Search in this book

CHAPTER

12 On the Philosophical Significance of Frege's Theorem 🔓

Crispin Wright

Abstract

Frege's theorem is the result that elementary arithmetic may be derived from the second- order sentence often referred to as 'Hume's Principle' (HP). In this essay, Wright considers whether 'abstraction principles' of this kind can give rise to a distinctively logicist philosophy of mathematics; in particular, he considers whether HP can underpin a satisfactory epistemology of arithmetic. After outlining the neo-Fregean project as it arises out of Frege's Grundlagen, Wright moves to consider the bad company objection: the distinctive abstractive form of HP is shared by numerous other second-order sentences, some of which (such as Frege's Basic Law V) are inconsistent. After examining various forms of this challenge and potential responses to them, Wright recommends that the neo-Fregean restrict attention to those abstraction principles which are conservative. The final sections of the essay survey prospects for extending the neo-Fregean story to other areas of mathematics.

Keywords: abstraction principle, Boolos, conservativeness, Dummett, Frege's theorem, Hume's Principle, logicism, mathematics

Subject: Philosophy of Language, Philosophy of Mathematics and Logic

Collection: Oxford Scholarship Online

My interest in Frege's philosophy of mathematics was kindled by studying *Grundlagen* as a visiting graduate student in Oxford under Michael Dummett's supervision in 1966–7. I don't think I had realized before that year how deep analytical philosophy can reach. The experience decided me to become a professional philosopher if I could. However, my debt in thinking about these matters is, of course, not just to Michael's teaching of thirty years ago, but to his encouragement, writings, and general example ever since. I would have liked to avoid the apparent ingratitude of contributing an essay to his Festschrift with which he will largely disagree! He will agree, anyway, about the interest and importance of the larger debate.

The essay which follows pursues the sympathetic reconstruction of Frege's philosophy of arithmetic embarked on in my *Frege's Conception of Numbers as Objects*.[1] It is mainly concerned with an intriguing line of difficulty for the neo-Fregean first clearly perceived by another of Michael's former students, George Boolos. As his comments in the present volume illustrate, I never did manage to persuade George that something akin to logicism about arithmetic might yet fly. But my interest in these matters has prospered in recent years under his sceptical prompting and I owe a great debt to his cognate researches into Fregean arithmetic and the attendant philosophical issues. His early death is a serious blow to contemporary philosophy of logic and mathematics. We shall sorely miss his considerable intellectual presence and collegial spirit.

I

In §§70–83 of Grundlagen, Frege gives an informal sketch of reasoning which, although he does not explicitly signal the point,[2] shows how the standard axioms for arithmetic can be derived within a suitable second-order logic with identity[3] to which what is nowadays often called *Hume's Principle*—informally:

> For any concepts F and G, the number of Fs is the same as the number of Gs if and only if F and G can be put into one–one correspondence,—

has been adjoined as sole additional axiom.[4],[5] Moreover the system consisting of such a logic and Hume's Principle can be shown to be consistent relative to second-order arithmetic.[6] So, save in neurotic moments, there should be no doubt that notwithstanding the inconsistency of the larger system of *Grundgezetse*, Frege did at least establish a new *mathematical* foundation for number theory: a subsumption of the laws for finite cardinal numbers under a single principle articulating the condition for the identity and distinctness of cardinal numbers in general. The Dedekind–Peano Famous Five can be reduced to One. Following a happy suggestion of Boolos,[7] I shall call this result *Frege's Theorem*.

My question here is what is the *philosophical* significance of Frege's Theorem? Sure, there is no general reason why such a mathematical simplification should carry any philosophical significance at all. But Hume's ↳ Principle, recall, is originally introduced in *Grundlagen* [8] as a *contextual definition* of numerical identity; and the putative drawback which causes Frege apparently to push it aside and move instead to explicit definitions of numerical singular terms—essentially, its vulnerability to the 'Caesar Problem': its inability to decide the truth-values of mixed identity-statements like 'The number one is identical to Julius Caesar'—is at worst a shortcoming of *insufficiency* rather than incorrectness. Hume's Principle, for all Frege says to the contrary in *Grundlagen*, might therefore be conceived as a correct *partial* definition of the use of certain numerical expressions. And if it is at least a correct partial definition, then Frege's Theorem shows that the fundamental laws of arithmetic—the Peano axioms—are logical consequences of a definition. Would not that be a vindication of a species of *logicism* about number theory?

We need some distinctions. Most stringently conceived, number-theoretic logicism ought to be the thesis that every truth of number-theory is a definitional transcription of a truth of logic—so that the use of what is ordinarily recognized as distinctively arithmetical vocabulary is inessential for the expression of such truths. There is no doubt that Frege's Theorem makes no case for such a view. Hume's Principle provides the resources for a transcription of arithmetical truths into what are uncontroversially truths of second-order logic only in one very special kind of case, viz. when the truths concerned are identity-statements linking a pair of numerical singular terms each formed by applying the numerical operator to an open sentence which itself contains no occurrences of that operator and can be expressed in second-order logic.

Take a simple example. Consider the statement

(a)   Nx: x ≠ x = Nx: x ≠ x

—the statement that the number of things which are not self-identical is identical to itself. Hume's Principle, taken as a contextual definition, will allow us to transform this into a theorem of pure second-order logic, to the effect that there is a one–one correspondence between the instances of the concept, not self-identical, and themselves.[9] But now consider the existential generalization of (a),

(b)   (∃ y) (y = Nx: x ≠ x).

This is an immediate consequence of (a). But its singleton occurrence of the numerical operator is one which the resources afforded by Hume's Principle provide no means to eliminate.

However, must the failure of the most stringent version mark the failure of number-theoretic logicism's principle thesis— that the laws of arithmetic lie within the deductive closure (or among the logical consequences) of (what can properly be regarded as) logic? If Hume's Principle permits us to regard (a) as a definitional equivalent of a truth of higher-order logic, then (a) is presumably itself a truth of logic. Since whatever follows from a truth of logic is likewise a truth of logic, we have it that (b) is a truth of logic too, notwithstanding the absence of any means to paraphrase it into purely logical vocabulary as ordinarily understood.

But *this* thought confronts what may seem to be a fatal objection. Suppose, what is presumably a perfectly legitimate option, we interpret Hume's Principle as genuinely *defining* the sentence (a) in terms of the second-order theorem which it gives as an equivalent; then on this interpretation the inference from (a) to (b) is unjustified, since (b) requires, what (a), so interpreted, does not, that the universe be non-empty.[10] Alternatively, suppose we accept the inference from (a) to (b) as a routine existential generalization; then in that case (a) has a content which, it appears, it cannot have acquired by being stipulated to be equivalent to the relevant second-order theorem. What we can't do is to adopt both points of view simultaneously, taking it *both* that (a) is defined by Hume's Principle *and* that it sustains the inference to (b).

This is a natural and very familiar type of objection. And it brings us up against one of the most fundamental and fascinating of the issues raised by the philosophy of mathematics outlined in *Grundlagen*—the status of the kind of abstraction principle of which Hume's Principle is one illustration. By way of a general characterization: we are speaking

p. 276   of principles which ↳ introduce a unary term-forming operator, Σ, on some familiar kind of expressions, $\alpha_1 \ldots \alpha_k \ldots$,— these may be singular terms, or first-order or even higher-order predicates (or relational expressions)—by fixing the truth-conditions of identity statements of the form, $\Sigma(\alpha_k) = \Sigma(\alpha_j)$, by reference to the obtaining of some equivalence relation, $\approx$, among the items denoted by $\alpha_1 \ldots \alpha_k \ldots$ etc. Call any such principle an *abstraction*.[11] In a first-order abstraction, like Frege's famous example of direction:

> The direction of line a is identical to the direction of line b if and only if a and b are parallel,

the new operator, 'The direction of . . . ', is applied to terms standing for objects. In Hume's Principle, by contrast, the operator, 'The number of . . . ', is applied to expressions standing for first-level concepts, and the relevant equivalence relation—henceforward the *abstractive relation*—is correspondingly second level. And in principle, such an abstraction may be of any order.

Accordingly, the objection just rehearsed comes, in its most general form, to the thought that, whatever else is said, no such principle,

$$(\forall \alpha_i)(\forall \alpha_j)(\Sigma(\alpha_i) = \Sigma(\alpha_j) \leftrightarrow \alpha_i \approx \alpha_j),$$

can be merely *stipulatively* true if taken in the manner intended. For to construe the syntax of its left-hand side at face value, as intended, so that genuine singular terms are involved, flanking what is genuinely an expression of identity, is to read into it an existential commitment which the righthand side—the statement of the abstractive relation's holding among the prior items—does not implicate. And the problem for one who would try to elicit some form of logicist moral from Frege's Theorem is that if Hume's Principle really is taken as a definition—something we may simply stipulatively lay down, without further epistemological presupposition—then nothing has been done to justify the kind of inference— that from (a) to (b) providing the simplest possible illustration—which we need to be able to make in order to found arithmetic upon it. But if it is interpreted in such a way as to sustain such inferences, then a substantial epistemological presupposition does appear to be made: we need to know that there really are the objects—cardinal numbers—which the truth of instances of Hume's Principle demands, and we need to know this, it seems, in some fashion independent of reliance on Hume's Principle.

Frege's discussion of Hume's Principle in *Grundlagen* moves very briskly to the Caesar problem and the apparent disqualification of the Principle from any foundational role—too briskly for any context to be established in which he

p. 277   might have had cause to confront this kind of worry. Nevertheless we can ↳ glean something, I believe, of how he might have responded to it. The key passage runs:

> 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 being identical. Admittedly this seems to be a very odd kind of definition, to which logicians have not yet paid enough attention; but that it is not altogether unheard of, may be shown by a few examples. (64) The judgement 'line a is parallel to line b', or, using symbols,
>
> a//b,

> can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'The direction of line a is identical with the direction of line b'. Thus we replace the symbol // by the more generic symbol =, through removing what is specific in the content [*Inhalt*] of the former and dividing it between a and b. We carve up the content in a way different from the original way, and this yields us a new concept.[12]

There is much to be said about this passage. But the vital point for our present purpose is that, in the metaphor of 'carving up' contents in different ways, Frege points to a way of receiving an abstraction which, if it is legitimate, completely defuses the stated problem. Consider again the abstraction for directions:

$$\mathrm{Da} = \mathrm{Db} \text{ if and only if } \mathrm{a}//\mathrm{b}.$$

The dilemma was that we either regard the left-hand side simply as a definitional transcription of the right, and thereby forfeit the possibility of taking its syntax at face value, of treating it as a genuine identity statement linking genuine singular terms in existentially generalizable position; or we take the principle as a substantial claim, to the effect that certain abstract objects—directions—are associated with lines in the way it describes, in which case we have no right simply to lay the principle down as a definition. But the key to Frege's view is that the dilemma is a false one—it is the thought, roughly, that we have the option, by laying down the Direction abstraction, of *reconceptualizing*, as it were, the type of state of affairs which is described on the right. That type of state of affairs is initially given to us as the obtaining of a certain equivalence relation—parallelism—among lines; but we have the option, by stipulating that the abstraction is to hold, of so reconceiving such states of affairs that they come to constitute the identity of a new kind of thing, directions, of which, by this very stipulation, we introduce the concept. The concept of direction is thus so introduced that that two lines are parallel *constitutes* the identity of their direction. It is in no sense a further substantial claim that their directions exist and are identical under the described ↳ circumstances. But nor is it the case that, by stipulating that the principle is to hold, we thereby forfeit the right to a face-value construal of its left-hand side, and thereby to the type of existential generalization which a face-value construal would license. When the abstraction principle is read in the way which Frege proposes, its effect is so to fix the concept of direction that there is absolutely no gap between the existence of directions and the instantiation of properties and relations among lines.

