories as true we can prove anything at all, which simply guts the notion of truth for that body of theory. While possibly harmless in fiction, contradiction is lethal in mathematics.

To return to the first point of criticism: it cannot be that, for the fictionalist, mathematics and its apparent quantification is fictional in the sense of an outright fictional quantification, because mixed contexts in which mathematical vocabulary is deployed are confirmably true of the empirical world, as when one claims:

> (3.4.4.) Baking soda ($NaHCO_3$) is often used as an antacid. It neutralizes excess hydrochloric acid (HCl) secreted by the stomach. Milk of magnesia, or $Mg(OH)_2$, is also used as an antacid. Which is the more effective per gram? To answer this question, we must determine the amount of HCl neutralized per gram of $NaHCO_3$ and per gram of $Mg(OH)_2$. Upon calculation, it turns out that 1 g of the

former neutralizes 1.19 x 10^{-2} mol HCl while 1 g of the latter will neutralize 3.42 x 10^{-2} of HCl. Thus, $Mg(OH)_2$ is a better antacid per gram than $NaHCO_3$.[19]

If the fictionalist were an *outright* fictionalist about (3.4.4.), the possibility of quantitative science would be open to serious doubt, and that kind of doubt is refuted by the observation of empirical success. But there are other avenues open to the fictionalist.

4.5. Didactic fiction

In addition to standard works of overt fiction, there are particular 'entities' (in fact, pretend entities) and stories involving them that we call *useful* fictions. Examples of these range from the idealizations so often put to use in science (proteins isolated from all other molecules in experimental settings, etc.) to morality plays, propaganda and parables. What groups these together is their purpose: in one way or another, they are all intended to be didactic, supplying idealizations for the purpose of transmitting information without regard to much contextual subtlety. I have in mind examples such as the Soviets' Stakhanov, the figure of Everyman in medieval theatre or Rawls's original position. They are much like point-masses: really, *nothing* is true of them, but they serve (or are thought to serve) a useful didactic purpose.

Quantification over useful fictions is not commitment-making either. But in these cases, its purpose is to express propositions that are true or false independently of the world of a canonical fiction in which they figure, while in cases of outright fiction the evaluation for truth or falsity of such sentences as 3.4.1. – 3.4.3. can only be made by reference to the fiction in which the entities over which they quantify figure. Useful fictions are different from outright fictions because their explicit purpose is to express

[19] Adapted from Zumdahl (1998: 71).

truths about the *actual* world: truths about the orbits of the actual planets or about how a good citizen should behave. Useful fictions stand as non-commitment-making *Ersätze* in the propositions in which they figure, *Ersätze* for the multiple possible substitution-instances that anchor the propositions to their applications. As Elgin (2004) points out, these *Ersätze* – roughly what she calls 'exemplifications' – are cognitively useful: we can retain a belief that we explicitly know is false because of its cognitive utility in smoothing out the rough edges of experience (as when we draw a smooth continuous curve through unevenly distributed data points on a graph), or illuminating some aspect of human relationships (as when we imagine the conditions that ought to govern a situation of rational bargaining). These are fairly clear cases of epistemic quarantine. So in the respect of fiction-as-*Ersatz*, the view is epistemically attractive: our cognitive functioning need not be tethered utterly to the literal truth of every belief in the cognitive web. But there is a difference here between the epistemically attractive and the metaphysically tenable. Suppose we have a sentence S in which figures an *Ersatz*: this sentence can well be retained in the cognitive docket for storing what we acknowledge is handy (with respect to certain purposes or functions), but false. Sentence S is not commitment-making with respect to the *Ersatz* it invokes. But it is here that the putative analogy between *Ersätze* and mathematical objects figuring in S breaks down: a mathematical S is not to be stored in the handy-but-false docket. It belongs with the handy-and-true in the case of a mixed-context belief, or with the just-plain-true in the case of a 'pure' mathematical belief (one without applications). Those beliefs, in turn, are indeed commitment-making. So, if we read Yablo's claim as the claim that mathematical objects can be treated as *Ersätze*, it still seems that commitment cannot be done away with.

4.6. Make-believe mathematics

Having discussed outright fiction and useful fiction, I turn now to some comments on the possible interpretation of mathematical talk as non-literal (or metaphorical or figurative) talk, which is really the primary thrust of Yablo (1998, 2000, 2001, 2002, 2005). Standard uses of metaphors are familiar:

(3.6.1.) Old Mother Hubbard keeps her kids on a short leash.

(3.6.2.) You are my working week and my Sunday rest.

(3.6.3.) *L'enfer, c'est les autres.*

There's no ontological commitment to be found here to leashes, hell or people numerically identical to weeks. The commitment-making logical form of these sentences is defused by their uses in contexts that are quite obviously non-literal; context insulates from commitment in these cases, and the logical form that we generally use as Quinean commitment-detectors is inert here. An important strand of fictionalism[20] contends that we should read a body of mixed-context mathematical talk (as in (3.4.4.) above, for example) as (a) metaphorical, (b) truth-apt, and (c) non-commitment-making, just as (3.6.1.) is non-literal (the kids aren't tied up), potentially true (is Mother Hubbard a strict disciplinarian?) and commitment-free (there's no such object as the leash in question). The applicability of numerical expressions in such talk as (3.4.4.) is accounted for by interpreting those expressions as prop-oriented in the context of a background collection of conventions of pretense, in the manner of Walton (1990). Here's an example of what this means:

(3.6.4.) I have slain the dragon with my sword!

[20] Of course I have in mind fictionalism in the philosophy of mathematics, but the metaphorical reading of ontologically problematic sentences is prevalent in strands of fictionalism about other ranges of talk as well (e.g. in moral or possible-worlds theory). See the papers collected in Kalderon (2005).

This is uttered in a context of children's play where 'the dragon' actually refers to a tree that participants in the game have agreed will serve as dragon-prop, and the 'swords' are sticks. To 'slay the dragon' is to swat the tree with a stick. Now, (3.6.4.) is used to express a propositional content which can be evaluated for truth or falsity with respect both to the arrangement of things in the actual world and the preestablished conventions of what will serve as props. The tree props up the pretend dragon, the sticks prop up pretend swords; but of course there are neither dragons nor swords. Commitment is in quarantine.

But consider now:

(3.6.5.) The power set $P(X)$ of a set X is the collection of all subsets of X.

This is a definitional sentence of set theory, expressing a proposition apparently committed to sets and power sets. Can (3.6.5.) legitimately be interpreted either as metaphorical in the sense in which (3.4.1.) – (3.4.3.) are metaphorical, or as pretense with respect to background conventions in the manner of (3.6.4.)?

Yablo makes the following point:

> Question: what makes real content [of what is uttered] real? Answer #1: It concerns *real things,* for instance, moons as opposed to numbers. Call this the *objectual reality* of real content. Answer #2: It is *really asserted*. Call this the *assertional reality* of real content.
>
> One way of putting the moral … is that objectual and assertional reality can come apart. Suppose that I as a nominalist declare the number of primes to be infinite. Assertional reality is not lacking. There is something that I am really saying, as opposed to just pretending to say. But there is no real content in the

> objectual sense. I am talking about the numbers as they are supposed to be imagined, not as they really are. (Yablo 2001: 85)

Drawing some analogies between modes of mathematical talk and figurative modes, Yablo gives examples such as "prime numbers are mostly odd" ≈ "apron-strings are short" and "the # of F's = the # of G's iff there are as many F's as G's" ≈ "my bottom line is the same as yours iff both of us are prepared to settle for such-and-such and neither is prepared to settle for anything less". Yablo writes further:

> [T]he [figurative] vocabulary's utility [for the purpose of expressing our interests and concerns] *does not depend* on conceiving of its referential-looking elements as truly standing for things. Those, if any, who take bottom lines and numbers dead seriously derive the exact same expressive benefit from them as those who adamantly deny their existence. And both of these groups derive the exact same expressive benefit as those who never gave the matter the slightest thought. (Yablo 2001: 87)

This is where assertional and objectual reality of expressed content come apart, on his view. If it is possible to establish that assertional reality, figuratively expressed or invoked, is sufficient for the interpretation of mathematical thought and talk, then there can be no motivation to drive home the platonist line. If we judge philosophical theories by their ability to explain relevant phenomena and solve puzzles better than their rivals, as in fact we should, then by Yablo's lights there can be no reason to prefer the elaborate machinery of platonistic objectual reality over the relatively simpler machinery of assertional reality. Yablo, I take it, is arguing *not* that nominalism is preferable on ontological grounds, but rather that since sticking to assertional reality is enough to give us all the explanatory weight we need.

My claim is that mathematical talk cannot be so interpreted. Here is the objection. From the definition of the power set, and the examination of some (finite or infinite) sets, we quickly see that it follows that for some set of given cardinality, its power set will have greater cardinality. For set Y, whose cardinality we express as $|Y|$, we know that $|P(Y)| > |Y|$. Consider the set of stringed instruments. There are certainly more duos, trios, quartets and so on of stringed instruments than there are bare instruments. Even confining our attention to sets of material objects with finite cardinality, the inference from (3.6.5.) to $|P(\textit{stringed instruments})| > |(\textit{stringed instruments})|$ is licensed. But this inference-ticket is not validated if we suppose that (3.6.5.) is a metaphor, a bit of let's-pretend confined to a level of assertional reality. To have truth-preserving inference, as we do, is to set out on an inferential road that begins with what we assent to. From a metaphor we can infer *nothing*, unless that metaphor does duty for a true proposition. This point finds additional support in the argument from constraint and semantic descent in the first chapter.

In the end, both fictionalism about, and metaphorical interpretations of, ontological commitments in mathematics cannot be coherently formulated unless the regulative question of whether or not those commitments are true is smuggled into a premise, which in turn undermines the fictionalist argument. The upshot is that we need availability for inference, which is cut off by the 'quarantine' proposal. The norm of truth still remains intact.