It is important to be clear that it would be a misrepresentation of this idea to view it as involving the notion that abstract objects are creations of the human mind, *brought into being* by a kind of stipulation. What is formed—created—by such an abstraction is rather a *concept*:[13] the effect is merely to fix the truth-conditions of identity statements concerning a new kind of thing, and it is quite another question whether those truth-conditions are ever realized. If we accept the concept-formation involved in the Fregean abstraction of Direction, the effect is not to define directions into existence but to coordinate the question of the existence of directions with that of the existence of lines; and the latter can remain, for all that is implicit in an acceptance of the abstraction, as objective and mind-independent a matter as you want. What is true is that when, as in the case of Hume's Principle, the abstraction is second order, it may happen that certain instances of the right-hand-side relation hold a priori, or even of logical necessity. And when that happens, the corresponding abstract objects will also exist a priori, or of necessity. This line of thought has struck some philosophers[14] as about as bad, in much the same way, as Anselm's Ontological Argument for the existence of God. But there is no real parallel. In particular, it is not, to stress, correct to regard the existence of numbers as generated purely out of the concept of number, as the existence of God is supposed, according to the Ontological Argument, to be generated out of His concept. Numbers are, rather, like directions, the output of a distinctive kind of *reconceptualization* of an epistemologically prior species of truth. Nothing like that happens in the Ontological Argument.[15]

Frege's idea is one of immense importance if it can be sustained. Abstract objects are sometimes thought of as constituting a 'third realm', a sphere of being truly additional to and independent from the concrete world of causal space-time. It is this conception of the abstract which generates the well-known ↳ epistemological problems to which nominalism and various forms of reductionism and structuralism attempt to respond.[16] But if anything like—or to the extent that—the Fregean conception can be sustained, there can be no epistemological difficulties posed by thought about and knowledge of any particular species of abstract objects which are not already present on the right-hand side of the abstraction which initially introduces the concept of that species. That is not to say that all such problems disappear. But it is a tremendous gain in manageability. And the blanket idea, that there has to be a general problem about thought and knowledge of *abstracta*, just in virtue of their being abstract, is quite undermined.

p. 278

p. 279

Suppose we grant that abstraction principles may in general legitimately be viewed in this kind of way, and that Hume's Principle is a legitimate special case. What in that case does Frege's Theorem show about the status of the laws of arithmetic? Does it after all establish a species of logicism? Must Hume's Principle itself be made out to be either a truth of logic or a strict definition if a version of arithmetical logicism is to be made good?

If so, then the prospects are unimproved. At any rate, it would be plausible to regard the limits of second-order logic as crossed just at the point where substantial principles concerning specific types of object begin to feature. And it would be plausible to regard the limits of strict definition as crossed just at the point where the statement of the definition is allowed to imply sentences containing occurrences of the supposed *definiendum* for which no eliminative paraphrase is provided. Even so, Frege's Theorem will still ensure the truth of the following: that the fundamental laws of arithmetic can be derived within a system of second-order logic augmented by a principle whose role is to *explain*, if not exactly to define, the general notion of identity of cardinal number; and that this explanation proceeds in terms of a notion which can be defined in terms of the concepts of second-order logic. If such an explanatory principle, in company with 'implicit definitions' generally, can be regarded as *analytic*, then that should suffice at least to demonstrate the analyticity of arithmetic. Even if that term is found troubling, as for instance by George Boolos,[17] it will remain that Hume's Principle— like any principle serving implicitly to define a certain concept—will be available without significant epistemological presupposition to one who has mastery of the other concepts it configures. So one clear a priori route into a recognition of the truth of its (recognizable) consequences—in particular, the fundamental laws of arithmetic—will have been made out. And if in addition the Principle may be viewed as a *complete* explanation—as showing how the concept of cardinal number may be fully understood on a purely logical basis—then arithmetic will have been shown up by Frege's Theorem not as ↳ part of logic, it is true, but as transcending logic only to the extent that it makes use of a *logical* abstraction principle—one whose right-hand side deploys only logical notions. So, always provided that concept-formation by abstraction is accepted, there will be an a priori route from a mastery of second-order logic to a full understanding and grasp of the truth of the fundamental laws of arithmetic. Such an epistemological result, it seems to me, would be an outcome still worth describing as logicism, albeit rather different from the conventional, more stringent understanding of the term.[18]

Reflect, at any rate, that Frege's intention for *Grundgezetse* can hardly have differed from the conclusion just mooted except in terms of generality. For Basic Law V is just a more general logical abstraction. Had this system been consistent—if such a counterfactual may be permitted—there would still have been a gap—for one inclined to see one—between second-order logic strictly so regarded and second-order logic augmented by the very general abstraction incorporated in Basic Law V. So, prima facie, the philosophical significance of Frege's Theorem cannot be less than this: that, at least as far as number theory is concerned, the more extensive epistemological programme which Frege hoped to accomplish in *Grundgezetse* is still a going concern.[19]

II

A number of fundamental misgivings about these neo-Fregean thoughts are expressed in Michael Dummett's recent book.[20] Chief among them are

(i)   concerns about the Caesar Problem;

(ii) doubts about the notion of reference which we are justified in ascribing to the singular terms which abstractions serve to introduce;

(iii) misgivings about the sense, if any, in which an ordinary understanding of the type of context that features on the right-hand side of an abstraction can be said to implicate the type of existential commitment that becomes transparent on the left; and

(iv) most importantly, the belief that the typical and necessary[21] impredicativity of second-order abstractions—the fact that objects of the newly introduced kind are treated as already lying within the range of the predicates and quantifiers that feature on the right-hand side of ↳ the abstraction;—the belief that this impredicativity completely vitiates the Fregean enterprise.

I believe that strong replies are available to each of these objections, but I will not be further concerned with them here.[22] Rather, I want to consider a distinct line of objection, also emphasized by Dummett, but which has been prominent in a

p. 280

p. 281

more developed form in recent independent discussions of George Boolos and Richard Heck.[23] Boolos, in particular, has been severely critical of the philosophical interest of abstraction principles, and his criticisms have, to the best of my knowledge, gone so far unrewarded by a Fregean response.

The basic line of objection is what we might term a *Bad Company* objection. Its gist is that the claims the Fregean wants to make on behalf of Hume's Principle will be seen to stand in need of, at best, major qualification when one reflects on the characteristics of other instances of the type of abstraction it exemplifies Basic Law V, to take the most immediate case, is, as we have noted, an abstraction of essentially exactly the same kind. Yet it is inconsistent. Other intuitively innocent-looking examples are in the same case.[24] Does there not have to be error in the Fregean conception of abstraction if it sanctions such radically unsatisfactory examples?[25]

Now, I do not think that the objection, taken just to this point, is impressive at all. Or better: if someone finds it impressive, it is probably because they are not thinking about the status of abstraction principles in the right kind of way. Such a principle, remember, is not an attempt to state an antecedently determinate truth. It is intended, rather, to fix a new concept by determining the truth-conditions of certain statements involving it. What of it, then, if certain such attempts—certain instances of the general shape—misfire, enjoining no consistent practice with the concept in question? Who ever doubted that explanations can misfire, that one may inadvertently pack inconsistency into a ↳ definition, for example, or a subtle clash into the rules of a newly invented game? The question has to be: do the particular failures show that the *whole genre* of purported explanations which they represent are actually abortive? And, unless inconsistency is taken to be a symptom of some deeper malady affecting even consistent examples of the genre,[26] the answer ought to be negative. Pending development of the objection in that direction, then, it is, for the Fregean's purposes, enough of a disanalogy between Hume's Principle and Basic Law V that the former is, as we observed, effectively known to be consistent. Indeed, *knowledge* of consistency is arguably more than is required. Explanations which seem to work well enough should surely be regarded as innocent until proved guilty. So it will be enough of a disanalogy if there is no extant reason to doubt the consistency of a second-order abstraction if the usual lines to contradiction do not succeed in its case.[27] To think otherwise is implicitly to suppose that we should look askance at any explanation, however apparently effective, whose general shape is one it is possible to exploit by inconsistent instances. I'll return to this theme below.

III

The literature illustrates, however, another readily foreseen and very simple kind of troublesome case.[28] Where R(F, G) is an equivalence relation on F and G, and P is any closed sentence,

$$P \lor R(F, G)$$

will likewise be an equivalence relation. So

$$(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow P \lor R(F, G))$$
$$(A)$$

will, for all we have said to the contrary, count as a possible abstraction. Since P is closed, it will not contain occurrences of 'F' or 'G' within the scope of A's initial quantifiers. So (A) will entail by routine logical moves:

$$\neg P \to (\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow R(F, G)).$$
$$(B)$$

[29] ↳ But the consequent of (B) is itself an abstraction. We have merely to reflect, then, that should this embedded abstraction be one of the inconsistent cases, like Basic Law V, then (B), and hence (A), will entail P.

Now P, recall, was any closed (second-order) sentence. So P could be, for instance, to the effect that there is a concept of Dedekind infinite extension; or that there is not; or that there are exactly six objects altogether; or uncountably many—all of which are second-order expressible claims. The neo-logicist wants to say that we may freely undertake the concept-formation embodied in Hume's Principle, and thereby be led to *the discovery* that the series of natural numbers, and hence the universe, is infinite. But now it appears a sceptic—Boolos—can retort that such 'freely undertaken' abstraction can lead us to the 'discovery' of the truth of *any* (second-order) P—which is just to say that there is something seriously amiss with the Fregean view of Hume's Principle.

Of course, this sceptical challenge will abort unless (A) can be a consistent abstraction even when the embedded abstraction—

$$(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow R(F, G))$$

—is not. But that reflection need give the sceptic no problem. Stipulate that, for any F, ΣF is to be some specific object—perhaps the null set—in the domain of the first-order quantifiers. Then (A) will hold true whenever P does, and hence will be satisfiable if P is.[30] We have to conclude that it is *not* sufficient for the acceptability of a second-order abstraction that it works with a well-defined equivalence relation and be consistent. The Fregean owes a distinction between these cases, whose implication of results about the cardinality of the domain is of no interest whatever, and cases like Hume's Principle whose implication of such results is, on the Fregean view, of the greatest interest.

IV

Can the needed distinction be drawn? It's natural to feel that the troublesome abstractive relations are a kind of cheat: that they merely point up the need to restrict abstraction to relations which are, as it were, *real* relations, rather than artefacts of language. Just as *your* having blue eyes is not, in any interesting sense, a real property of mine, so P—if it makes no mention of F or G—need not, in the intended spirit, involve a real condition on F and G.

p. 284   Of course, this kind of contrast is notoriously difficult to substantiate. But suppose it made out somehow; and suppose we insist that the characteristics which are ingredient in a proposed abstractive relation must be real—contrast: merely *Cambridge*—properties and relations,—that the holding, or not, of a bona fide abstractive relation between F and G can never depend on considerations which do not concern real characteristics of F and G. Then we would certainly expect often to be able to outlaw the type of equivalence relation involved on the right-hand side of instances of (A). For in case that not R(F, G), F and G may cease to exemplify the relation just by dint of a change from truth to falsity by P, which change need implicate no change in the facts concerning exactly what things are, or are not, F, or G, or anything else which ought intuitively to count as a real characteristic of those concepts.