4.7. Tacit modal disclaimers

The modal fictionalist strategy, finally, is the proposal that mathematical thought and talk be interpreted as prefaced by a tacit modal disclaimer: 'There are possible worlds in which the following is the case, and everything else proceeds as usual. The quantification in the disclaimer is interpreted as ranging over possible worlds construed as fictions, and therefore the mathematical speaker (or writer or thinker) is insulated from the commitments to possible worlds and their contents that would otherwise follow from the disclaimer realistically interpreted *à la* Lewis (1986). The disclaimer could, say, be printed on a bookplate and glued to the inside cover of all published works of mathematics, and apparent objects of commitment would be in quarantine in a fictional limbo. Novels do not need such a bookplate because (generally) readers already jump in with the expectation that they are being told a yarn.

The trouble with this modal fictionalist strategy is one of deciding where to tack on the disclaimer. As I have pointed out, it is disingenuous to pretend to pry apart the mathematical and physical or empirical components of work in quantum mechanics. It does seem, though, that this same objection does *not* apply to much empirically useful mathematics because it is acknowledged that the mathematics represents an idealization that can only approximate what really goes on; think, for instance, of a text in population genetics which describes the rates of diffusion and propagation of a favourable gene, or the mathematics of stochastic processes put to use in finance and notorious for being so idealized that it can't predict anything. In population genetics and finance and the like, it is clear that the *applications* for which the mathematics is used are so messy that the formulas so put to use are, empirically, largely false: this is indeed to show that the mathematical idealizations are useful fictions. In cases of these kinds the disclaimer would be salutary: 'don't expect the *actual* world to behave this way'. But the trouble is that the modal fictionalist strategy thoroughly insulates the actual world from the non-actual possible worlds in which the mathematical claims are true of the objects they purport to be true of, and this goes for the most recondite reaches of unapplied mathematics as well as for the mathematically idealized description of the empirical world. It cannot be, on this view, that all the mathematics we know is true of the actual world, because that would not be fictionalism (and the point is to dodge commitment); but it must also be the case that *none* of the mathematics we know is true of the actual world, because there is no criterial way of deciding what needs to have the disclaimer attached and what doesn't. The fictionalist thesis only offers a postponement of the ontological and epistemic issues that we confront in exploring the datum that is mathematics. Pressing the thesis further brings us right back into the very issues that motivate a philosophical examination of the right semantics for mathematics, and its ontological consequences.

I have argued that by examining the ontological commitments of our mathematical talk, we discover commitments to mathematical entities which are indispensable in the deployment of natural science; we should extend that indispensability beyond the sphere of the empirically applicable; we are genuinely committed to the norm of truth as an ideal of inquiry in mathematics as much as in any of the mind's pursuits; and deflationism through fiction is ultimately unsuccessful. Berkeley wrote that 'whether you argue in language or symbols, the rules of right reason are the same' (Berkeley 1734: §15): much the same spirit has animated my discussion here. Right reason is inevitably governed by the norm of truth, and that norm entails commitment.

5

Structure and mathematical knowledge

It is usually better to present a problem that is too big and then reduce it, than to start with the problem too small. For in that case, we may never be able to expand it enough.
-- Richard Levins, 'Strategies of Abstraction'

5.0. Introduction: genesis and justification

The story so far is this. First, constraint and the resources indispensable for explaining it jointly entail commitment to mathematical objects. Second, Frege shows that mathematics cannot be construed as an uninterpreted formal system: meaning matters. Third, the appropriate theory of meaning is one on which mathematical thought and talk is truth-apt and literal, not figurative in any of the senses assessed in the previous chapter. Now, in this chapter, I offer a modest, and entirely programmatic, proposal about how mathematical knowledge might be accounted for. I do not claim that what I say here settles the issue, nor that it provides a general account of *a priori* knowledge, nor even that it is a descriptive account of how mathematical knowledge actually is achieved in practice. The virtues I claim for the proposal are that it at least sketches plausible avenues of inquiry, that it is consistent with the foregoing metaphysical and semantic picture I have developed in previous chapters, and that it is able to cope with some familiar puzzles.

I begin by dividing the broad question of mathematical knowledge into two, which I will call the genetic and the justificatory questions. Consider the familiar schema:

S knows that *p*

in which 'S' stands for the knowing subject and '*p*' for a proposition. In a very broad range of cases, the process whereby one comes to have the belief that *p* is the very same process that justifies one's knowledge that *p*. Such is the case when '*p*' stands for an empirical or sensory proposition, for instance: in the case of 'S knows that her mug is full of coffee', the account of how the belief that her mug is full of coffee is generated is also the account of how that belief is justified, and thus transformed from mere belief into knowledge. The genetic story of how she comes to have the belief at all, and the justificatory story about why she is entitled to that belief, are the same: one process covers both. The process in question here will invoke the proper functioning and reliability of S's perceptual apparatus, her causally describable contact with the mug and its coffee, suitable lighting conditions, etc. To put it another way, the question of whether her belief gets to count as a genuine piece of knowledge turns on the story that's told about how that belief was generated: justification will emerge out of a suitable process of genesis. If the generative process is of some appropriate kind – generally one that is spelled out in causal terms – then the belief is *thereby* justified and is secured as knowledge. The task of epistemology in cases of these sorts is primarily that of spelling out the generative processes that get to count as justificatory: direct causal acquaintance, testimony from reliable sources supplemented by an account of reliability, strategies of evidence-probing in laboratories or in law courts, and so on. The fundamental thesis in this province of epistemological work is that what causes beliefs is also, and *in virtue of causing them*, potentially their justification, as long as the process is appropriate. In this vein I have in mind such important work as Goldman's *Knowledge in a Social World* (1999), Coady's *Testimony* (1992), Longino's *Science as Social Knowledge* (1990) and *The Fate of Knowledge* (2002), Solomon's *Social Empiricism* (2001), and many others (and here I cite only a little sampling of books in a very rich field). The matter of the justification of a piece of knowledge is generally formulated as a matter of fathoming its origins, examining whether those origins are suitably knowledge-productive, and exploring how the relevant knowledge-productive practices might be shored up in order to secure genuinely warranted beliefs.

There is also a substantial strand of the epistemological literature that examines specifically *a priori* knowledge from this same genetic vantage point. Historically, John Stuart Mill presents the most fully worked-out account; I will take Philip Kitcher (1983, 2001) as the principal contemporary representative. He expresses his criterion for a piece of knowledge to be *a priori* as follows:

> *X* knows a priori that *p* iff *X* knows that *p* and *X*'s knowledge that *p* was produced by a process that is an a priori warrant for *p*.
>
> α is an a priori warrant for *X*'s belief that *p* just in case α is a process such that for any sequence of experiences sufficiently rich for *X* [to acquire the concepts invoked in] *p*:
>
> (a) some process of the same type could produce in *X* a belief that *p*;
>
> (b) if a process of the same type were to produce in *X* a belief that *p*, then it would warrant *X* in believing that *p*;
>
> (c) if a process of the same type were to produce in *X* a belief that *p*, then *p*. (Kitcher 2001: 67)[21]

Note, in the first sentence of this passage, the melding of the genetic and the justificatory: causal origin *is identical to* warrant, subject to the additional conditions (a) through (c) on the warrant-conferring process α, on Kitcher's account. Mathematical knowledge then emerges as a body of knowledge that is not *a priori*. How come?

[21] This passage rehearses the very same condition that he set out in Kitcher (1983: 24), where the specific body of knowledge at issue was mathematics. The bit in square brackets is my gloss to clarify Kitcher's meaning. See also Parsons (1986: 130-131). In the quotation I keep Kitcher's convention of not italicizing 'a priori'.

In *The Nature of Mathematical Knowledge*, Kitcher rightly stresses the social dimension of mathematical knowledge-producing practices. A mathematician working by herself in her garden shed may have great confidence in the difficult proof of a not-very-obvious theorem that she has finally managed to produce, but her degree of epistemic confidence will inevitably be increased when she shows it to other competent people and they are brought, by her proof, to affirm the result (see clauses (a) and (b) above, and substitute 'other competent people' for '*X*'). Confirmation by others of the adequacy of one's proof process is a significant barrier to scepticism, and ratchets up the degree of warrant for this theorem *p* (make the same substitutions in clause (c)). It is moreover the case that a vast amount of mathematical knowledge is held – even by mathematicians themselves – on the authority of others. The social dimension is not incidental; rather, it is crucially and properly *epistemic* in the degree that it serves as a defeater of scepticism, a block to the opportunity for doubt. Do these features of α, the warranting process, meet the conditions set out for α to be authentically *a priori* warrant? Not so, according to Kitcher. I borrow the following gloss from Parsons:

> [In Kitcher's view] an *a priori* warrant for *X*'s belief that *p* is a process such that, no matter what *X*'s experience, provided that it suffices to acquire the concepts in *p*, some process of the same type could produce the belief that *p*, and in the case of any such process the belief would be true and warranted.

Furthermore, in the light of the foregoing observations about the actual bases of mathematical practice:

> It follows that nearly all mathematical knowledge is causally dependent on experience which goes beyond the experience necessary to acquire the concepts involved. (Parsons 1986: 130-131).

Mathematics, then, is not a body of knowledge achieved *a priori* because the warranting processes α rely on experiences that outstrip the requirements of concept acquisition for the concepts involved in just about any particular mathematical proposition of some minimal complexity. The idea, I take it, is that the warrant that one has for nearly any mathematical belief is widely distributed over, and reliant on, a web that over-reaches one's own concept-acquiring capacities, a web whose nodes are essentially causally linked to a range of experience that one could not possibly have.

The requirements for apriority that Kitcher lays down are stern indeed, subsequently leading him into a very sophisticated version of empiricism about mathematics. It is an attractive view at least in so far as it is consonant with a naturalistic, causal theory of knowledge; but I do not propose to discuss the specifics of Kitcher's version. Rather, what I want to challenge is the epistemic picture that lies upstream from the conclusion, which conflates genesis and justification.[22] I take Kitcher's theses merely as illustrative of that conflation, which I want to take apart. The taking apart of genesis and justification, and the consistency of this with the ontological realism I have defended previously, are all I aim to defend as positive theses in this chapter. My claim will be that, in the case of mathematics, the cause of beliefs comes apart from the justification of those same beliefs, and that it is more fruitful to think of these questions of genesis and justification separately. The genetic question asks 'what is the causal origin of our having any mathematical beliefs at all?', while the justificatory question asks 'why is it rational to hold the mathematical beliefs that we do hold?'. As I said above, to answer the first is, in very many important cases of knowledge, to answer the second in the same breath; but not in the case of mathematical knowledge. What, then, motivates such a claim?