The sceptic about abstraction has better weapons than schema (A), however. Let φ be any property of concepts for which coextensiveness is a congruence.[31] Then

$$\varphi F \& \varphi G \quad \vee \quad (\forall x)(Fx \leftrightarrow Gx)$$

is an equivalence relation.[32] It can also be, presumably, for very many choices of φ, a bona fide, non-Cambridge relation, each of its disjuncts depending, in what seems the requisite intuitive sense, on real characteristics of the related concepts. Such choices could include: has exactly six instances, has infinitely many counter-instances, is true of no elephant, is true of most singers, and so on. Consider therefore this schematic abstraction:

$$(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow (\varphi F \& \varphi G) \vee (\forall x)(Fx \leftrightarrow Gx)).$$

(D)

Any instance of (D) will entail

$$(\exists F)\varphi F.$$

p. 285 [33] ↳ and φ, remember, can be any second-order expressible property[34] which is intuitively a real characteristic of concepts. So it could be, for instance, 'is Dedekind infinite or has a Dedekind infinite complement', or 'is finite and has a finite complement', or 'has exactly six things which fall under it or under its complement', and so on. But the possession by any concept of a property of these kinds is tantamount to some claim about the cardinality of the universe. So judiciously selected instances of (D) will apparently generate just the same range of *ad hoc* resolutions of questions to do with the cardinality of the universe as (A), even when all components of their abstractive relations are bona fide, non-Cambridge characteristics.

What of consistency? Instances of (A), we noted, will be satisfiable if P is. The same line of reasoning will deliver that an instance of (D) will be satisfiable in case *every* F has the relevant property φ: again, merely stipulate that for each F, Σ F is to be the same favoured object selected from the domain. But now reflect that, for a wide class of unwelcome interpretations of φ, all concepts will be φ if any are; this holds, in particular, of all the interpretations of the kind latterly illustrated. So one should have no doubt about the consistency of the relevant instance of (D) in any case where φ so corresponds to what one considers to be a possible state, in point of cardinality, of the universe.

Can the Cambridge proposal—as we may call it—be refined to handle this troublesome class of interpretations of (D)? Here is one ploy that would work thus far. Suppose some range of properties of things of a given kind are *intrinsic* to them. And characterize a Cambridge property as one whose instantiation by such things does not supervene upon their intrinsic characteristics—so that an item may vary in a Cambridge respect across worlds in which its intrinsic properties are the same. Now plausibly, when concepts are viewed extensionally—as Fregean Concepts—their extensions—the actual range of items which instantiate them—are such an intrinsic feature of them. Indeed, it is hard to see that they have any other intrinsic feature. In this ↳ respect Fregean Concepts are like sets. So we may say that a property, or relation, on Concepts is merely Cambridge if it is possible for a Concept, or n-tuple of Concepts, which exemplify it to cease to do so without change in their intrinsics, that is, in their actual extensions.

p. 286

By this standard, the abstractive relations involved in instances of (D) when φ is taken to be a property like 'is finite and has a finite complement', or 'has exactly six things which fall under it or under its complement', etc., will standardly be Cambridge. For if both φ F and φ G hold but F and G are not coextensive, then they may cease to exemplify the relevant abstractive relation just by dint of one or the other ceasing to be φ; and that change need implicate no change in which things are F, or G but may be constituted just by a change in which things are *not*—a change in the extensions of one or both of their respective complements.

So that is one possible tactic for the Fregean. Any stable gain from it would have, of course, to depend on the possibility of showing the *relevance* of that real–Cambridge distinction in the present context; i.e. of showing that equivalence relations branded by such an account as merely Cambridge can be *independently* appreciated to be unsuitable for abstraction. Perhaps something could be accomplished in that direction—by amplifying the thought, maybe, that it is only when one is concerned with *real* relations on the Concepts in question that it is appropriate to believe in the *reality* of the abstracted objects. Let's put the matter on ice for the moment.

V

There is a second, quite different line of response. The Fregean proposes to regard abstraction merely as one legitimate form of concept formation. The objection, as it stands at present, is to the effect that (second-order) abstractions may not be so regarded if, for all we have said so far about what an abstraction principle is, examples can be given which have unforeseen, undesired, and mutually contradictory contingent consequences. What the Fregean may reply is, as in the case of the objection from the inconsistency of Basic Law V *et al.*, that it is not merely concept-formation by abstraction which is open to this kind of difficulty. For other, much less controversial principles for concept or sense-fixing are familiarly prone to the same sort of abuse.

Consider for instance the principle that we may fix the sense of a sentence by stipulating that it is to have a certain truth-condition. Suppose we stipulate that 'P' is to be a sentence whose truth-condition is that if the sentence inscribed on the

last page of my notebook is true, then the universe is finite. That may seem a perfectly determinate truth-condition. But of course I can now mischievously inscribe 'P' itself (and only 'P') on the last page of my notebook. And once I've done so, the truth-condition for 'P' becomes:

> If 'P' is true, then the universe is finite.

Whence, given only that any sentence with determinate truth-conditions will satisfy the disquotational scheme:

> (DS) 'S' is true if and only if S,

we have all the materials we need, following the reasoning of Curry's well-known paradox, for a proof of the finitude of the universe.[35] Since we are also apparently free to stipulate that 'Q' is to be a sentence whose truth-condition is that

> If the only sentence written on the penultimate page of my notebook is true, then the universe is infinite,

only my refraining from a similar piece of mischief stands between our stipulations and the proof of a contradiction.

Now, it would surely be absurd to conclude that the idea that content can be given to a sentence by assigning it what appears to be a well determined truth-condition is simply mistaken. A balanced response will grant that the principle may stand in need of some qualification—for instance, that not every apparently well-determined truth-condition need be so for all circumstances—but insist that, *in advance of* our having a generally satisfactory restriction, there is nevertheless no reason to doubt the principle in the ordinary run of cases.

Or again, consider the principle that we may introduce a property of objects of a certain kind by stipulating under what conditions such objects possess that property. Such a principle is in difficulty as soon as we consider the property of predicates,

> . . . is not true of itself,

and proceed to form the concept *heterological* in the well-known manner:

> (∀ F) (F is heterological if and only if F is not true of itself).

If somebody proposed that the principle needed refinement just to exclude the case where inconsistency results, they would be confounded by the following stipulation:

> (∀ F) (F is *innocently heterological* if and only if the universe is finite or F is not true of itself).

For we immediately obtain

> 'Innocently heterological' is innocently heterological if and only if the universe is finite or 'innocently heterological' is not true of itself,

which will straightforwardly generate, once again, a proof of the finitude of the universe. Yet we should not, presumably, conclude that the whole idea that properties can be defined by the stipulation of satisfaction conditions is misbegotten. Rather, some kind of restriction is wanted. Nor, crucially, pending such a restriction, should we suspend judgement about what appear to be perfectly innocent examples of such a procedure. The sensible response is rather that there is a distinction to be drawn which we are, perhaps, not clear how exactly to draw but which, once drawn, will safeguard the vast majority of cases where we fix a property by stipulating satisfaction conditions.

Why should it be any different with Fregean abstraction? We have known for a long time how, in the sort of way illustrated, to use the semantic paradoxes to make trouble for what appear to be innocent forms of definition and concept introduction, either directly by generating contradictions or indirectly, by so insinuating a paradox into an instance of such a form as to generate a consistent but utterly arbitrary and inappropriate resolution of a contingency (the classic example is the Cretan). If Fregean abstraction is a legitimate mode of concept-formation, it is only to be expected that it too will be vulnerable to exploitation in such fashion.

That reflection does not excuse the Fregean the work of offering some principled restriction whereby instances of (D) may be excluded, just as the procedure of definition of predicates by satisfaction conditions needs some restriction to exclude both the heterological paradox and the kind of disjunctive exploitation of it illustrated. But it is in the interest of clarity and, as it were, intellectual good housekeeping that such work is wanted. No doubt should be raised by these examples about the *very possibility* of defining a property by giving it satisfaction-conditions. By the same token, one initially disposed to accept Fregean abstraction as a legitimate mode of concept-formation has no reason to regard the existence of directly, or 'disjunctively' paradoxical cases as somehow casting doubt over the entire principle.

VI

Now we come to the real action. The foregoing are all points and moves which it is as well to bring into the open. But it would be premature to conclude that the balance of the game is tending towards the Fregean. For there is, unfortunately, another class of cases: a range of consistent, second-order abstractions which involve nothing analogous to the kind of disjunctive exploitation of paradox in (A) and (D), nor any play with Cambridge ↳ relations, yet still arguably yield results inconsistent with Hume's Principle. The lines of response so far canvassed on the Fregean's behalf are quite inadequate to handle this remaining class.

p. 289

It ought all along to have seemed an open possibility that such awkward cases might arise. We are concerned with principles which take us from equivalence relations on concepts to identity statements concerning certain associated kinds of object. These principles in effect partition the universe of concepts into cells—equivalence classes under the relation in question—and then call for a one–one correlation between these cells and the objects yielded by the abstraction, hence a mapping of the cells into the original population of objects. Now as we have already had occasion to remember, Frege's Concepts are extensional—are, save in point of 'saturation', just like *subsets* of the original domain of objects. By Cantor's Theorem, there will therefore be 2^k such Concepts if there are k objects in the original domain. *One* problem with Basic Law V is that its abstractive relation—coextensiveness—partitions the universe of Concepts into singletons. Since 2^k is always greater than k, Basic Law V therefore demands more objects to serve as the referents of its left-hand-side singular terms than can possibly be provided by the original domain.[36] There is therefore going to be a general constraint on second-order abstractions to the effect that the cells into which the relevant equivalence relation partitions the universe of Concepts must not outrun the population of objects which constitute the range of the first-order variables in the abstraction principle—at least, not if we want the new kind of objects to lie within that range.

One way of understanding the proof of the infinity of the number series that Hume's Principle generates is to reflect that this constraint is violated if we take the original population of objects to be finite, since one–one correspondence will generate k + 1 cells of Concepts on any population of objects of cardinality k. That is no problem if, but only if, k is infinite. So Hume's Principle is satisfiable only in infinite domains. But now the risk is evident: might there not be (non-Cambridge) second-order definable equivalence relations on Concepts whose effect is just the reverse—equivalence relations the (non-paradox-embedding) abstraction based on which is satisfiable in any *finite* domain, but which generate too many cells of Concepts as soon as the population of objects is taken to be infinite?

And the answer is, apparently, 'yes'. Consider the relation on F and G,

Only finitely many objects are either F-and-not-G or G-and-not-F, and write this as

$$\Delta(F, G)$$

p. 290 $\Delta(F, G)$ is an equivalence relation.[37] It is also, at least by the lights of the earlier proposal, a non-Cambridge relation.[38] Moreover finitude is a second-order definable property. So $\Delta(F, G)$ would seem to be a bona fide—non-Cambridge, non-paradox-embedding—second-order equivalence relation, inviting the abstraction of a range of objects which I will call 'nuisances' and designate by terms of the form, '$\nu(F)$':

$$(\forall F)(\forall G)[\nu(F) = \nu(G) \text{ iff } \Delta(F, G)].$$

Now suppose the universe of objects to be finite. Then all Concepts are finite too, so $\Delta(F, G)$ will hold universally. Hence we can ensure the satisfaction of the Nuisance Principle in any finite domain simply by stipulating, once again, that $\nu(F)$ is, for each Concept F, to be some particular, favoured object.

Suppose by contrast that the universe of objects is infinite, specifically, that it is of infinite cardinality k. How big are the cells into which $\Delta(F, G)$ partitions the associated Concepts? That is: for any particular F, how many distinct Gs will there be such that

$$(F\&\neg G) \vee (G\&\neg F)$$

is finite? Two lemmas are now germane. First, any domain of objects of infinite cardinality k will generate exactly k *finite* Concepts.[39] A fortiori, ↳ there are at most k finite Concepts of the particular form,

$$(F\&\neg G) \vee (G\&\neg F).$$

Next, consider any particular such Concept. Could this be the same (i.e. true of the same objects) as

$$(F\&\neg H) \vee (H\&\neg F)$$

even though G and H are distinct Concepts, i.e. have different objects fall under them? The second lemma says no.[40] Thus for any particular F, there are at most k Concepts of the form,

$$(F\&\neg G) \vee (G\&\neg F)$$

which are finite; and no pair of these Concepts coincide unless their values for 'G' coincide. It follows that for any particular F, the cell of Concepts constructed around it by $\Delta(F, G)$ contains at most k elements.

And that spells trouble. Reflect that, since the domain of objects is k-fold, there are at most k nuisances. And each nuisance is, by the reasoning just sketched, the nuisance of at most k Concepts. So at most k times k, = k, Concepts get to be assigned a nuisance. But the Nuisance Principle demands that *every* Concept has its nuisance. Since there are 2^k Concepts altogether, and since 2^k is always greater than k, it follows that the Nuisance Principle, although satisfiable in finite domains, is unsatisfiable if the original population of objects is infinite.

Hume's Principle is not, as we have noted, a definition. But if, as the Fregean proposes, it is akin to a definition in serving merely to fix a new concept, then its truth should be a trivial consequence of the nature of the concept so fixed, together with the other concepts which it configures, and the principle ought therefore to rank, if not strictly as a logical truth, then at least as analytic—as a conceptual truth. But if that is right, then the Nuisance Principle ought, *pari passu*, to rank as analytic of the concept of nuisance. And analytic principles cannot be inconsistent with each other, can they? So which is the analytic truth in this case? Failing some breakage in the analogy, ought we not to conclude that neither is? Thus, essentially, runs Boolos's clever objection.[41]

VII

There may seem to be a simple if rather flat-footed reply. If we are to regard abstraction principles as *formative* of the concepts they introduce, then the analyticity of such a principle is consequent on our assigning to it just such a role. If, therefore, we assign such a role to Hume's Principle, then we have so organized our concepts that the Nuisance Principle is *analytically false*. That is: if it is in the nature of the concept of number, as fixed by Hume's Principle, that it has infinitely many instances, then it is a conceptual, hence analytic, truth that there are infinitely many objects; and any principle which entails that there are not is consequently false on purely conceptual grounds.

This line of reply, however, carries a high cost. Imagine that we have so far adopted neither Hume′s Principle nor the Nuisance Principle. Then, for all that has so far been said to the contrary, we have the option of assigning to the Nuisance Principle the status which the Fregean wants to give to Hume′s Principle. And then we shall have so organized our concepts that it follows that the universe is finite—and it will be the *finitude* of the universe that is the analytic truth. No doubt someone could take the line that that was indeed an option, although one which has since been foreclosed. But the effect of this way of looking at the matter is implicitly to surrender the Fregean idea that the infinity of the series of natural numbers, and hence of the universe, comes by way of a *discovery*. The idea, as I have emphasized repeatedly, was to be that Hume′s Principle should be viewable merely as fixing the truth-conditions of statements about numbers, whose satisfaction is then left for determination by how relevant matters independently and objectively stand. When Hume′s Principle is laid down, that is to say, as analytic of the concept of number, nothing is supposed to be happening inconsistent with the general picture that thinking can *invent concepts* under which objects may or may not fall but that it cannot in general *invent the objects* which do or do not fall under them. The sortal concepts we actually choose to employ may be only one among many possible groups; but if others are possible, then—the picture is—there already are, or are

p. 293 not, objects falling under them, whether or not we ever come to employ those concepts. The cost of the mooted response ↳ to the Nuisance objection is that it becomes impossible to sustain this kind of thinking about the abstract realm. Rather, we shall have to say that how many objects there are, and hence which objects of which kinds there are, is something which is relative to the scheme of concepts we happen to employ; so that in the abstract realm, our adoption of a particular conceptual scheme affects not merely which objects we shall *recognize* to exist, as in the concrete case, but which objects *actually* exist. That is not perhaps an incoherent view; at least, it will not be the work of an instant to show it incoherent. But it is utterly foreign to the Fregean spirit which the new logicism was supposed to safeguard.

Boolos writes:

> How to demarcate those contextual definitions that should count as logical truths in some extended sense of the expression from those that should not would seem to be a philosophical problem we have no hope of solving at present.[42]

Rephrase the problem as: how to demarcate those abstractions—or principles having the form of an abstraction, for all we have said to the contrary—that may permissibly be laid down as conceptual explanations from those that may not. Do the Nuisance Principle, and its ilk, warrant the conclusion that, as Boolos suggests, the problem is hopeless?

VIII

Somebody might be encouraged to think not by the reflection that the inconsistency between Hume′s Principle and the Nuisance Principle was elicited, not purely logically but in a manner depending on substantial assumptions about the arithmetic of infinite cardinals and its sphere of proper application. The reasoning cannot go through, for instance, without the thesis that $k^2 = k$ for *any* infinite cardinal k, proving which, in full generality, will require the unrestricted Axiom of Choice. Likewise an analogue of Cantor′s Theorem will be needed for the Concepts associated in any particular domain of objects. Any form of limitation on the scale of the domains in which Choice, or the comprehension assumptions required by Cantor′s Theorem, are to hold, will have the effect that the proof of the inconsistency of Hume′s Principle and the Nuisance Principle can be blocked, just provided one is willing to accept that the universe of objects exceeds that scale.

A related point is that the demonstration given that the Nuisance Principle cannot be satisfied in any infinite domain
p. 294 presupposes a certain conception of what the infinity of a domain of objects can consist in. ↳ Briefly: it presupposes that if a domain of objects is infinite, it will be measured by some specific infinite cardinal, k, and that 2^k Concepts (subsets) will be associated with the domain. Now, a long tradition of thought about the foundations of set theory would argue, in effect, that this cannot be the right way to think about Frege′s intentionally all-inclusive domain of objects: that Cantor′s paradox shows, in effect, that there can be no universal *set*—no absolutely all-embracing totality which is subject, for example, to the operations and principles which provide the proof of Cantor′s Theorem. That is not to say that unrestricted first-order quantification is illegitimate—that concession would be fatal to Frege′s whole project. Rather it is to say that the objects that lie in the range of such unrestricted quantification do not compose a *determinate* totality: that the totality in question is, rather, in Michael Dummett′s phrase, 'indefinitely extensible'—a totality such that any attempt to view it as a determinate collection of objects will merely subserve the specification of new objects which ought, intuitively, to lie

Downloaded from https://academic.oup.com/book/26137/chapter/194208803 by University of St Andrews Library user on 02 February 2026

within the totality but cannot, on pain of contradiction, be supposed to do so.[43] One way in which this tradition has expressed itself is in the contrast between sets and ultimate classes. But whatever we think of particular attempts to make it concrete, it is clear that to treat the reasoning from the Nuisance Principle to the finitude of the universe as conclusive is simply to suppose that there is no point to this tradition. Otherwise the option remains of concluding that the reasoning effects a *reductio* not of the supposition that the domain of the first-order quantifiers is infinite, but of the supposition that it can be assigned any *determinate* infinite cardinality. An indefinitely extensible totality will be infinite in the sense that infinite sets, of determinate cardinality, will be composable from its elements. But no such set will be comprehensive; and no determinate cardinal number will be possessed by the totality.

That then is one direction of response—to this particular example. But a cautious reader is not likely to repose much confidence in it.[44] Boolos's observation that Basic Law V is inherently inflationary[45]—that it calls for more objects to be courses-of-values on a domain than any such domain can possibly contain—is likewise dependent on model-theoretic

presuppositions and assumptions about the arithmetic of infinite cardinals. But someone ↳ innocent of Russell's paradox who imagined that Basic Law V might possibly be defended by making privileged exception to those assumptions on the basis of the indefinite extensibility of Frege's all-inclusive objectual domain would rapidly be confounded by the contradiction which it yields *purely logically*. These are matters for further study, but who is to say that the model-theoretic inconsistency between Hume's Principle and the Nuisance Principle outlined above does not likewise prefigure a purely logical inconsistency between them?[46]

And after all, Nuisances are only one example. Others may be given.[47] It is surely essential for the Fregean to have some more general form of response: some principled way of barring *any* abstraction of the relevant unwelcome kind. That may seem a tall order. But I believe there is one.

IX

Suppose that, innocent of Hume's Principle and of any conception of the infinity of the series of natural numbers, and free of any belief in infinite totalities, we *had* accepted the Nuisance Principle. In that case—let it now be granted—we would have been committed to thinking of the universe as finite, and hence of all bona fide sortal concepts as at most finitely exemplified within it. And this would require in particular that *all* sortal concepts, abstract and concrete, are at most finitely exemplified. The point will cover not just concepts like 'zebra' but, more specifically, concepts like 'sometime, someplace zebra'—all the zebras that ever have or ever will exist anywhere will have to be finite in number. No doubt they may be. But could anything legitimately regarded *merely* as a mode of concept-formation—a fixing of the truth-conditions of identity statements concerning *abstracta* of a novel kind—contain the resources to settle such a matter a priori?

The situation is not at all to be compared to the fashion in which Hume's Principle settles the infinity of the universe. Hume's Principle does not require there to be infinitely many objects save by requiring that there be infinitely many objects of the very kind—numbers—with which it is specifically concerned. And to stress: it is, in a way, misleading to think of it as requiring even that. If an abstraction is really to be merely a principle whereby a certain concept is determined, then—as I keep repeating—all it can legitimately do is to lay down *truth-conditions* for statements of the kind which it serves to explain; it has to be an independent matter whether those truth-conditions are satisfied. In the case of Hume's Principle, that 'independent matter' is, in the first instance (the case of zero), a second-order logical matter; and

one should not confuse its being a logical truth that the ↳ truth-condition assigned by Hume's Principle to

$$Nx\!:\!x \neq x = Nx;\ x \neq x$$

is met with the idea that Hume's Principle all of itself, as it were, settles the existence of Nx: x ≠ x. A legitimate abstraction, in short, ought to do no more than introduce a concept by fixing truth-conditions for statements concerning instances of that concept. In a limiting case, it may associate such statements with truth-conditions of such a kind that they are necessarily fulfilled—that is just what Hume's Principle sometimes does. But that should not blow away the distinction between concept-formation on the one hand and mere axiomatic stipulation about existence on the other.

The Nuisance Principle, for its part—still granting that it is indeed inconsistent with Hume's Principle—is on the wrong side of that distinction. And the reason it is so is because it carries implications for the extensions of concepts, especially

Downloaded from https://academic.oup.com/book/26137/chapter/194208803 by University of St Andrews Library user on 02 February 2026

concrete sortal concepts, quite unconnected with the concept it aims to introduce. The Nuisance Principle cannot be viewed merely as introducing that concept, by fixing truth-conditions for statements concerning instances of it, and then leaving it to the world to settle which if any of those truth-conditions are satisfied; if that were all it served to do, it could not possibly carry any import for the cardinality of the extensions of concepts which are quite unconnected to the concept of Nuisance—whose extensions, that is to say, do not even partially overlap with that of Nuisance—and whose explanation proceeds quite independently. How many sometime, someplace zebras there are is a matter between that concept and the world. No principle which *merely* assigns truth-conditions to statements concerning objects of a quite unrelated, abstract kind—and no legitimate second-order abstraction can do any more than that—can possibly have any bearing on the matter. *Tant pis*, then, if the Nuisance Principle does have a bearing on the matter: so much the worse for its claim to be, as the Fregean proposes to view Hume's Principle, merely a legitimate principle of concept-formation, a legitimate second-order abstraction.