[22] Kitcher is perfectly happy to acknowledge this conflation, and he repeatedly calls his own approach 'psychologistic'.

The motivation comes partly from reflecting on what is required in order to specify the content of a mathematical – more particularly, arithmetical – belief. On the basis of the first two chapters above, and the argument against fictionalism from the third, mathematical thought emerges as thought *de re*, and not *de dicto*: it is object-dependent, and thus its content will require an externalist specification making appeal not just to other components of the subject's conceptual apparatus but also to content-fixing extra-mental resources, elements of the world. But to specify the content is not always, and in particular not in the arithmetical case, to supply that content's warrant. Consider an instance of a veridical demonstrative perceptual belief: 'that's red'. What causes the belief is the very same as what justifies it, and what specifies the content of the belief will invoke its cause – some red thing in a suitable causal relation to the knower – in addition to whatever is required to possess the concept RED. What structuralism can offer us (as I will also suggest below) is a very good picture of the content-fixing *causes* of at least elementary arithmetical thoughts. In essence, genetic structuralism holds that the content of elementary arithmetical beliefs is fixed (say in children learning addition) in part by structures, perceived as commonality of pattern and relationships among the elements composing pattern.[23] I think this is an adequate genetic story about the fixation of arithmetical belief. But should it then also be the case that what causes our (elementary, at least) beliefs can also justify them, as is so often the case in the general run of belief? I will argue that to blend together the causes and the justifications for mathematical belief is to impoverish the resources of justification unnecessarily. The question of 'why is this belief rational?' will not be answered by appealing to a causal relation to some portion of structure. Richer resources are needed.

[23] I say "in part" because there are also recurring relational functions to master.

To pick an elementary example, take a child's belief that $1 + 1 = 2$. The child is plausibly brought to believe this by being shown matchsticks: one matchstick concatenated with another yields two matchsticks, and from this the child abstracts the generalization that wherever there is a pattern of one thing and another, there are consequently two things, discerning this pattern in the two wheels of her bike, the two eyes in her head, the two bees on the peach, and so on. The cognitive capacity of pattern-recognition does this work. But that cognitive capacity does not supply the rigour of warrant that arithmetic demands. The *proof* that $1 + 1 = 2$, given from the Peano axioms and invoking the recursive successor operation, is both a lot less obvious than the proposition so proved, and certainly not causally responsible for the child's belief. Though her belief is both true and partially, inductively warranted by her perception of common pattern, the belief that $1 + 1 = 2$ is not *mathematically* justified by her cognitive capacity combined with perceived structure. We need mathematical justifications to justify mathematical propositions.

In the next three sections I will discuss versions of structuralism in turn, with a view to defending *in re* structuralism as a plausible theoretical support for the genetic externalist thesis in epistemology, a thesis about how it is that mathematical beliefs may be acquired in the first place. This will serve as the logic to underwrite my account of genesis. Justification, we shall see, is another matter; my only substantive claim here is that it is fruitful to separate these questions in order to capture the requisite features of mathematical epistemology.

5.1. Epistemic structuralism

The thesis of epistemic structuralism is that ultimately mathematical knowledge is acquired thanks to a psychological faculty of pattern recognition which allows human knowers to abstract general, multiply-modellable structure from the gradual perceptual experience of patterns repeated over and over in the environment. Rows of cars parked in a lot, the buttons down the front of a shirt and the signs '1, 2, 3, 4, 5, …' or '|, ||, |||, ||||, …' written up on a blackboard are perceived as sharing a common structure defined by relational positions: there's a first, a second, a third and so on, and the learner comes to recognize that what these patterns have in common is an ordering of abstract slots that can be filled by almost anything. The root of mathematical knowledge is the ability to abstract the general concepts of relational mathematical structure from the particular experiences of objects ordered in recurring patterns. The notion is that elementary mathematical beliefs are acquired through the perception of structure and order, of patterns, in the world. This is a cognitive ability that explains why elementary arithmetic is so obvious. The proposition that 2+2=4 is very obvious indeed, much more so than any formal justification (proof) of that proposition could ever be.

It may well be that the faculty of pattern recognition is largely innate; it may also be that it is partly due to human beings' enculturation or initiation into a body of simple practices that are wont to make certain kinds of patterns salient. Whatever the balance of factors – innateness or initiation – there is no fundamental tension between the two. Human minds bathe in structured patterns, so to speak; and just as an accomplished swimmer needs both a certain kind of organism and a lot of initiation to be able to swim well, the result is that facility with elementary mathematical beliefs becomes as much a part of unconscious second nature as a good swimmer's technique.

An epistemological theory for mathematics, we have seen, needs to answer two questions. The first is genetic: how does a bit of knowledge come to be acquired at all? What is its road into the mind? The second is justificatory: why is it reasonable, or rational, to believe as we do? A structuralist story about pattern-recognition is a genetic story; there have been those, such as Hilbert, who looked to this kind of structuralism for justifications as well, but, as I argue elsewhere in this essay, this cannot be the case. The justification of our mathematical beliefs lies partly in the Fregean arguments that emerge from the semantic analysis offered in chapter 2. A further argument establishes more directly that epistemic structuralism can generate but not justify. The argument is just this: perception can only deliver knowledge of contingent truths; be the perceived regularities ever so regular, they still fall within the causal order and whatever we express about them is bound to be contingently true if it's true at all.[24] But epistemic structuralism *is* adequate with respect to the genetic problem.

Let me flesh this out a bit. My being able to count, say, four shapes on the board, and the attendant belief that there are four shapes on the board, is an ability that I can only have if I am prepared to say that since there are four shapes, there are at least three and no more than four, and at least two, and so on. The objection is that since my ability to count implies a mastery of the concept of how many *x*s there are, and mastery of that concept is constituted in part by mastery of the concept's material inferential consequences, then in having learned to count four shapes on the board I must also have logically prior, though tacit, mastery of 'if there are *n* things, then there must be at least *n* – 1 things.' The objection is not to be dismissed on the grounds that I can have a

[24] Here one might object that one can indeed just *see* impossibilities, *a posteriori*, as when one perceives two entwined solid three-dimensional rings and *sees* that it is impossible to separate them without breaking them. But this objection, it seems to me, is mistaken, because a claim of impossibility must be the conclusion of an inference which will invoke premises about metaphysical necessities.

concept of alcohol without having the concept of, say, CH_3OH, because being CH_3OH is not an inferential consequence of being alcohol (though from Kripke (1980) we have learned that it is a metaphysical consequence), whereas there being at least three things *is* a material inferential consequence of there being four things. This objection rightly supposes that we need to master some conceptual and perceptual-recognitional capacities, in addition to logic, in order to achieve mathematical knowledge. But then, says the objector, this seems to impart a deplorable flavour of contingency to mathematical beliefs: if their acquisition depends on perceptual-recognitional faculties (the ability to individuate objects, for instance), then if our perceptual faculties were set up differently, we would have entirely different mathematical beliefs. Suppose we were beings inhabiting an entirely fluid world, with nothing corresponding to what we call 'discrete objects'; mathematics there would be unimaginably different from what it is here, and surely this is a consequence to be avoided.

Indeed; but we need to be sensitive to the distinctive tasks of genesis and justification. The genetic structuralism sketched above is not a justificatory theory meant to account for mathematical necessity and apriority. We need a foot in the door to get mathematical knowledge started; there can be many ways in. Justification takes place from the inside; the genetic story just offers a route by which elementary arithmetic may enter the mind, a route consonant with causal requirements. Let's call the picture I have proposed a genetic externalism. It is externalist to the extent that it requires perceptual impetus from the mind-independent world, equipped with its patterns, to trigger the genesis of mathematical beliefs in subjects like us.

It is worth discussing genetic externalism in relation to another externalist epistemology: Gödel's. On Gödel's view we have mathematical knowledge thanks to some faculty of mathematical perception, akin to sensory perception but not reliant on objects available to the five senses as conceived of in the traditional classification. The idea here is that since we have mathematical knowledge, we must have a mode of acquiring it. We know, for example, that most of us human beings have eyesight, and we demonstrate that we have the faculty of sight in certain ways (e.g. dodging solid objects, admiring paintings, etc.). That's a good argument for saying of S that she has sight. In the same way, a person's knowledge that 2 + 2 = 4 is a good argument for the claim that that person has a faculty of mathematical perception of abstract objects that delivers mathematical truths. I raise this possibility here to note some similarities and differences. Gödel's epistemology seems to demand a *causal* (since it's perceptual) relation to abstract objects without saying anything positive about what the nature of that relationship might be. There could, perhaps, be a Gödelian analysis of causation. But a yet more fundamental objection to Gödelian epistemology raises the spectre of circularity: Sellars apparently called the Gödelian view the "Canadian Mountie" picture because it requires immediate apprehension (the story is reported in Bonevac (1982: 9)), and Sellars's comment is accurate to the extent that it underlines Gödel's reliance on self-evidence. But when we ask 'to whom is proposition *p* self-evident?', the answer is presumably that *p* is self-evident to any rational creature, where 'rational creature' is defined as any creature who obeys such-and-such laws, one of which is the very law *p* that is meant to be self-evident. Though this circularity may not be vicious, it doesn't have the ring of benign retrograde confirmation that, for example, we can get from assuming *p* and checking its consequences; so I think that agnosticism is the most appropriate attitude to adopt toward Gödelian epistemology. On the other hand, genetic externalism with respect to structure as a road to the acquisition of elementary beliefs, without the claim that the beliefs are justified by their causal origins, offers the advantages of Gödelian epistemology without the somewhat mystery-invoking element.

In the next two sections I evaluate the epistemic commitments of two kinds of structuralism along the genetic line, and argue that *in re* structuralism squares better with the picture given above; my attention is confined to its *epistemic* attractions. Both the *in re* and *ante rem* theses are comprehensive philosophies of mathematics, addressing ontology, semantics and knowledge; in what follows my purpose is to address these only in their epistemic dimensions.