What is at stake in this disanalogy is, in effect, *conservativeness* in (something close to) the sense of that notion deployed in Hartry Field's exposition of his nominalism.[48] For Field, pure mathematical theories are typically literally false, since there are no abstract mathematical objects. But he holds that a false mathematical theory, M, can nevertheless be *good* provided it is conservative over any nominalistically acceptable theory, N, in the sense of generating, when added to N, no claims about the objects comprising the ontology of N which are not already consequences of N alone. Field's interest in this

notion, of course, has nothing to do with the status of putative definitions, or other forms of conceptual explanation. ↳ But anything which *is* merely a definition, or other type of concept-formation, must be required to be conservative, in just that intuitive sense, with respect to any theory to which it is added. And of course, since it might be added to theories which are antecedently neutral on the matter of the finitude of certain of the concepts they concern, the Nuisance Principle fails the test. It therefore has no claim to whatever form of analyticity, or other epistemological advantages, are properly ascribed to statements whose role is merely to determine concepts.[49]

Hume's Principle, by contrast, is conservative if, as Field believes, all classical mathematics is; indeed, it is conservative if classical second-order arithmetic is. That issue can hardly be any simpler, epistemologically, than the corresponding question of consistency. But I do not think that many would hold the matter to be in serious doubt.

If this contrast is soundly drawn, then its underlying principle may be used to default all of the troublesome examples which we have so far reviewed, and hence obviates any strict need for the responses I canvassed earlier. Any principle which entails that the universe is less than countably infinite will be non-conservative and so open to the objection. It follows that all second-order abstractions which stay within the modest brief which the Fregean view assumes Hume's Principle discharges—the brief of merely fixing truth-conditions for a certain kind of statement, leaving it to independent considerations to determine whether those truth-conditions are met—all such abstractions will be consistent, at least as far as cardinality is concerned, with Hume's Principle. And in that case, we have the needed a priori guarantee that there can be no cogent objection of the kind which Boolos tried to develop.[50]

X

I contend, then, that—at least as far as number theory is concerned, and at least as far as the 'Bad Company' objection is concerned—the prospect is still open of successfully completing a programme of essentially the philosophical significance which Frege hoped for.[51] Might such a prospect be extended further? Is there any hope of encompassing Analysis, for instance, within the scope of a neo-logicist project? This and the final section will offer a rapid overview of some of the relevant points and problems.

The natural first move in any attempt to go further is to call for some reconstructed notion of *set*—specifically, a notion of set which is (i) yielded by abstraction, (ii) consistent, and (iii) reasonably strong—i.e. the comprehension principle implicit in the associated abstraction must be reasonably powerful. Frege's notion of extension, or more generally course-of-values, met the first and third of these conditions but—alas—crucially omitted to satisfy the second. By contrast, the modern iterative conception of set may reasonably safely be taken to meet the second condition, and certainly meets the third. The trouble is it does not meet the first: its sets are not yielded by an abstraction as the natural numbers are yielded by Hume's Principle, but are merely *postulated*. So no account is offered of why such things should be supposed to exist. Is there any hope for a conception of set which meets *all three* constraints?

There were three elements in Frege's thinking which went to compose the recipe for the disaster which overtook *Grundgesetze*. The first is one which Dummett has emphasized in his criticisms of Frege: the conception of a single, unrestricted domain of objects which constitutes the range of the first-order quantifiers of his theory and within which have to be found all entities yielded by logical abstraction. The second is the belief that abstraction is permissible on an arbitrary (higher-order) equivalence relation—that any such equivalence relation can play the role of an abstractive relation. This generality is in effect encapsulated in Basic Law V.[52] For the reasons we have been reviewing, a descendant of Frege's programme cannot, it would seem, allow itself this luxury, even in consistent cases. The third component is the absence altogether of any kind of predicativity restrictions on higher-order quantification. Frege's ontology comprises a single type of objects and successive levels of Concepts. If we were to superimpose on to it the distinctions *at* a given level effected by, for example, the ramified theory of types, the usual derivation of Russell's paradox would be blocked at the critical step where the Russellian predicate:

$$(\forall F)(y = \{x: Fx\} \rightarrow \neg Fy),$$

has to be taken to lie within the range of its own quantifier.

One possible direction, then, for a Fregean to take is to investigate the status and strength of systems in which Basic Law V, overtly quite unmodified, is harnessed to a fully predicative second-order logic. I have not myself worked on this kind of proposal, and do not know in detail how things might develop.[53] What I want to close by considering is a quite distinct direction, first noted, as it happens, by Boolos himself.[54] It is in the tradition of 'limitation of size', and thus might well have a philosophical motivation in the considerations to do with indefinite extensibility and how best to conceive of the objectual universe briefly canvassed above in connection with Nuisances.

Now 'limitation of size' could, in the context of second-order abstraction, suggest a number of things. The simplest would be a restriction of the scope of admissible abstraction to Concepts which are 'not too big'—Concepts under which fall only limitedly many things—the exact form of limitation remaining to be specified. Naturally, a restriction of this kind will be ineffectual as a response to the paradoxes unless we can take it that the key Concepts involved in their generation—for example the Russellian Concept cited above—will turn out to lie on the wrong side of the restriction. But there is cause to be sanguine about that. Intuitively, the Russellian Concept ought to turn out 'too big' since all unit sets, of which there ought to be no fewer than there are objects, ought to fall under it.

Independently of that constraint, however, it's initially hard to see how any such proposal could fly. For it would seem that, in order to work with the proposed restriction, whatever its detail—in order to know which if any Concepts are not too big—we will need to know *first* what objects are contained within the universe. Any abstraction which has to work within the restriction is therefore liable to be altogether emasculated as a comprehension principle. If Basic Law V, for instance, were modified so as to assume this shape:

$$V^*: (\forall F)(\forall G)[F, G \text{ are not too big} \\ \rightarrow \{x:Fx\} = \{x:Gx\} \leftrightarrow (\forall x)(Fx \leftrightarrow Gx)],$$

then a derivation of the existence of $\{x: x \neq x\}$, for example, would require that the Concept, not self-identical, not be too big. But it *will*, presumably, be too big, however exactly that notion is defined, if the universe is *empty*. So we shall need to know that is not so *before* we can put V^* to any effective use. ↳ Hence V^* could not function as a pure comprehension principle—the burden of proofs of existence would have to be carried elsewhere.

Boolos's suggestion neatly side-steps this problem. Provided the property, 'is not too big', is expressible in second-order logic, it will follow, in accordance with a fact noted earlier,[55] that the relation

$$(F \text{ is too big} \& G \text{ is too big}) \vee (\forall x)(Fx \leftrightarrow Gx)$$

will be an equivalence relation. Boolos's idea is to consider an abstraction on *this* relation, with 'F is too big' defined as meaning that F is equinumerous to the Concept, self-identical:

$$(\forall F)(\forall G)[\{x: Fx\} = \{x: Gx\}$$
$$\leftrightarrow ((F \text{ is too big} \& G \text{ is too big}) \vee (\forall x)(Fx \leftrightarrow Gx))].$$

Now there is no problem about comprehension. Any Concept satisfies the condition of self-coextensiveness, and so verifies the right-hand disjunct of the abstractive relation. So for every F, the new principle allows us to derive the existence of an object

$$\{x: Fx\}.$$

But the reasoning which leads to Russell's paradox merely channels itself harmlessly into a proof that the Russellian Concept,

$$(\forall F)(y = \{x: Fx\} \rightarrow \neg Fy),$$

is too big. In general, for Concepts which are not too big, the new principle supplies objects which behave like naïve sets; for all Concepts as big as the universe, however, it supplies a single, non-extensional object.

The question immediately arises whether in the light of earlier considerations this principle—henceforward to be referred to as VE[56]—can possibly rank as an acceptable abstraction. The concern is its analogy of structure with some of the unwelcome cases which we reviewed. Specifically, VE is an instance of (D):

$$(\forall F)(\forall G)(\Sigma(F) = \Sigma(G) \leftrightarrow (\varphi F \& \varphi G) \vee (\forall x)(Fx \leftrightarrow Gx)).$$

Was it not the upshot of our earlier discussion to cast doubt on the credentials of any abstraction of this form?

Not quite. Section V argued that the possibility of constructing instances of (D) which yield unmotivated resolutions of unconnected (contingent) matters by exploitation of the, so to say, paradoxical disjunct, is no more reason for rejecting abstraction *in general* as a legitimate form of concept-formation than is the 'heterological' paradox, or variants of it, a reason for rejecting predicate definition by the stipulation of satisfaction-conditions. That leaves open the question whether *some* abstractions of this shape may nevertheless be acceptable, even if not all. But which ones?

p. 301 Well, we know that it will be possible to exploit the inconsistency of Basic Law V to obtain from any (D)-type abstraction a proof that some Concept meets the condition on F and G detailed by the left-hand disjunct of its right-hand side. So we will be able to exploit the inconsistency of Basic Law V to derive from VE a proof that some Concept is indeed too big. What is salient, however, is that, in the present case, that is something that we can prove *independently*, as a theorem of second-order logic. For it follows from the definition of 'too big' that self-identity is itself too big.

That suggests that a particular D-type abstraction, or other paradox-embedding instance, may be acceptable provided it meets another conservativeness constraint: roughly, that any consequences which may be elicited by exploiting its paradoxical component should be, a priori, in independent good standing. Theorems of logic are so *par excellence*. But 'independent good standing' might also reasonably be taken to cover the case where a consequence elicited from such an abstraction by 'fishy'—paradox-exploitative—means can also be obtained not from logic alone but, as it were, innocently from additional resources provided by that very abstraction.

To make this a little clearer, reflect on the Field-type conservativeness constraint used to spike the Nuisances, and the series of D-type abstractions obtained by taking φ as respectively 'is at least finite', 'is at least countably infinite', 'is at least uncountably infinite', 'has at least as many instances as there are subsets of the real numbers', and so on. Each of these is—presumably—consistent and each entails stronger results than all its predecessors about the cardinality of the universe. Moreover none is in contravention of the Field-type conservativeness constraint since, for all they say to the contrary, the lower bounds they respectively ascribe to the size of the universe may (and in the case of the first, must) be attributable to the very population of objects they serve to introduce. It remains that they ought, even so, to impress as unsatisfactory if

these cardinality consequences can be elicited *only* by exploitation of their embedded paradoxes. For if these are reckoned as decent abstractions, then proofs for cheap are thereby foisted on us for any 'at least' cardinality claim, however inflated. Notice however that, in the case of the first and second examples[57]—

$$(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow (\text{at least finite } (F)$$
$$\&\text{at least finite } (G)) \vee (\forall x)(Fx \leftrightarrow Gx)),$$

and

$$(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow (\text{at least countably infinite } (F)$$
$$\&\text{at least countably infinite } (G)) \vee (\forall x)(Fx \leftrightarrow Gx))$$

the relevant consequences (respectively, that some F is at least finite and that some F is at least countably infinite) may be exhibited just by drawing in an intuitively quite innocent way on the resources they supply.[58] Unless, therefore, some other form of paradox-exploitative reasoning serves to elicit further consequences from these examples not in that benign case, it does not seem that their exemplification of the D-pattern provides any reason to rule them inadmissible.