5.2. ***Ante rem* structuralism.**

The structuralist thesis, on both the ontological and epistemic fronts, is born of an argument from the theory of reference that has been around for a long time. Structuralisms *ante rem* and *in re* tend to emphasize different aspects of the argument: the former lean on the thesis of ontological relativity to a background language or theory in the manner of Carnap and Quine, and the latter prop up their position on a quite radical version of the causal theory of reference. I will discuss the first in this section and second in the next.

The ontological relativism of the *ante rem* structuralist goes like this. (In this section it is principally Shapiro (1997) that I have in mind.) Reference is a process of reference-in-a-language. The pairing of linguistic expressions to parts of the world is a pairing that is always – or very nearly – open to a variety of different interpretations which are all adequate to the relevant collection of facts or phenomena. In the case of empirical science these interpretations can be mutually exclusive and yet empirically adequate, as in the case of the several possible current interpretations of quantum mechanics (e.g. Copenhagen vs. the 'many-worlds' hypothesis). If an empirical reason for preferring one over the other emerges, then the question is settled: one is superior in explanatory power to the other. If, on the other hand, there could not in principle be an empirical way of establishing the correctness of one to the detriment of its rival, then the tendency is to think that in fact the supposedly rival theories are isomorphic: really, they are the same. They share a structure, and though from the 'inside' of each structure it seems that the theory's objects are determined in a way that is incompatible with its rival's objects, since they have the same explanatory pay-off this is to be explained by their sharing a structure which is doing the explanatory work, and the particular way of determining the objects delivered by one theory or another is irrelevant to the really important issue of what does the explaining in the face of phenomena. Ontological commitments are therefore relative to a background theory, conceived of as a structure capable of multiple but ultimately equivalent interpretations (if the theory is true, anyway).

The argument of the *ante rem* structuralist is analogous for the case of strictly mathematical theories which need have no empirical component whatsoever. An *ante rem* structure is a 'freestanding' one over many, to borrow Shapiro's (1997) phrase, not categorical with respect to its objects but determinate only with respect to the system of relations between objects that it embodies. The 'objects' in question are positions in the structure, characterized exclusively by their mutual relationships. The perfect example is given by the Peano postulates, whose characterization of arithmetic is one that invokes relations among positions in the structure defined, rather than among determinate objects. Different kinds of objects can occupy structural positions of the right kind, and so there is never any principled reason (though there might be pragmatic ones) to privilege one system of objects over another. Any system fulfilling the conditions is isomorphic to any other and there is nothing to choose between them. Note that *ante rem* structuralism is a form of mathematical platonism: structures are real abstract objects whose existence is not contingent upon being instantiated by anything in particular.

But this is a kind of realism that embodies a suspicious dualism of scheme and content. The thesis of referential relativity to a scheme or structure that is independent of its particular content runs up against objection of the kind offered by Davidson's (198?) to scheme/content dualism in general, which is applicable to the *ante rem* structuralists' case as well. The idea is this: in order to interpret each others' natural language speech, we are forced to assume that every speaker shares an overwhelming majority of ontological assumptions about how the world is carved up with every other speaker. This is because in order to understand each other we must be able to discern the truth conditions for particular utterances (or inscriptions), and the process of interpretations would never be able to get going unless we assume a massively shared background of beliefs. The very existence of fluently multilingual people is a confirmation instance for the Davidsonian argument: if we were to suppose that semantic content is substantially dependent upon a schematic background, we would not be able to discern the speaker's assumptions about truth conditions required for interpretation because since those assumptions are relative to a background 'scheme' that we the interpreters do not share, interpretation and subsequent communication would become impossible. Now, in the mathematical case, the structuralists (both *ante rem* and *in re*) claim that semantic content is irrelevant to the communication of mathematical thoughts, and that structure alone is able to account for the shareability of mathematical thoughts. But it is far from clear that it is possible to do mathematics without doing semantics, at least tacitly. Structuralism is not coextensive with the old Hilbert-style formalism (discussed in the opening of chapter 2) but structuralism does share the ambition of removing the semantic from the mathematical to the extent that it wants to declare particular assignments of objects to positions irrelevant to the practice of mathematics, that is, irrelevant to the setting out of the truth conditions for mathematical claims. It is true to claim that there is an interpretation of the *symbols* '0, 1, 2, 3, 4, …' on which they fulfill the same role in Peano structure as the *symbols* '2, 4, 6, 8, 10, …', but for the purposes of generating valid mathematical inferences and communicating them to each other, we cannot possibly suppose referential relativism. On the contrary, we have to suppose that a shared *meaning* assigned to the symbols is indispensable for the communication of mathematical results and for the assignment to them of truth conditions that will allow us to account for truth-preserving inference in mathematics. *Ante rem* structuralism, while platonist on the ontological side, is semantically eliminativist, and that eliminativism leaves too many open questions to be theoretically satisfactory.

5.3. *In re* structuralism

In re structuralism, on the other hand, is eliminativist *tout court*. Mathematical objects are accounted for as positions in structures, but the structures themselves have no standing independently of whatever currently, and contingently, happens to occupy positions that we human minds order under certain relational predicates.[25] The fundamental argument is again one that is drawn from the theory of reference, this time from a stringent version

[25] The structuralism I have in mind here is developed by Chihara (2004), which is the fullest statement and defence I know of.

of the causal theory. It comes out well in an examination of Chihara's 'constructibility theory', which begins by establishing some 'new' quantifiers. Thus Chihara:

> Constructibility quantifiers are sequences of primitive symbols: either (C –) or (A –), where '—' is to be replaced by a variable of the appropriate sort. Using 'Ψφ' to be short for 'φ satisfies Ψ', '(Cφ)Ψφ' can be understood to say:
>
> It is possible to construct and open-sentence φ such that φ satisfies Ψ
>
> whereas '(Aφ)Ψφ' can be understood to say:
>
> Every open-sentence φ that it is possible to construct is such that φ satisfies Ψ
>
> (Chihara 2004: 170)

He goes on to say that the usual rules of quantifier negation apply to the constructibility quantifiers, that 'open-sentence' should be read as 'open-sentence *token*', i.e. as a physical inscription, and finally 'to say that an open-sentence of a particular sort is constructible is not to imply or presuppose that any such open-sentence token actually exists or, indeed, that anything exists.' (Chihara 2004: 171) The subsequent machinery requires only a level-zero of objects and then a regimented hierarchy of symbols for open-sentences of levels 1, 2, and so on up, plus connectives. In fact what Chihara develops is a Tarski-hierarchy physicalistically interpreted, of much the kind that Field (1972) proposed, which takes only the notion of satisfaction as primitive. Chihara's (C –) and (A –) quantifiers are respectively the existential and universal substitutional quantifiers Σ and Π, taken to range only over tokens, or at any rate over possible constructible tokens: this is to say that the universe that is presupposed on Chihara's view is exclusively that of possible constructions, in particular the universe of possible open-sentence constructions. On this basis Chihara (2004: 170 - 184; 1990 *passim*) works out a logic for the foundations of mathematics that does not make appeal to any entities of level-zero other than constructible open-sentence tokens themselves.

From the epistemological point of view I am indeed very sympathetic to this approach, though I think that substitutional quantification is not the right framework for a discussion of mathematical ontology and the commitments that come with it. But Chihara's substitutional logic is able to code the kind of elementary, material forms of inference that get psychologically inculcated into learners in the process of pattern-recognition. The logic is able to code how, from perceived patterns, we can express the structure of inference (for example in ordinary addition and multiplication) which is gradually mastered in this process. The structure of substitutional quantification displays a theoretical systematization of what happens in mastering elementary arithmetical inference, how patterns can be generalized and made available in thought as *recurring* structures that are multiply instantiated. Furthermore, the notion of a possible construction of an open sentence avoids the objection standardly raised against Field's (1972) construal of Tarski's theory of truth and which could be applied to substitutional quantification generally, which is to ask what we do with terms and predicates newly introduced into the language, because it (i.e. Chihara's notion) appeals only to possible constructions and 'possible' is interpreted in the sense of 'conceptual possibility', and thus includes tokens that are infinitely long or otherwise physically unrealizable. So Chihara places no restrictions on the zero-level, other than that the objects there invoked be constructible in the broadest sense of 'possibility': here we deflate a surface commitment to unsuspected modalities by making those modalities explicit and subsequently defusing them by looking to *conceptual* possibility, without being committed to any *actual* realization of the conceptual possibility. Chihara in fact develops the purest nominalism, and his technique is a success, so long as we accept his construal of modality and the idea that physically-inscribed tokens and structural patterns of these are the *genetic* roots of elementary arithmetical beliefs, even supposing a stringent causal theory of knowledge that holds that only those objects with which we are in causal-perceptual contact can be responsible for the generation of veridical beliefs.

But now we run into a familiar problem, raised by Benacerraf. I discuss this problem in the following section, and show how the contrasts between genesis and justification can open an avenue for a satisfactory response.

5.4. Benacerraf's challenge

Suppose for the sake of argument that we have good reason to think that Fregean realism about mathematical objects is the right way of giving a truth-conditional semantics for mathematical language and accounting for the truth of true mathematical propositions. In fact, let's pretend that the arguments I've offered in chapter 1, and especially in chapter 2, are sound. Bring in Benacerraf's (1973) challenge to the ontological platonist: if you hold that putative mathematical objects are the truthmakers for mathematical propositions, and that such objects are outside the causal order, then how do we come to acquire mathematical beliefs at all? I propose in this chapter to hold the previous chapters' ontology of arithmetical and objective truthmakers constant, and try to solve for the problem of our knowledge of them. But in any case, I think that what I claim in this section is not contingent upon any particular ontology.

A couple of easy responses to Benacerraf should be cleared out of the way from the start. The first possible response is to claim that Benacerraf presupposes that the causal theory of knowledge is our best epistemological bet, but that the causal theory can't be a true bill because there are counter-examples that defeat it. For instance, James Robert Brown (1990) invokes the Einstein-Podolsky-Rosen problem: if we imagine a two-particle entangled system in which the two particles lie outside of each others' light-cones, and make a measurement on one of them, then the other is bound to have the opposite measurable property. If, say, particle α measures 'white', then particle β will necessarily measure 'black'. Since there can be no causal relation between the two particles (unless causation can somehow operate at greater than the speed of light), says Brown, and following Bell's result showing that there is no hidden variable operative in the system, then we have a case in which we gain knowledge of one of β's properties even though we have no hope of causal contact with β. This is proposed as a counter-example to the causal theory of knowledge as a global epistemology.