A second conservativeness proposal, then, might be developed along the following lines. Consider any consistent abstraction A which, in the intuitive fashion which concerns us, embeds a paradox. That involves that there will be a formal expression of A other instances of which can be shown to be non-conservative in the Field-type sense. Let A^* be such an instance, and C^* any consequence whose being entailed by A^* renders the latter non-conservative in that sense. Then the additional conservativeness constraint which A has to meet is that, for each such C^*, the corresponding C, i.e. the reinterpretation of C^* in just the respects in which A is a reinterpretation of A^*, must be exhibitable as a consequence of A in a non-paradox exploitative fashion: that is, there must be a demonstration of C from A of a form of which no other instance is a demonstration of the non-conservativeness of an abstraction.[59]

I do not know whether VE meets that condition in full; but at least it jumps the first hurdle.

XI

Say that an object, {x: Fx}, yielded by VE is a *Set* just in case F is not too big. As Boolos shows, an interesting amount of work can be done with VE. The axioms of General Set Theory—Extensionality, Adjunction, and *Aussonderung*—are all derivable for *Sets* within the system—*SOLVE*—consisting of second-order logic plus VE, and number theory may be

developed therein.[60] ↳ Boolos himself has seemed inclined to play down the philosophical significance of the point, partly no doubt because of his low opinion of 'contextual definitions' in general, but more specifically because

> like Frege's revision of Basic Law V, and unlike Hume's Principle and the original Basic Law V, its drawback is that there is no notion, and certainly no unquestionably logical notion, that the patched axiom can be thought to be analytic *of*.[61]

But that is not true. There *is* a notion that VE is analytic of: it is analytic of the notion which stands to it as *cardinal number* stands to Hume's Principle and *direction* stands to the abstraction for directions. And this is a logical notion in precisely the sense that, like Hume's Principle, its abstractive relation is definable in second-order logic. What is true is that there is no *prior*, no intuitively entrenched notion, no notion given independently, which VE is analytic of. But maybe it gets as close as any consistent notion can to the naïve, non-iterative conception of set which Frege thought was a notion of logic; and maybe a limitation-of-size component is something which *that* notion should have incorporated in any case, paradoxes apart. In any case, to suppose an abstraction needs the sanction of faithful codification of such an antecedent, entrenched notion is to miss what I have been urging all through: that success for an abstraction is success as an *explanation*, not as a *description*.

Any idea, however, that VE might contain the key to a full reinstatement of Frege's project is dashed by the realization that classical Analysis cannot be forthcoming on this basis. The block is that, as Boolos observes,[62] nothing equivalent to the

standard set-theoretic axiom of Infinity is derivable in SOLVE. This may seem bewildering—I just remarked that arithmetic is constructible in SOLVE, so what about the set of the natural numbers thereby constructed? The explanation, of course, is that while the domain of any model for SOLVE will indeed be infinite, the existence of an infinite *Set* among the objects within that domain—that is, a 'well-behaved' object, obeying Extensionality in particular—is quite another matter. We are assured that the Concept, Natural Number, will be infinitely instantiated in any such model; but there is no assurance that the object, {x: *Nat*(x)}, which VE thereby gives us, is a *Set*. On the contrary, SOLVE has models in which all the *Sets* are finite.[63] ↳ Since the usual kind of construction of the real numbers depends essentially on the formation of well-behaved infinite sets—initially, sets of pairs of natural numbers—it is certain that construction cannot go through in SOLVE.

Hard on the heels of that disappointment comes the corollary that the Cantorian theory of the transfinite must likewise be unobtainable in SOLVE. Actually, as Boolos notes,[64] neither the Power Set axiom—in the form that to every *Set* corresponds a *Set* of all its subsets—nor, consequently, Cantor's Theorem can be proved in SOLVE. But even if they could, the unobtainability of infinite *Sets* would ensure that they could not be exploited to force open the door to Cantor's Paradise.

To one interested primarily in the mathematical capabilities of SOLVE, these limitations will simply invite the addition of supplementary axioms. We could add, for instance, a pair of axioms stipulating that *Natural Number* is not too big, and that provided F is not too big, neither is 'is a subset of {x: Fx}'. But for the neo-Fregean, such a move is tantamount to folding one's hand. The project was to explain how a recognition of the existence of the domains of objects called for by classical mathematical theories might be accomplished. The mooted additional axioms would contribute nothing to that project: they merely amount to stipulations about the size of the intended domains—the fundamental question, how the existence of such domains might be recognized in the first place, would be left wholly unaddressed.

Might these difficulties be overcome by some quite different 'New V', perhaps unconnected to the principle of limitation of size? There are two main strategies to consider. Naïve sets, {x: Fx} and {x: Gx}, are the same just when their defining characteristics, F and G, are coextensive: that is the principle fatally incorporated into Basic Law V. One strategy, illustrated by Frege's own proposal in the face of Russell's paradox, is to weaken the requirement of coextensiveness to that of *all-but-coextensiveness*—coextensiveness with respect to all objects save possibly certain designated exceptions (in Frege's case, {x: Fx} and {x: Gx} themselves).[65] However the technical shortcomings of Frege's own proposal are well known. And it seems likely that similar difficulties will afflict any principle of this kind.[66]

The alternative is to set the criterion of identity for the abstracts not as some form of all-but-coextensiveness but as coextensiveness proper for some Concepts and as some other relation for others. Essentially, that brings us ↳ back to something of the shape of VE. Now such an abstraction might work with all kinds of conditions besides too-bigness, of course. But on reflection, it would seem to be this general shape, rather than the detail of any particular proposed condition, which is responsible for at least one of the problems we have noted. Any abstraction of the form:

$$(\forall F)(\forall G)(\{x\colon Fx\} = \{x\colon Gx\} \leftrightarrow (\varphi F \& \varphi G) \vee (\forall x)(Fx \leftrightarrow Gx)),$$

must invite the question: how many Concepts in general over a given domain of objects can be expected to meet the condition φ? Let k be the cardinality of the domain. Then if more than k Concepts fail to meet φ, the abstraction will be inflationary—will call for more objects than the domain contains. So all but at most k Concepts must be φ. But in that case at most k objects will be yielded which behave like *Sets*. So it will not always be true that a *Set* has a power set which is also a *Set*.

These remarks are inevitably sketchy. But what they suggest is a moral close to one which Boolos also draws:[67] that while many of the operations—Unit Set, Pairing, Separation, Union, Choice—characteristic of classical set theory may well be salvageable on the basis of a pure abstractive notion of set, free of the troubles which beset the naïve notion, the *distinctive ontology* of that theory is not: the justification of the belief in uncountable domains is not to be sought on the basis of an abstraction on anything like the model of VE.

Still, there are other models. In particular, there is the shining example of Hume's Principle itself. Hume's Principle is a case of an abstraction which is inflationary with respect to any finite domain but stable thereafter, which is conservative

(in the first of the senses we distinguished—the second, of course, does not apply), and which is independent of the notion of set. Whence the idea that the ontology for Analysis has to be secured by sufficiently powerful principles concerning sets? Might there not be some abstraction to take one *directly* to the reals as Hume's Principle—restricted to finite Concepts—takes one to the naturals, without any play with the notion of set at all? Such a principle would have to be inflationary with respect to any finite or countably infinite domain but stable thereafter. To be interesting, it would have to meet at least the type of conservativeness constraint with which we have been concerned. The hard question is what the abstractive relation, and its terms, might be. Effectively, whether or within what framework such an abstraction might be possible is a matter so far almost entirely unresearched.

There is another possibility. The sought-after abstraction does not have to lead *directly* to the reals. *Any* uncountable population will provide a backcloth against which Natural Number will be not too big in the sense of VE. So SOLVE would, in the context of such a background ontology, provide the ↳ resources for a standard set-theoretic construction of Analysis. In other words, augment SOLVE by any otherwise acceptable abstraction—it need not even be a logical abstraction[68]—which demands an uncountable population of objects and you will have the resources to carry through a set-theoretic construction of Analysis along tried and tested lines[69] without making any direct use of the new objects at all. In such a scenario, the reals could still be viewed as logical objects—since they would be identified with certain of the abstracts introduced by VE—but the recognition of their existence, in contrast with that of the naturals, would not be grounded in a purely logical abstraction.[70]

Enough speculation. Almost all the best modern commentary, even Michael Dummett's, has effectively written Frege's philosophy of mathematics off.[71] Let me seize this opportunity to leave the reader with a sense, to the contrary, of how open the problems and possibilities for Fregean, or quasi-Fregean foundations for classical mathematics remain.

Notes

1   Wright (1983).

2   It is just possible, therefore, though of course highly unlikely, that he never realized the fact. For a detailed demonstration of how the proofs of the fundamental laws of arithmetic given in *Grundgesetze* can be reconstructed without recourse to abstracts for extensions and courses-of-values, see Heck (1993).

3   The qualification is unnecessary for anyone content with the usual second-order definition of identity: $x = y \leftrightarrow (\forall F)(Fx \leftrightarrow Fy)$. I shall henceforward omit it.

4   A detailed sketch of the derivation of the Peano axioms from this principle, there referred to as $N^=$, is given in Wright (1983: 158–69). An outline is given in the Appendix to Boolos (1990*a*). The derivability of Frege's Theorem is first explicitly noted in Charles Parsons (1964: see remark 194). My own 'rediscovery' of the theorem was independent.

5   Formally:

$$(\forall F)(\forall G)\{Nx{:}Fx = Nx{:}Gx \leftrightarrow (\exists R)[(\forall x)[Fx \to (\exists y)(Gy\&Rxy\&(\forall z)(Gz\&Rxz \to z = y))]\&(\forall y)[Gy \to (\exists x)(Fx\&Rxy\&(\forall z)(Fz\&Rzy \to z = x))]]\}.$$

This principle is attributed to Hume at *Grundlagen* §63. Readers of Book 1, part III, sect. 1 of the *Treatise* may feel Hume is flattered by the attribution. But Hume's Principle has achieved some currency among the leading writers on these matters. I will stick to it here despite Michael Dummett's excoriation of the term in the Appendix to Dummett (1998). The truth is that its use should no more encourage misapprehension than does the customary use of 'sunrise' (an example of van Fraassen's).

6   This has been noted independently by George Boolos (see the discussion in Boolos 1987*b*: 6–10), and John Burgess's review of *Frege's Conception* (Burgess 1984); also Hodes (1984: see remarks at 138 concerning an equivalent of $N^=$ involving branching quantifiers). For a detailed proof, see the first appendix to Boolos and Heck (1998).

7   See Boolos (1990*a*: 268).

8   §§62–3.

9   Formally:

$$(\exists R)[(\forall x)[x \neq x \to (\exists y)(y \neq y\&Rxy\&(\forall z)(z \neq z\&Rxz \to z = y))]$$
$$\&(\forall y)[y \neq y \to (\exists x)(x \neq x\&Rxy\&(\forall z)(z \neq z\&Rzy \to z = x))]],$$

a sentence whose validity is overdetermined by the considerations that x ≠ x necessarily has no instances and that identity is available as an interpretation for R.