I imagine that the defender of the causal theory would reply that we have *inferred* a property of β on the basis of an observation made on α, and that inference should somehow be construed as a causal relation, perhaps by arguing that inference to necessary conclusions is a properly causal relation inheritable from observable phenomena to their consequences. Such is the strategy adopted by Cheyne (2001), for instance. This kind of strategy has obvious empiricist presuppositions, and it might be that expanding our notion of causation to include inferential consequences is not much of a stretch, but it seems to me that there are convincing arguments for thinking that the normative cannot be reduced to the contingencies of the causal. For instance, if the brain of every logician in the world were implanted with a microchip that made the logicians believe that *modus ponens* is not a valid form of argument – a chip that made them believe that *modus ponens* is a fallacy on a logical par with affirming the consequent, for example – we would still not be right in saying that *modus ponens* is an invalid form, even though the strict empiricist would be committed to this view, since it would be just one of the facts of the empirical world.

In any case, I want to set aside the argument by counter-example to the causal theory of knowledge, not on the grounds that it fails, but on the grounds that it really doesn't tell us much. A proper response to Benacerraf's epistemological challenge should be constructive and say something positive about how we actually do come to have knowledge of mathematical truths and the capacity to acquire mathematical beliefs.

The second possible 'easy' response goes like this: let's say that, Gettier notwithstanding, knowledge is best analyzed as justified, true belief. And we do have mathematical beliefs; there's no doubt about that. We also can justify those beliefs: that is the role of *proof*. Admittedly, proof is a very special mode of justification with a limited sphere of application (it only works for some propositions of logic and mathematics), but it does the job, subject to a few well-known constraints, like Gödel's Theorems. The only problem left over, then, is the problem of truth, and we solve that by appealing to an ontology of mathematical truthmakers.

That by itself won't do. A suitable epistemology for mathematical truths faces some challenges that don't arise for other cases of claims to knowledge. When I claim to know that the cat is on the mat, I straightforwardly *believe* that the cat is on the mat. I can *justify* the claim by appealing to veridical perception of cats and mats in ideal conditions, and by further appeal to my knowledge of how '… is on ___' functions in the language. And I can say that the claim is made *true* by the right cat, the right mat and the right relation between them, with the respective rightnesses accounted for by some suitable context-sensitive semantics. So I have a good case for a knowledge-claim. But in this cat-and-mat-based case I also have a convincing story about how I came to acquire the belief in the first place: that story has to do with how I learned to individuate cats and mats, and use relational predicates like "… is on ___'. Benacerraf's challenge to the mathematical platonist is so hard precisely because it does *not* attack either the truth or the justification of mathematical knowledge-claims. It's hard because it asks the platonist to design a convincing argument to show how mathematical beliefs are in some way *connected* to their truthmakers.

This is what I propose at least to sketch in sections 4.5. and 4.6. below. But first, a few more negative remarks. I've indicated that two "easy" solutions of Benacerraf's problem aren't viable. There is a third: naturalism, in one form or another. Naturalism, in all its incarnations, has the great advantage of accounting for all the desirable features of a theory of the acquisition of mathematical beliefs in a way that is continuous with a standard story about how beliefs of any kind at all are acquired. Some naturalists are platonists of a kind (Bigelow), some are realists (Maddy), some are flat-out nominalists (Field); without surveying the very complex territory that different positions occupy, I want to present a general argument that applies to any form of epistemic *justificatory* naturalism in the philosophy of mathematics.

Structuralists are in the happy position of being able to explain how we acquire mathematical beliefs in a way that squares neatly with naturalistic commitments: they do so by saying that we just see mathematical objects in the ordinary perceptual way. Seeing a dozen eggs, for instance, involves seeing an egg and an egg and an egg and an egg and an egg and an egg and an egg and an egg and an egg and an egg and an egg and an egg. Being able to use the term 'dozen' or, at a new level of abstraction, 'twelve', is just the ability somehow to read 12 off things iterated twelvefold, and that ability derives from the prior ability to individuate things or particulars or whatever. Essentially, naturalism claims that mathematical knowledge is derived from the perception of difference between empirically available objects of sensory perception. If S knows that 2 + 2 = 4, S knows that two things (particulars) and two more things (particulars) make four things.

Now, I've said that Benacerraf's challenge is so hard because it demands that the realist be able to account for how we acquire mathematical beliefs in the first place. But the challenge is also hard because it demands that we account for the mathematical beliefs that we *do* have, and this makes it a challenge not only for the realist, but for the causalist as well. In fact, things are worse for the naturalist, because a causalist epistemology seems to demand that we reject a vast swathe of mathematical beliefs as unacceptable, on the grounds that we can't generate the objects of our mathematical beliefs out of what the senses deliver.

Suppose that the naturalist is right, and that only what is empirically available to the senses is available to the mathematician. This includes, of course, every relation between the empirically available bits. The motivation for such a view is extremely attractive, for the following reason. If I claim that the cat is on the mat, and I understand the sentence 'the cat is on the mat' because I understand what must be the case in the world in order for the definite description to be true, then I know how to go about verifying that the cat is indeed on the mat. My understanding of 'the cat is on the mat' is explained by a prior grasp of the language in which talk of cats and mats and being-on makes sense, and a truth-conditional semantics explains why I go about checking for the truth of this sentence by hunting about for a cat and a mat and a property said to hold between them. Extending this picture to mathematical language, the naturalist has a *prima facie* appealing account of how mathematical sentences come to be understood and where to look to verify the claims they make. Ultimately, mathematical language is as responsible to the physical world, and only the physical world, as the rest of physical language. We have causal contact with the referents of mathematical terms in the very same ways that we have causal contact with more run-of-the-mill referents.

The problem that, in my view, defeats epistemic naturalism as a global account of mathematical knowledge lies in some significant explanatory failures, and those failures are of two kinds: epistemological on the one hand, and semantic on the other. In the next section I sketch some details of both problems.

5.5. Nature's reach

Genetic externalism, then, holds that it is physical-world structures and the perception of these that are causally responsible for the generation of elementary mathematical beliefs: this is their route into the mind. This is a process whose logical architecture is usefully framed in the terms of Chihara quantifiers because, while remaining ontologically square with naturalistic commitments, it can yet offer an account of the road from perceived structure to the coding of the elementary material inferences into which those perceived pieces of structure fit. In effect, in this context we need a theory of inference that from eyesight to conceptualization and availability for inference, and Chihara's story does this duty in a way that is entirely within naturalism's bailiwick. But in this section I will show that while a structuralist answer to the genetic question can be easily squared with naturalism, the same does not go for the justificatory question. Naturalism's reach is limited to genesis; it is silent on mathematical justification. And indeed the fact that these two questions are not tractable under the canons of a single method is a further motivation for splitting them apart.

'Naturalism' is a term that divides its reference in multiple ways, but a characterization of its commitments across various of its incarnations might fairly be articulated as follows. Naturalism is a thesis about which explanatory strategies are legitimate, a circumscription of legitimate methodology across any province of inquiry that claims to offer authentic explanations, whether that province is economic history, quantum mechanics or the best way to grow zucchini. Naturalism adheres to at least these two principles: (1) a methodological one, which enjoins us to appeal only to those resources that are indispensable to adequate explanations, and (2) a principle of exclusivity, which holds that only causal explanations are to be counted adequate. My claim in this section will be that, while principle (1) is perfectly sensible, principle (2) does not follow from (1), and is indeed false. The epistemic import of this claim is that, when mathematical knowledge becomes the subject matter which we are interested in accounting for, we see that naturalism is inadequate to the justificatory task and that different explanatory resources, not countenanced by devotees of naturalism, are required.

The first principle, an inheritance from the pragmatists and Quine, has the advantage of being broad enough to avoid prejudging the issue in favour of an empiricist epistemology and allows, for example, for postulating innate knowledge on properly naturalistic grounds (as we find in Chomsky, for instance). It also allowed Quine himself, again on naturalistic grounds, to postulate abstract objects such as sets. The second principle is more restrictive and I do not claim that it is characteristic of every strand of naturalism in the literature. Nevertheless, a substantial part of the naturalists' claim is that the methods and explanations furnished by the natural sciences are exemplars of our *best* theories, and those methods are to be imitated as far as possible in suitable theorizing.

The commitment here is not just that all adequate explanations are causal (i.e. inscribed in the natural order of relations between successive events in space and time), but also that the first principle *entails* the second. In what follows I want to challenge that entailment.

Borrowing the account of causal explanation given in Lewis (1986b), we find:

> (CE) An adequate explanation of an event is a probabilistic (and thus contingent) function relating *explanans* and *explanandum* in temporal succession, a function that makes the event to be explained more likely than it would have been in the absence of the previous event serving as *explanans*.

Causal explanation, being probabilistic, is contingent, and it is a relation between events unfolding in time. I do not want to challenge the (CE) statement; I want, rather, to challenge its claim to exclusivity over explanations, and challenge the supposed transition from principle (1) to (CE).

Begin with an example, one that's also raised by Pincock (2007) though for rather different purposes. The example is the famous 'bridges of Königsberg' problem, and it goes like this. The first figure (overleaf) sketches an overview of the seven bridges of the Prussian city of Königsberg as they stood in the mid-eighteenth century: Here we have a river, two islands in the river, and seven bridges connecting different parts of the city. Now we ask the following question:

> (BK) Is it possible to walk, or trace, a path through the city, starting from a given point and crossing each bridge once and once only?

The answer is *no*: no such path is possible. Why not? The answer is this. We can model the problem by considering the connected, non-Eulerian graph (fig. 2):

Illustrative figures

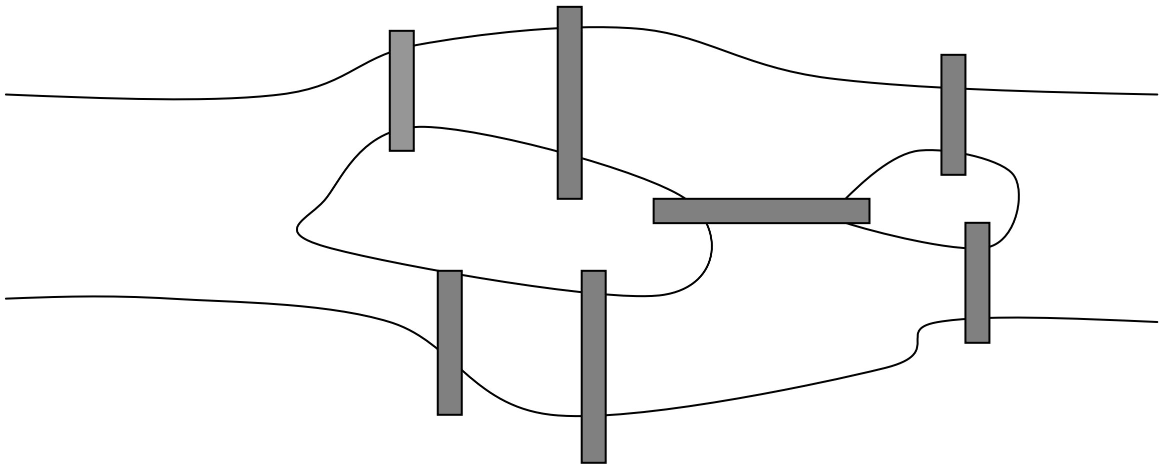

Fig. 1: The seven bridges of Königsberg

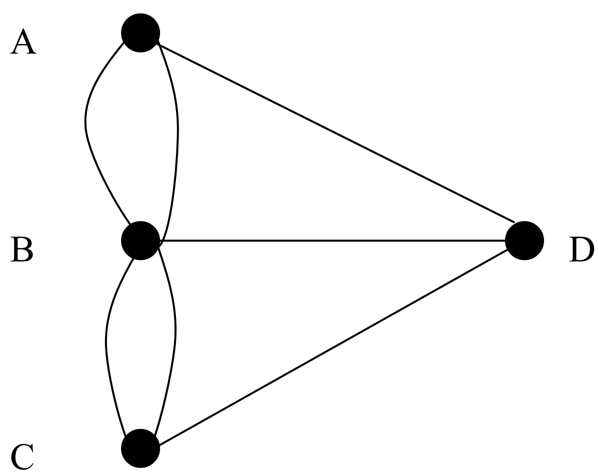

Fig. 2: Non-Eulerian graph of the bridges of Königsberg

In the graph, A and C are points in the city on the riverbank, and B and D are points on the islands. The connecting lines are bridges. Each point is a vertex and each line an edge. The number of edges joined at a vertex we'll call that vertex's degree. Call a path of the kind described in (BK) an 'Eulerian path'. Now, loosely speaking, the reason you can't walk an Eulerian path through Königsberg is that, if it *were* possible, there would have to be one way out of each point for every way into that point. That is, every vertex would have to have even degree (one out for one in); but, looking at the graph, all four vertices have odd degree. This is also the sense in which the graph above is *non*-Eulerian; an Eulerian graph is one in which an Eulerian path can be traced. In fact, more precisely, an Eulerian path on a connected graph is possible if there are *at most two* vertices of odd degree; in other words, those vertices would be either the beginning or the end of the path: a way in, a way out, and a way back in and then you stop. But this graph has more than two vertices of odd degree.

This is the explanation of why you can't walk around Königsberg a certain set way, and it is a non-causal one: it does not relate events probabilistically in temporal order. Indeed, this is an instance of the argument from constraint, although we should note that it does not do the same work as was done by the transcendence of π in the first chapter because this bit of graph theory can well be accommodated by Field's geometric nominalization strategy described in § 1.2. above, in the vocabulary of arrays of regions. Nevertheless, this argument from graph theory is not an explanation of causal relationships between points in the array; it is an explanation of a constraint given by the degree of vertices.

I conclude on this basis, then, that principle (1), which enjoins us to make ontological appeal only to those resources which are indispensable for the adequate explanation of phenomena, does not entail the exclusivity principle (2): we need not restrict ourselves to those resources which are causal in the sense laid out by (EC). But what, now, is the epistemological import of this argument?

For the purposes at issue here, it is this. We want an explanation of how mathematical knowledge is possible. I have argued that such an explanation really needs to tackle two questions: the genetic one, which can be handled by the causally describable features of naturally occurring structures and the perception of these via a faculty of pattern-recognition; and the justificatory one, about which I have so far been silent, except to argue that one process cannot cover both the first and second questions because the causal resources invoked to answer the first can deliver only partial warrant for inductively-established contingent claims, which are insufficient for accommodating the necessity of metaphysical constraint and the necessity of inferential relations that stand (as established by mathematical induction, say) among mathematical propositions. But by the argument in this section, this should not be especially worrisome from the epistemological point of view, because, as we have just seen, there are in fact convincing grounds for rejecting causalist strictures on explanation. The path is then open to begin replying to the justificatory question non-causally.

5.6. How do you get there from here?

Having gone some distance down one of the paths that split at Benacerraf's fork, having pressed the genetic question and its distinctness from the justificatory, what can now be said to have this path rejoin the other, the one where the abstract mathematical truthmakers live? How do you get there from here? By the argument from the previous section, it is open to us to offer without qualms a non-causal reply to the justificatory question, one which can both be sensitive to the requirements of mathematical thought and talk as invoking ineliminable abstract objects, *and* begin to supply an account of how they might be known. That we have any elementary mathematical beliefs at all is accountable to what nature affords us as perceptible structure; but that we are able to justify the full, rich range of necessary mathematical propositions cannot be cashed out by the contingencies of the perceptual apparatus and the partial structures embodied in the physical world.

While the generation of mathematical beliefs, at an elementary level, is causal, their justification is *normative*. In justificatory thought about the philosophical foundations of mathematics there is of course a good deal of disagreement over which norms of reasoning in particular ought to be operative. Constructivists, for instance, spurred by verificationist worries, are keen to place certain characteristic normative constraints on allowable constructions. The constructivist denies, for instance, that pure existence proofs are legitimate ways of forming mathematical beliefs. Take an example to illustrate the point: in 1874, Cantor proved that there are transcendental numbers by showing that since there is an injective function from the algebraic reals to **N**, but the cardinality of **R** is greater than the cardinality of **N**, there must be non-algebraic, i.e. transcendental, reals. His pure existence proof is non-constructive: it does not give an instance of a transcendental number, but shows rather that there must be such things, though we know not which.[26] The constructivist worry is that unless and until you can point to an instantiation of a particular property, say the property of being transcendental, you can have no reliable beliefs about the putative objects of belief because you don't know what they are. In essence the constructivist says that if you can't construct (give an example

[26] As a matter of fact Liouville gave a constructive proof in 1851. But until Cantor, transcendentals were thought to be pretty rare. As it turns out, practically all the members of **R** are transcendental.

of) the object of a belief, you don't know what you're talking about and have thus violated an overarching norm of reasoning that places epistemic constraints on evidentiary canons: you have failed to live up to a norm of inquiry. Whether or not there are such objects of belief as transcendentals is an open question until an instance is produced. The underlying intuition of constructivism as a norm of mathematical reasoning is not peculiarly mathematical, but more broadly epistemic, with wide application across all provinces of rational inquiry. Let's suppose that some evolutionary biologist determines that this proposition:

(P) there is a species intermediate between eubacteria and eukaryota

coheres with the whole, or at least with all the relevant bits of, well-established theory. (P) in itself is perhaps a reason for hunting about for evidence of the intermediate species, but coherence by itself does not warrant the *assertion* of (P). If I'm an evolutionary biologist interested in the earliest forms of life on this planet and I learn that (P) is a thesis coherent with the rest of what I accept, then I might be inclined to do my best to find a specimen of the putative intermediate species. But until I find an instance, I can have no reliable beliefs about the intermediate species because nothing about its supposed members has been said, save that they exist. A great many existence-claims are coherent with any body of theory, but coherence is not the criterion, the regulative norm, by which we judge the truth of an existence-claim in the natural sciences. Coherence is a reason to look for instances of a putative kind of entity, but in the natural sciences it is not evidence of existence; no more should it be, says the constructivist, in mathematics.

Now, not all constructivist views need necessarily ride on a supposed continuity between mathematical and non-mathematical norms of verification.[27] If we take the constructivist objection to be about which norms should be operative in justifying the mathematical beliefs that we actually do have, rather than about how we acquire any mathematical beliefs at all, then there is something deep about Dummett's claim that 'the only thing which will justify the assertion of a mathematical statement is the existence of a proof, and, when "existence" is not interpreted platonistically, this is something of which we cannot be unaware.' (Dummett 1982: 259)

McDowell (1989) puts a related objection nicely:

> The platonist … has it that our understanding of the relevant [mathematical] sentences consists in a conception of what it would be for them to be true that might outrun what we could display in sensitivity to proofs and refutations – even including those that could in principle be devised. But what other aspects of our use of arithmetical sentences could be cited as potentially manifesting the supposed residue of understanding? … [P]latonism apparently makes no concession to the thought that someone's understanding of a sentence must be able to be made fully overt in his use of it. The trouble is that this leaves it a mystery how one person can know another person's meaning. (McDowell 1989: 347)

This sort of objection turns on the nature of the norms governing mathematical thought, the norms that govern the give-and-take of reasons and evidence in mathematical exchanges. If the platonist is left with a commitment to some kind of 'residue' dangling in the third realm which finds no resonance in thinkers' sensitivity to proofs and

[27] Though *de facto*, most seem to. See e.g. Prawitz (1998) for an example of constructive verificationists' remarkable reluctance to move from even accepted low-order observations to their consequences.

refutations, then platonism is in that respect surplus to requirements. There can be no commitment that overreaches epistemic entitlement.