10   One who presses this objection is unlikely to regard it as any mitigation of the point that the non-emptiness of the universe is a theorem of both first- and second-order logic as standardly formulated. That merely reflects the fact that our *interest* in both cases is in sentences which are valid in any non-empty domain. However, in one good sense of 'truth of logic', the truths of logic have to be appreciable *a priori. And while (a), viewed as genuinely defined by Hume's Principle, can be known a priori, that simple consideration may seem evidently insufficient to determine that (b) can.*

Alternatively, as Richard Heck has observed (private communication), the objector may make the intended point by working not with (a) but with:

$$(\exists y)(Nx:x \neq x \neq Nx:x = y). \quad (c)$$

Hume's Principle will allow us to derive (c) from:

$$(\exists y)\neg(\exists R)[(\forall x)[x \neq x \rightarrow (\exists z)(z = y \& Rxz \& \quad (\forall w)(w = y \& Rxw \rightarrow z = w))]$$
$$\& (\forall u)[u = y \rightarrow (\exists x)(x \neq x \& Rxu \& (\forall z)(z \neq z \& Rzu \rightarrow z = x))]],$$

which will hold in any non-empty domain. However (c), standardly interpreted, implies that there are two distinct objects, and so does not.

11   So we depart from the rather unhappy terminology of 'contextual definition' of e.g. Heck and Boolos.

12   *Grundlagen* §§63–4.

13   I am here, of course, using 'concept' in the usual informal philosophical way, and will continue to do so. But I shall henceforward use 'Concept' as a term of art to denote the unsaturated, extensional entities which Frege viewed as the *Bedeutungen* of predicates and relational expressions, and also sometimes when speaking of the items that comprise the range of the predicate- and relational variables in second-order logic, which I will assume at various points in later arguments to be extensionally individuated.

14   Notably Hartry Field. See e.g. Field (1989: ch. 5, 'Platonism for Cheap').

15   For further dissolution of the parallel, see my 'Field and Fregean Platonism', this volume, Essay 6, Sect. 3.

16   The classic statement of these problems, which has enormously influenced so much recent work, is of course Benacerraf (1973).

17   See Boolos (1997).

18   The understanding evinced, for instance, by Boolos's remark, 'the view Wright describes [is] . . . logicism only if it is claimed that Hume's principle is a truth of logic' (Boolos 1990*a*: 274).

19   So why did not Frege himself fall back on this proposal in response to Russell's Paradox? For one interesting line of thought on the matter, see Heck (1997*a*).

20   Dummett (1991*a*).

21   Necessary, at least, if it is to be possible to iterate in usual ways—for instance, to number the numbers meeting a certain condition, as Frege's strategy for proof of the infinity of the finite cardinals requires, or to form sets of sets.

22   For perceptive discussion of each of the four lines of objection see Bob Hale's 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', this volume, Essay 8. The second objection was anticipated and treated in sect. 4 of my 'Field and Fregean Platonism', this volume, Essay 6. The fourth objection is discussed in detail in my 'On the Harmless Impredicativity of $N^=$ (Hume's Principle)', this volume, Essay 10.

23   Boolos (1990*a*); Heck (1992).

24   Another inconsistent example, noted by Hodes (1984: 138), is provided by the abstraction for relation-numbers which affirms that a pair of relations have the same relation-number just in case they are isomorphic.

25   Thus Dummett writes:

> 'If the context principle, as expounded by Wright, is enough to validate the 'contextual' method of introducing the cardinality operator, it must be enough to validate a similar means of introducing the [class] abstraction operator.
>
> This is why the mere fact that, on his view, it is unnecessary to define the cardinality operator in terms of classes or of value-ranges does not entitle Wright to ignore the problem of the abstraction operator. For Frege's method of introducing [the latter] was, notoriously, *not* in order. It rendered his system inconsistent; and that inconsistency forced him eventually to acknowledge that his entire enterprise had failed. If the context principle, as stated by Wright, was sound, there could have been no inconsistency.' (Dummett 1991*a*: 188)

A somewhat less emphatic expression of the same worry occurs in Burgess (1984).

26   As in effect Dummett does take it—his real objection is the fourth listed above, concerning impredicativity.

27   Thus I reject the suggestion that a problem must lurk for the Fregean in the fact that, as observed by Heck (1992), the consistency of second-order abstractions will not in general be effectively decidable.

28   First noted, I think, in Heck (1992).

29   Here are the moves:

1	(1) $(\forall F) (\forall G)(\Sigma F = \Sigma G \leftrightarrow P \lor R(F, G))$	Assumption
2	(2) $\neg P$	Assumption
1	(3) $\Sigma F = \Sigma G \leftrightarrow P \lor R(F, G)$	(1), UE
4	(4) $\Sigma F = \Sigma G$	Assumption
1,4	(5) $P \lor R(F, G)$	(3), (4), prop. logic
1,2,4	(6) $R(F, G)$	(2), (5), prop. logic
1,2	(7) $\Sigma F = \Sigma G \rightarrow R(F, G)$	(4), (6), prop. logic
8	(8) $R(F, G)$	Assumption
1,8	(9) $\Sigma F = \Sigma G$	(3), (8), prop. logic
1	(10) $R(F, G) \rightarrow \Sigma F = \Sigma G$	(8), (9), prop. logic
1,2	(11) $\Sigma F = \Sigma G \leftrightarrow R(F, G)$	(7), (10), prop. logic
1,2	(12) $(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow R(F, G))$	(11), UI twice (1 and 2 are 'F' and 'G' free)
1	(13) $\neg P \rightarrow (\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow R(F, G))$	(2), (12), prop. logic

30   As Heck observes.

31   That is, for which, for any F and G, ϕ F and $(\forall x)(Fx \leftrightarrow Gx)$ jointly entail ϕ G. Of course, if concepts are Concepts—cf. fn. 13 above—the qualification is unnecessary. And I shall drop it henceforward. The omission need not, however, at this point be taken as a commitment to the extensionality of the items in the range of the second-order variables. It is justified by the fact that at least as far as any second-order context, ϕ . . . , is concerned, it will be a theorem of second-order logic that $(\phi$ F $\& (\forall x) (Fx \leftrightarrow Gx)) \rightarrow \phi$ G. In other words, second-order logic is blind to any distinctions which non-extensionally individuated but coextensive concepts might exhibit.

32   Proof:

> Symmetry: trivial.
>
> Transitivity: assume $(\phi$ F $\& \phi$ G$) \lor (\forall x) (Fx \leftrightarrow Gx)$ and $(\phi$ G $\& \phi$ H$) \lor (\forall x) (Gx \leftrightarrow Hx)$. To see that $(\phi$ F $\& \phi$ H$) \lor (\forall x) (Fx \rightarrow Hx)$ is entailed, reflect that if both the lhs disjuncts in the assumptions hold, then ϕ F $\& \phi$ H follows; that if both rhs disjuncts hold, then $(\forall x)$ (Fx $\leftrightarrow$ Hx) follows; and that if, third, ϕ F $\& \phi$ G and $(\forall x)$ (Gx $\leftrightarrow$ Hx) hold, then ϕ H, and hence ϕ F $\& \phi$ H, follows by the point noted in fn. 31; *mutatis mutandis* for the fourth case.

33   Proof:

1	(1) $(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow (\phi$ F $\& \phi$ G$) \lor (\forall x)(Fx \leftrightarrow Gx))$	Assumption
2	(2) $(\forall F)(\forall G) \neg(\phi$ F $\& \phi$ G$)$	Assumption
1	(3) $\Sigma F = \Sigma G \leftrightarrow (\phi$ F $\& \Phi$ G$) \lor (\forall x)(Fx \leftrightarrow Gx)$	(1), UE twice
2	(4) $\neg(\phi$ F $\& \phi$ G$)$	(2), UE
5	(5) $\Sigma F = \Sigma G$	Assumption
1,2,5	(6) $(\forall x)(Fx \leftrightarrow Gx)$	(3), (4), (5), prop. logic
1,2	(7) $\Sigma F = \Sigma G \rightarrow \forall x(Fx \leftrightarrow Gx)$	(5), (6), prop. logic
8	(8) $(\forall x)(Fx \leftrightarrow Gx)$	Assumption
1,8	(9) $\Sigma F = \Sigma G$	(3), (8), prop. logic

1	(10) $(\forall x)(Fx \leftrightarrow Gx) \rightarrow \Sigma F = \Sigma G$	(8), (9), prop. logic
1,2	(11) $\Sigma F = \Sigma G \leftrightarrow (\forall x)(Fx \leftrightarrow Gx)$	(7), (10), prop. logic
1,2	(12) $(\forall F)(\forall G)(\Sigma F = \Sigma G \leftrightarrow (\forall x)(Fx \leftrightarrow Gx))$	(11), UI twice (1 and 2 are 'F' and 'G' free)
1,2	(13) Contradiction	(12) is Basic Law V
1	(14) $\neg(\forall F)(\forall G) \neg(\phi F \& \phi G)$	(2), (13), prop. logic
1	(15) $(\exists F)(\exists G)(\phi F \& \phi G)$	(14), higher-order logic
1	(16) $(\exists F) \phi F$	(15), higher-order logic

34   For which coextensiveness is therefore a congruence.

35   In case any reader is unfamiliar with the Curry paradox:

> Consider any significant declarative sentence, Q, and let P be a sentence with the truth-condition: if 'P' is true, then Q. Then

1	(1) P	Assumption
1	(2) 'P' is true	(1), DS
1	(3) If 'P' is true, then Q	(1), truth-condition of P
1	(4) Q	(2), (3) modus ponens
	(5) If P, then Q	(1), (4) conditional proof
	(6) If 'P' is true, then Q	(5), DS, sentential logic
	(7) P	(6), truth-condition of P
	(8) Q	(7), (5) modus ponens

36   This is essentially Boolos's diagnosis of the provenance of Russell's Paradox in Boolos (1993). But, while the substantive point is of great importance, I am not sure it is entirely happy *qua* diagnosis since unsatisfiability is not, in second-order logic, the same as inconsistency.

37   Proof:

> Symmetry: trivial.
>
> Transitivity: Suppose for *reductio* that the Concepts, (F-and-not-G)-or-(G-and-not-F) and (G-and-not-H)-or-(H-and-not-G), are both finite, but that (F-and-not-H)-or-(H-and-not-F) is infinite. Then each disjunct of the first and second must be finite, but at least one of the third's disjuncts must be infinite. Let F-and-not-H be such a disjunct. Only finitely many F-and-not-Hs can be not-G, or F-and-not-G would be infinite. So infinitely many of them must be G. But then there are infinitely many G-and-not-Hs, contrary to hypothesis. *Mutatis mutandis* if we assume the other disjunct, H-and-not-F, to be infinite.

38   Proof that Δ(F, G) is not Cambridge:

> Suppose F-and-not-G-or-G-and-not-F was finite but becomes infinite. Then one or both of F-and-not-G and G-and-not-F was finite but is now infinite. Let that be true of F-and-not-G. Suppose F has not changed its extension. Then infinitely many Fs which were G are now not-G. So G has changed its extension.
>
> Likewise, suppose F-and-not-G-or-G-and-not-F was infinite but becomes finite. Then one or both of F-and-not-G and G-and-not-F was infinite but is now finite. Let that be true of F-and-not-G. Suppose F has not changed its extension. Then infinitely many Fs which were not-G are so no longer, hence are now G. So, again, G has changed its extension.
>
> Thus any transformation from a situation where Δ(F, G) obtains to one where it does not, or conversely, requires either F, or G, to change in extension. Taking the extension of a Concept to be a real characteristic of it, it follows that Δ(F, G) is a non-

Cambridge relation on F and G.