Is platonism so committed? Not so, in my view, for this reason. Entitlement does not halt when *constructive* evidence gives out. Recall the argument from methodological continuity given in the first chapter: if it is rational to believe some thesis T (along with its commitments) on the grounds of some method M, and M also establishes an extension T+ of T, then it is equally rational to believe T+ on those same grounds M. The ingredients of M are the set of justificatory norms of reasoning that have emerged out of mathematical practice. If it is rational to follow Cantor from a collection T of theses about sets through to T+ about the existence of transcendentals via methods invoking injective functions and relatively larger cardinalities, then we have excellent normative grounds for believing in transcendentals. Having prised apart the questions of the genesis and justification of mathematical belief, we can envisage a mode of reconnecting the two through norms of mathematical reasoning: beliefs that are causally generated are taken up, justified and extended thanks to the norms of thought operative in mathematical practice. How those norms are structured, I have not even begun to say. But I conclude, on the basis of what has been argued in this essay, that abstract objects are indispensable in the right metaphysics of what is required for mathematical thought.

Bibliography

--Austin, J. L. (1956), 'A Plea for Excuses', in his *Philosophical Papers*, ed. J. Urmson & G.J. Warnock, Oxford University Press, Oxford, 1970.

--Azzouni, Jody (1997), 'Applied Mathematics, Existential Commitment and the Quine-Putnam Indispensability Thesis', *Philosophia Mathematica* (3), Vol. 5, 1997.

--Azzouni, Jody (2004), *Deflating Existential Consequence*, Oxford University Press, Oxford.

--Baker, Alan (2005), 'Are There Genuine Mathematical Explanations of Physical Phenomena?', *Mind*, Vol. 114, 454, April 2005.

--Balaguer, Mark (1995), 'A Platonist Epistemology', *Synthèse* 103.

--Balaguer, Mark (1998), *Platonism and Anti-Platonism in Mathematics*, Oxford University Press, New York & Oxford.

--Beaney, Michael (1997), ed., *The Frege Reader*, Blackwell, Oxford.

--Benacerraf, Paul (1965), 'What Numbers Could Not Be', *Philosophical Review*, 74, 1965.

--Benacerraf, Paul (1973), 'Mathematical Truth', *Journal of Philosophy*, 70, 1973.

--Berkeley, George (1710), *The Principles of Human Knowledge*, in *Berkeley's Philosophical Writings*, ed. D.M. Armstrong, Macmillan, New York, 1965.

--Berkeley, George (1734), *The Analyst*, in *Berkeley's Complete Works,* Vol. III, ed. A.C. Fraser, Oxford University Press, Oxford, 1901.

--Boghossian, Paul & Christopher Peacocke (2000), eds., *New Essays on the A Priori*, Oxford University Press, Oxford.

--Bonevac, Daniel (1982), *Reduction in the Abstract Sciences*, Hackett, Indianapolis.

--Boolos, George (1998), *Logic, Logic, and Logic*, Harvard University Press, Cambridge, Mass.

--Brown, James Robert (1990), 'π in the Sky', in *Physicalism in Mathematics*, ed. A. D. Irvine, Kluwer Academic Publishers, Dordrecht, 1990.

--Burge, Tyler (2005), *Truth, Thought, Reason: Essays on Frege*, Clarendon Press, Oxford & New York.

--Burgess, John P. & Gideon Rosen (1997), *A Subject With No Object: Strategies for Nominalistic Interpretation of Mathematics*, Clarendon Press, Oxford.

--Burgess, John P. (2004), 'Mathematics and *Bleak House*', *Philosophia Mathematica*, Vol. 3, No. 2, 2004.

--Carnap, Rudolf (1950), 'Empiricism, Semantics and Ontology', repr. in Carnap (1956).

--Carnap, Rudolf (1956), *Meaning and Necessity: a Study in Semantics and Modal Logic*, University of Chicago Press, Chigaco, 1964.

--Cheyne, Colin (2001), *Knowledge, Cause and Abstract Objects*, Kluwer Academic Publishers, Dordrecht.

--Chihara, Charles (2004), *A Structural Account of Mathematics*, Clarendon Press, Oxford.

--Coffa, J. Alberto (1991), *The Semantic Tradition from Kant to Carnap: To the Vienna Station*, Cambridge University Press, Cambridge.

--Colyvan, Mark (2001), *The Indispensability of Mathematics*, Oxford University Press, New York.

--Demopoulos, William (1995), ed., *Frege's Philosophy of Mathematics*, Harvard University Press, Cambridge, Mass.

--Dennett, Daniel (1981), 'Reflections on Hofstadter', in Hofstadter & Dennett, eds, *The Mind's I*, Basic Books, New York.

--Donnellan, Keith (1966), 'Reference and Definite Descriptions', *The Philosophical Review,* LXXV, July 1966.

--Dummett, Michael (1982), 'Realism', *Synthèse* 52. Reprinted in Dummett (1993).

--Dummett, Michael (1991), *Frege: Philosophy of Mathematics*, Harvard University Press, Cambridge, Mass., 2nd printing 1995.

--Dummett, Michael (1993), *The Seas of Language*, Oxford University Press, Oxford.

--Dummett, Michael (1994), 'Wittgenstein on Necessity: Some Reflections', in *Reading Putnam*, ed. P. Clark & B. Hale, Blackwell, Cambridge, Mass., 1995. Also in Dummett (1993).

--Elgin, Catherine Z. (2004), 'True Enough', *Philosophical Issues* 14, 2004.

--Evans, Gareth (1982), *The Varieties of Referencec*, ed. John McDowell, Clarendon Press, Oxford.

--Feynman, Richard (1967), *The Character of Physical Law*, M.I.T. Press, Cambridge, Mass.

--Field, Hartry (1972), 'Tarski's Theory of Truth', in Field (2001a).

--Field, Hartry (1980), *Science Without Numbers*, Princeton University Press, Princeton, N.J.

--Field, Hartry (1989), *Realism, Mathematics & Modality*, Basil Blackwell, Oxford & New York.

--Field, Hartry (1994), 'Disquotational Truth and Factually Defective Discourse', in Field (2001a).

--Field, Hartry (2001a), *Truth and the Absence of Fact*, Oxford University Press, Oxford, 2001.

--Field, Hartry (2001b), 'Apriority as an Evaluative Notion', in Field (2001a).

--Fine, Kit (2002), *The Limits of Abstraction*, Clarendon Press, Oxford.

--Floyd, Juliet & Sanford Shieh (2001), eds., *Future Pasts: The Analytic Tradition in Twentieth-Century Philosophy*, Oxford University Press, Oxford.

--Frege, Gottlob (1879), *Begiffsschrift und andere Aufsätze*, ed. Ignacio Angelelli, Georg Olms Verlagsbuchhandlung, Hildesheim, 1964. [See also Beaney (1997) and van Heijenoort (1967).]

--Frege, Gottlob (1884), *The Foundations of Arithmetic*, trans. J.L. Austin, Northwestern University Press, Evanston, Illinois, 1980.

--Frege, Gottlob (1885), 'On Formal Theories of Arithmetic', in *On the Foundations of Geometry and Formal Theories of Arithmetic*, trans. & ed. Eike-Henner W. Kluge, Yale University Press, New Haven, Conn., 1971.

--Frege, Gottlob (1892a), 'On Concept and Object', trans. P.T.Geach, in Geach & Black (1952).

--Frege, Gottlob (1892b), 'On Sense and Meaning', trans. Max Black, in Geach & Black (1952).

--Frege, Gottlob (1893), *The Basic Laws of Arithmetic*, trans. M. Furth, University of California Press, Berkeley & Los Angeles, 1967.

--Frege, Gottlob (1894), 'Review of Husserl's *Philosophie der Arithmetik'*, trans. P.T. Geach, in Geach & Black (1952).

--Frege, Gottlob (1897/1969), 'On Herr Peano's *Begriffsschrift* and My Own', *Australasian Journal of Philosophy*, Vol. 47, No. 1, May 1969.

--Frege, Gottlob (1914), 'Logic in Mathematics', in Frege (1979).

--Frege, Gottlob, (1918), 'Thoughts', in Frege (1980).

--Frege, Gottlob (1980), *Philosophical and Mathematical Correspondence*, ed. Gottfried Gabriel, Hans Hermes et al., trans. Hans Kaal, University of Chicago Press, Chicago.

--Frege, Gottlob (1979), *Posthumous Writings*, ed. Hans Hermes et al., Basil Blackwell, Oxford.

--Frege, Gottlob (1984), *Collected Papers on Mathematics, Logic, and Philosophy*, ed. Brian McGuinness, Basil Blackwell, Oxford.

--Friedman, Michael (2001), 'Tolerance and Analyticity in Carnap's Philosophy of Mathematics', in Floyd & Shieh (2001).

--Furth, Montgomery (1968), 'Two Types of Denotation', in *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, No.2, ed. N. Rescher, Basil Blackwell, Oxford, 1968.

--Gabbay, Dov & Franz Guenther (1983), *Handbook of Philosophical Logic*, D. Reidel, Dordrecht.

--Gaukroger, Stephen (1992), 'The Nature of Abstract Reasoning: Philosophical Aspects of Descartes's Work in Algebra', in *The Cambridge Companion to Descartes*, ed. John Cottingham, Cambridge University Press, Cambridge, 1992.

--Geach, Peter & Max Black (1952), eds., *Translations from the Philosophical Writings of Gottlob Frege*, Basil Blackwell, Oxford, 3rd edition 1980.

--Geach, Peter (1971), *Mental Acts*, reprinted by Thoemmes Press, Bristol, 1992.

--Hale, Bob (1987), *Abstract Objects*, Blackwell, Oxford

--Hale, Bob & Crispin Wright (1997), eds., *A Companion to the Philosophy of Language,* Blackwell, Oxford.

--Hale, Bob & Crispin Wright (2000), 'Implicit Definition and the A Priori', in Boghossian & Peacocke (2000).

--Hale, Bob & Crispin Wright (2001), *The Reason's Proper Study*, Clarendon Press, Oxford.

--Hazen, Allen (1983), 'Predicative Logics', in Gabbay & Guenther (1983).

--Hughes, G.E. & M.J. Cresswell (1968), *An Introduction to Modal Logic*, Methuen & Co., London.

--Jacquette, Dale (2002a), ed., *Philosophy of Logic: an Anthology*, Blackwell, Oxford.

--Jacquette, Dale (2002b), ed., *Philosophy of Mathematics: an Anthology*, Blackwell, Oxford.

--Jubien, Michael (1977), 'Ontology and Mathematical Truth', *Noûs*, 5, 11.

--Kalderon, Mark Eli (2005), ed., *Fictionalism in Metaphysics*, Clarendon Press, Oxford.

--Kant, Immanuel (1787), *Critique of Pure Reason*, trans. Norman Kemp Smith, MacMillan & Co., London, 1958.

--Keefe, Rosanna & Peter Smith (1999), eds., *Vagueness: A Reader*, MIT Press, Cambridge, Mass.

--Kitcher, Philip (1983), *The Nature of Mathematical Knowledge*, Oxford University Press, New York.

--Kitcher, Philip (2000), 'A Priori Knowledge Revisited', in Boghossian & Peacocke (2000).

--Kripke, Saul A (1980), *Naming and Necessity*, Harvard University Press, Cambridge, Mass.

--Kripke, Saul A. (1982), *Wittgenstein on Rules and Private Language*, Harvard University Press, Cambridge, Massachusetts.

--Lamarque, Peter & Stein Haugom Olsen (1994), *Truth, Fiction, and Literature*, Clarendon Press, Oxford. Reprinted 2002.

--Leibniz, Gottfried Wilhelm (1765), *Nouveaux Essais sur l'entendement humain*, GF-Flammarion, Paris, 1990.

--Lewis, David (1970), 'Nominalistic Set Theory', reprinted in Lewis (1998).

--Lewis, David (1986a), *On the Plurality of Worlds*, Basil Blackwell, Oxford.

--Lewis, David (1986b), 'Causal Explanation', in his *Philosophical Papers, Vol. II*, Oxford University Press, New York & Oxford, 1986.

--Lewis, David (1993), 'Mathematics is Megethology', reprinted in Lewis (1998).

--Lewis, David (1998), *Papers in Philosophical Logic*, Cambridge University Press, Cambridge.

--Lindemann, Ferdinand von (1882), 'Über die Zahl π', *Mathematische Annalen*, Vol. 20, No. 2, June 1882.

--Linsky, Leonard (1967), *Referring*, Humanities Press, Atlantic Highlands, N.J.

--Linsky, Leonard (1971), ed., *Reference and Modality*, Oxford University Press, Oxford.

--Ludlow, Peter (1997), ed., *Readings in the Philosophy of Language*, MIT Press, Cambridge, Mass.

--Malament, David (1982), 'Review of Field (1980)', *Journal of Philosophy*, Vol. 79, No.9, September 1982.

--Marion, Mathieu (1998), *Wittgenstein, Finitism and the Foundations of Mathematics*, Clarendon Press, Oxford.

--Marion, Mathieu (MS), 'Wittgenstein on Mathematics: Constructivism or Constructivity?', manuscript.

--McDowell, John (1989), 'Mathematical Platonism and Dummettian Anti-Realism', *Dialectica* 43. Reprinted in McDowell (1998).

--McDowell, John (1998), *Meaning, Knowledge & Reality*, Harvard University Press, Cambridge, Mass.

--Parsons, Charles (1964), 'Infinity and Kant's Conception of the "Possibility of Experience"', in Parsons (1983).

--Parsons, Charles (1965/1995), 'Frege's Theory of Number', in Demopoulos (2001).

--Parsons, Charles (1971a), 'Ontology and Mathematics', *Philosophical Review,* Vol. 80, 1971.

--Parsons, Charles (1971b), 'A Plea for Substitutional Quantification', in Jacquette (2002b).

--Parsons, Charles (1979/1980), 'Mathematical Intuition', *Proceedings of the Aristotelian Society*, New Series, Vol. LXXX, The Aristotelian Society, London, 1980.

--Parsons, Charles (1982), 'Objects and Logic', *The Monist*, 65, 1982.

--Parsons, Charles (1983), *Mathematics in Philosophy*, Cornell University Press, Ithaca, NY. Reprinted 2005.

--Parsons, Charles (1986), 'Review of (Kitcher's) *The Nature of Mathematical Knowledge*', *Philosophical Review*, XCV, No.1, January 1986.

--Parsons, Charles (1990), 'The Structuralist View of Mathematical Objects', *Synthèse* 84, 1990.

--Pincock, Christopher (2007), 'A Role for Mathematics in the Physical Sciences', *Noûs* 41:2, 2007.

--Potter, Michael (2000), *Reason's Nearest Kin: Philosophies of Arithmetic From Kant to Carnap*, Oxford University Press, Oxford. (Reprinted 2004).

--Prawitz, Dag (1998), 'Truth and Objectivity From a Verificationist Point of View', in *Truth in Mathematics*, ed. H.G. Dales & G. Oliveri, Clarendon Press, Oxford, 1998.

--Priest, Graham (1998), 'What is so Bad about Contradictions?', *Journal of Philosophy*, Vol. 95, No. 8, August 1998.

--Putnam, Hilary (1956), 'Mathematics and the Existence of Abstract Entities', *Philosophical Studies*, Vol VII, No.6, December 1956.

--Putnam, Hilary (1967), 'Mathematics Without Foundations', reprinted in Putnam (1979a).

--Putnam, Hilary (1971), *Philosophy of Logic*, Harper & Row, New York.

--Putnam, Hilary (1975), 'What is Mathematical Truth?', in Putnam (1979a)

--Putnam, Hilary (1979a), *Mathematics, Matter and Method* (2nd Edition), Cambridge University Press, Cambridge. Repr. 1980.

--Putnam, Hilary (1979b), 'Analyticity and Apriority: Beyond Wittgenstein and Quine', *Midwest Studies in Philosophy, Vol IV*, ed. French, Uehling & Wettstein, University of Minnesota Press, Minneapolis, 1979.

--Quine, W.V. (1947), *Mathematical Logic*, Harvard University Press, Cambridge, Mass.

--Quine, W.V. (1951a), 'Two Dogmas of Empiricism', in Quine (1953).

--Quine, W.V. (1951b), 'On Carnap's Views on Ontology', in Quine (1966).

--Quine, W.V. (1953), *From a Logical Point of View*, Harvard University Press, Cambridge, Mass., 2nd revised edition, 1980.

--Quine, W.V. (1960), *Word and Object*, M.I.T. Press, Cambridge, Mass., 18th printing 1992.

--Quine, W.V. (1964), 'Ontological Reduction and the World of Numbers', in Quine (1966).

--Quine, W.V. (1966), *The Ways of Paradox and Other Essays*, Random House, New York.

--Quine, W.V. (1969), 'Existence and Quantification', in *Fact and Existence*, ed. J. Margolis, Basil Blackwell, Oxford, 1969.

--Ramsey, F. P. (1925), 'Universals', in Ramsey (1978).

--Ramsey, F.P. (1978), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, ed. D.H. Mellor, Routledge & Kegan Paul, London and Henley.

--Rayo, Augustin & Stephen Yablo (2001), 'Nominalism Through De-Nominalization', *Noûs* 35:1, 2001.

--Resnik, Michael D. (1985), 'Ontology and Logic: Remarks on Hartry Field's Anti-platonist Philosophy of Mathematics', in *History and Philosophy of Logic*, 6, 1985.

--Resnik, Michael D. (1998), 'Holistic Mathematics', in Schirn (1998).

--Reynolds, Steven L (2006), 'Realism and the Meaning of "Real"', *Noûs*, 40:3, 2006.

--Rosen, Gideon (1990), 'Modal Fictionalism', *Mind,* Vol. XCIX, No. 395, July 1990.

--Russell, Bertrand (1903), *The Principles of Mathematics*, Routledge, London. Reprinted 1997.

--Russell, Bertrand (1923), 'Vagueness', in Keefe & Smith (1999).

--Schirn, Matthias (1998), ed., *The Philosophy of Mathematics Today*, Oxford University Press, Oxford.

-- Shapiro, Stewart (1997), *Philosophy of Mathematics: Structure and Ontology*, Oxford University Press, Oxford.

--Sober, Elliott (1993), 'Mathematics and Indispensability', *The Philosophical Review*, Vol. 102, No. 1, January 1993.

--Swartz, Norman (1991), *Beyond Experience: Metaphysical Theories and Philosophical Constraints*, University of Toronto Press, Toronto.

--Tarski, Alfred (1936), 'On the Concept of Logical Consequence', in Tarski, *Logic, Semantics, Metamathematics*, trans. J H Woodger, 2nd ed. John Corcoran, Hackett, Indianapolis, 1983.

--Tharp, Leslie (1971), 'Truth, Quantification and Abstract Objects', *Noûs* 1971, pp 363-372.

--van Heijenoort, Jean (1967), ed., *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, Cambridge, Mass.

--Walton, Kendall L. (1990), *Mimesis as Make-Believe*, Harvard University Press, Cambridge, Mass.

--Wang, Hao (1961), 'Process and Existence in Mathematics', in *Essays on the Foundations of Mathematics*, ed. Y. Bar-Hillel *et al*, Hebrew University of Jerusalem, Magnes Press, 1961, 2nd edition 1966.

--Wang, Hao (1974), *From Mathematics to Philosophy*, Routledge & Kegan Paul, London.

--Wigner, Eugene (1967), *Symmetries and Reflections: Scientific Essays of Eugene P. Wigner*, Indiana University Press, Bloomington & London.

--Wittgenstein, Ludwig (1929), 'Some Remarks on Logical Form', in Ludlow (1997).

--Yablo, Stephen (1998), 'Does Ontology Rest on a Mistake?', in *Proceedings of the Aristotelian Society: Supplementary Volume LXXII*, Aristotelian Society, 1998.

--Yablo, Stephen (2000), 'Apriority and Existence', in Boghossian & Peacocke (2000).

--Yablo, Stephen (2001), 'Go Figure: A Path Through Fictionalism', *Midwest Studies in Philosophy*, XXV, 2001.

--Yablo, Stephen (2002), 'Abstract Objects: A Case Study', *Philosophical Issues, 12, Realism and Relativism*, 2002.

--Yablo, Stephen (2005), 'The Myth of the Seven', in Kalderon (2005).

--Zumdahl, Steven S. (1998), *Chemical Principles*, 3rd edition, Houghton Mifflin Co., Boston.