39 Proof by induction on the size of the Concepts in question. Base: there will be k singletons. Step: suppose there are exactly k n-fold finite Concepts. Then each can be extended into an n + 1-fold Concept in exactly k ways. So there are at most k^2, = k, n + 1-fold Concepts. So there are at most k + k + k . . . , = ω.k, = k, more-than-n-fold finite Concepts if there are k n-fold finite Concepts. So there are exactly k finite Concepts altogether.

40 It suffices to establish that $(\forall x)[(Fx \,\&\, \neg\, Gx) \lor (Gx \,\&\, \neg\, Fx) \leftrightarrow (Fx \,\&\, \neg\, Hx) \lor (Hx \,\&\, \neg\, Fx)] \rightarrow (\forall x)\,(Gx \leftrightarrow Hx)$.

41 Boolos (1990*a*: 273) first called attention to such examples. There he offered the abstraction:

$$(\forall \mathrm{F})(\forall \mathrm{G})(\mathit{\Sigma}(\mathrm{F}) = \mathit{\Sigma}(\mathrm{G}) \ \text{ iff the number of things which are F or G but not both, is even}).$$

He calls the associated objects Parities. The above reasoning is essentially just a transposition of Boolos's reasoning (or what I take to be its nerve—his eight-line sketch is actually rather difficult to follow) to a somewhat simpler case. Boolos also offers another example (1990*a*: 227, n. 15) abstracting

on the equivalence relation,

(Equ) If at least two things are F, or at least two things are G, then F and G are coextensive; Prima facie, this case belongs with the less challenging group generated by schema (D) discussed earlier:

$$(\forall \mathrm{F})(\forall \mathrm{G})(\mathit{\Sigma}(\mathrm{F}) = \mathit{\Sigma}\mathrm{G} \leftrightarrow ((\varphi\mathrm{F}\&\varphi\mathrm{G}) \lor (\forall \mathrm{x})(\mathrm{Fx} \leftrightarrow \mathrm{Gx}))),$$

(take φ F as: less than two things are F). But that assimilation doesn't do it justice. For note that the reasoning schematized in fn. 33 above merely delivers the harmless:

Some F has fewer than two instances,

whereas Boolos shows that the abstraction on Equ is actually unsatisfiable in any domain containing three or more members.

42 Boolos (1993: 228). Boolos is actually referring at this point to the problem posed by instances of the schema (A).

43 Dummett first introduced this notion—which of course ultimately derives from Russell—in Dummett (1963). The notion looms large in the concluding chapter of Dummett (1991*a*), and in his 'What is Mathematics About?' in Dummett (1993*a*).
It is consistent with reservations about Dummett's willingness to regard the totalities of natural and real numbers in particular as indefinitely extensible—cases where no actual contradiction results from the attempt to suppose that there is a definite, comprehensive collection—and with misgivings about Dummett's attempt to deploy the notion in defence of a distinctively intuitionistic form of logicism, to see it is incorporating an important insight into how an unrestricted domain of objectual quantification must be conceived.

44 Note, by the way, that it would call for restrictions of the range of applicability of the cardinality operator: there could presumably be no Nx:x = x—no number of absolutely everything—for someone who pursues this line.

45 This apt term is Kit Fine's.

46 I know of no definite finding on this matter at present.

47 The interested reader should review the detail of the second Boolos example cited in n. 41.

48 Field (1980).

49 Drawing on Field's exposition for his own, quite different purposes (Field (1980: ch. 1, 13 ff.)), a precise conservativeness requirement which abstractions should meet could be formulated as follows. (But note that Field, who wished—for reasons I do not fully understand—to define a notion whereby a platonistic mathematical theory may be conservative with respect to a nominalistic theory which is actually inconsistent with it (because it e.g. affirms that there are no non-spatio-temporal objects) proceeds somewhat differently in detail.) Let

$$(\forall \alpha_i)(\forall \alpha_j)(\mathit{\Sigma}(\alpha_i) = \mathit{\Sigma}(\alpha_j) \leftrightarrow \alpha_i \approx \alpha_j),\ (\Sigma)$$

be any abstraction. Introduce a predicate, Sx, to be true of exactly the referents of the Σ-terms and no other objects. Define the Σ-*restriction* of a sentence, T, to be the result of restricting the range of each objectual quantifier in T to non-S items,—thus each subformula of T of the form, (∀ x) Ax, is replaced by one of the form, (∀ x) (¬ Sx → Ax), and each subformula of the form, (∃ x) Ax, is replaced by one of the form, (∃x)(¬Sx & Ax). Let θ be any theory with which Σ is consistent. Then Σ is conservative with respect to θ just in case, for any T expressible in the language of θ, {θ ∪ Σ} entails the Σ-restriction of T only if θ entails T. The requirement on acceptable abstractions is, then, that they be conservative with respect to any theory with which they are consistent.
This way of approaching the matter must rest, of course, on the good-standing of the predicate, ¬ Sx, and thus—as do so many points in the Fregean programme—demands a generalized solution to the Julius Caesar Problem.

50 It will be speedily apparent to the reader that the only abstractions which can comply with the mooted constraint will be ones whose implications, if any, for the cardinality of the range of their objectual variables are of the 'there are at least . . . ' rather than 'there are exactly . . . ' or 'there are at most . . . ' variety.

51   What—and how much—significance that might prove to be is the question which the collapse of Frege's own system has distracted us from.

52   Since, of course, any objects which are the referents of terms introduced by abstraction on an abstractive relation R may be identified with equivalence classes under R.

53   However, it strikes me as something of a dead hand on the prospects in this direction that the derivation of the normal recursive clauses for addition and multiplication in second-order logic with Hume's Principle makes seemingly unavoidable use of the impredicative ancestral. More specifically, any predicates—like 'z is the sum of x and y' or 'z is the product of x and y'—which will receive explicit, ancestral-style definitions in the logicist project are obviously going to present problems for a predicative version of it if we later want to take them as predicates for an induction—as we will in proving e.g. the functionality of addition—since there will be no reconciling the second-order quantifiers contained in their *definientia* with the quantifier in the statement of the induction principle.

54   See Boolos (1987*a*: 145–51) and, for a related proposal, Boolos (1993: 230–1).

55   See n. 31 above.

56   For 'V enlightened'. Boolos calls the principle New V.

57   But, so I believe, in none of the larger cases.

58   The point is trivial for the first example. That there can be no finite models for the second abstraction may be 'innocently' seen as follows. Obtain Σ x: x ≠ x. Show that x ≠ x and y = Σ x: x ≠ x are finite and hence, from the abstraction, that

$$\varSigma \mathrm{x}\colon \mathrm{x} \neq \mathrm{x} = \varSigma \mathrm{y}\colon(\mathrm{y} = \varSigma \mathrm{x}\colon \mathrm{x} \neq \mathrm{x}) \leftrightarrow (\forall \mathrm{y})(\mathrm{y} \neq \mathrm{y} \leftrightarrow \mathrm{y} = \varSigma \mathrm{x}\colon \mathrm{x} \neq \mathrm{x});$$

whence, since (∃ y) (y = Σ x: x ≠ x),

$$\varSigma \mathrm{x}\colon \mathrm{x} \neq \mathrm{x} \neq \varSigma \mathrm{y}\colon(\mathrm{y} = \varSigma \mathrm{x}\colon \mathrm{x} \neq \mathrm{x}).$$

Define unit set, {x}, as: Σ y: y = x. Show that, for any x, y = x is always finite and hence, by the abstraction, that {x} = {y} ↔ x = y. Likewise, show that Σ x: x ≠ x ≠ {y}, for any y. Then Σ x: x ≠ x, {Σ x: x ≠ x}, {{Σ x: x ≠ x}}, {{{Σ x: x ≠ x}}}, . . . is an ω-series.
This does not establish, of course, that (∃ F) (F is at least countably infinite) can be established *sans* paradox-exploitation *within* e.g. second-order logic plus the relevant abstraction. What it shows is rather that the paradox-exploitative proof is of a conclusion whose soundness, in the context of that abstraction, may be appreciated independently.

59   'Demonstration' must here be interpreted broadly so as to cover exhibitions of both formal and semantic consequence; cf. n. 58.

60   See Boolos (1987*a*) for details. A still fuller discussion of the set theory issuing from VE and a comparison with the Iterative conception is to be found in Boolos (1989). The appendix to the latter paper usefully summarizes the details.

61   Boolos (1993: 231).

62   Boolos (1989: 19).

63   One such construction due to Boolos (private communication) is as follows. Take some arbitrary non-set, α, and consider the series of sets formed as follows:

$$\mathrm{A}_0 \quad = \{\mathrm{a}\}$$
$$\mathrm{A}_{n+1} = \text{the set whose members are exactly a and the subsets of An (n, of course, finite).}$$

A model of VE may now take as its domain the set, D, consisting of the union of each such A_n. It will be apparent that, with the exception of the object α, the members of D are exactly its own finite subsets. Since D is countably infinite, a Concept, F, defined on D is not too big just in case finite. Accordingly, if F is finite, fix the reference of '{x: Fx}' to be the set of elements in D which fall under F; otherwise, let its reference be to the object α. Then VE is satisfied and every referent of a term, '{x: Fx}', whose embedded Concept—so to say—is not too big, is a finite set.

64   Boolos (1989: 19).

65   Actually, for any particular F and G, the exceptions *have* to include the objects, {x: Fx} and {x: Gx} if the principle is to block Russell's paradox—the verification of which claim I leave as a diversion for the reader. (An immediate consequential worry for this strategy would concern the apparently inevitable presence of the terms such an abstraction is supposed to introduce in the very statement of the abstractive relation on its right-hand side. But there are other problems—see n. 66.)

66   For relevant considerations here, see M. Resnik's helpful discussion of Frege's proposal in Resnik (1980: esp. 221).

67   Boolos (1993: concluding remarks).

68   i.e. one whose abstractive relation is definable in second-order logic.

69   Though perhaps not along lines Frege himself would have approved. For discussion of the likely course of Frege's preferred form of construction, see Simons (1987), and the ipsonymous ch. 22 in Dummett (1991*a*).

70   Indeed there is no cause to assume that only an *abstraction* can do the job here. Any well-grounded route into recognition of the uncountable would supply the necessary background. It would be interesting, for instance, to review Hartry Field's reasons for confidence in the uncountability of (what he conceives as concrete) space-time points. It seems to me that there is no compelling advance reason why a reputable epistemology for Analysis should not draw on such materials. To the contrary: it is entirely plausible that our fundamental beliefs about continuity should be grounded in thought specifically about space and time.

71   I am not forgetting that Dummett (1991*a*) compliments Frege as 'the greatest philosopher of mathematics yet to have written' and

credits him with many important insights and achievements; for instance,

(a) a successful polemic against formalism;

(b) a rightful emphasis on the applications of pure mathematics as something intrinsic to its content, and significant strides towards a satisfying account of the content of statements of number theory and analysis which meets this constraint;

(c) a crucial suggestion, via the idea that a single thought may be conceived under a variety of logical forms, about how analytic truths and deductive inference can come as news;

(d) the insight, contained in the Context Principle, that reference to abstract objects need not *per se* be problematical;

(e) the overarching logicist conception of the general character of pure mathematics (see esp. 308–9 of Dummett's text).

The fact remains that the single most important idea in Frege's philosophy of mathematics: the conception that the objects of the fundamental theories of classical mathematics—the real and natural numbers—may be given to us, and the fundamental laws concerning them justified, by reference to impredicative principles of abstraction, drawing only on logical conceptual resources, is dismissed by Dummett as a complete confusion, which he views as undermining Frege's entire approach and responsible for the collapse of his system.