Diálogos, 66 (1995) pp. 7–20.

AXIOM V AND HUME'S PRINCIPLE IN FREGE'S FOUNDATIONAL PROJECT

MATTHIAS SCHIRN*

1. Introductory remarks

In this paper, I want to discuss some points in Paul Benacerraf's article 'Frege: The Last Logicist' (1981) and Harold Hodes' article 'Logicism and the Ontological Commitments of Arithmetic' (1984). In particular, I want to examine critically what both authors say about Axiom V and courses-of-values in Frege's *Basic Laws of Arithmetic* on the one hand and his attempted contextual definition of the term 'the number which belongs to the concept F' (symbolically: '$N_xF(x)$') in the *Foundations of Arithmetic* on the other. Axiom V is the exact formal analogue of Frege's contextual definition of '$N_xF(x)$', and it is of considerable interest to analyze the relationship between the axiom and the attempted definition in the light of Frege's logicist programme. The definition was designed to introduce numerical terms while the task of the axiom was to introduce course-of-values terms. The definition is consistent while the axiom proved to be inconsistent. Nevertheless, both, the definition and the axiom, suffer from the same major defect: they fail to fix uniquely the references of abstract singular terms of a certain kind. Now, before I turn to Benacerraf and Hodes, it will be useful to sketch briefly Frege's second and third attempt to define number (*Anzahl*) in the *Foundations* as well as his introduction of courses-of-values in the *Basic Laws*.

* I am grateful to Roberto Torretti and Guillermo E. Rosado Haddock for much interesting discussion after each of the two talks entitled "Frege: Objetos Lógicos e Indeterminación de la Referencia (I + II)" which I gave at the University of Puerto Rico, Río Piedras on 12 and 14 April 1994 at the invitation of the Department and Seminar of Philosophy. Special thanks are also due to Álvaro López Fernández for organizing this rewarding meeting.

Frege's first attempt to define the individual cardinal numbers as objects in § 55 of the *Foundations* fails for reasons which I need not discuss here. The second likewise heuristic attempt to define number is intended to provide a criterion of identity for cardinal numbers ('E' is to abbreviate 'equinumerous'):

(I) $N_xF(x) = N_xG(x) :\equiv E_x(F(x),G(x))$

Frege discusses the difficulties involved in the transition from an equivalence relation to an identity by taking parallelism and identity of directions as his paradigm for illustration. In what follows, I shall transfer the main points of his discussion to the case of equinumerosity and numerical identity.

Frege points out that (I) raises three logical doubts. The first two can be met. The third doubt is this. The proposed criterion of identity embodied in (I) fails to cover all conceivable cases. It enables us to determine the truth-value of only those equations in which the expressions flanking the identity-sign are both of the form "$N_xG(x)$". Yet, the criterion is powerless to decide whether, say, Julius Caesar is the number of planets. We may call this "the Julius Caesar problem". If we already had a definition of the concept *n is a number* satisfying the requirement of sharp delimitation, we could stipulate: if q is not a number, then "$N_xF(x) = q$" is false; if q is a number and if it is given to us appropriately, then (I) will settle the truth-value of "$N_xF(x) = q$". At this stage of the inquiry, however, the indeterminacy arising from (I) cannot be removed by setting up the definition

(II) $N(n) := \exists\varphi(N_x\varphi(x) = n)$

since the numerical operator "$N_x\varphi(x)$", forming a part of the *definiens*, has not yet acquired a determinate meaning. In other words, in order to apply (I), we would first have to know in each case whether "$n = N_xF(x)$" is true or false.

At any rate, Frege rejects (I) only on the ground that his third objection has proved sound. Of course, to deprive (I) of its status as a definition of the term "$N_xF(x)$" does not amount to dispensing with formula (T) "$N_xF(x) = N_xG(x) \equiv E_x(F(x),G(x))$" altogether. (What I refer to as (T) is now often called "Hume's principle".) On the contrary, as a provable theorem, (T) is to play a crucial role in the intended logical construction of number theory (cf. FA, GLA, § 73). Furthermore, in § 104, Frege declares the determination of the sense of a recognition-judgment as a guiding principle for defining fractions, irrational numbers, and complex numbers in purely logical terms. A glance at the structure of Axiom V of the *Basic Laws* shows clearly that the transformation of an equivalence relation into an identity was and remained for Frege the means *par excellence* for introducing logical objects. Note that the equivalence relations embodied in Hume's principle and Axiom V are both of second level.

Facing the predicament resulting from his third objection to (I), Frege suggests a way out of it by eventually defining the number which belongs to the concept F as the extension of the (second-level) concept *equinumerous with the concept F,* that is, as an equivalence class of the relation of equinumerosity:

(III) $N_xF(x) := \lambda\varphi(E_x(\varphi(x),F(x)))$

This explicit definition is obviously designed to conform to Hume's principle. Yet, it rests on the *questionable* assumption that we intuitively know what the extension of a concept is. No doubt, at the time when Frege wrote the *Foundations,* he could not rely on a commonly accepted view of the nature of extensions of concepts, let alone of the nature of numbers. And to be sure, the rigorous standards that Frege applies to his own inquiry into the foundations of arithmetic require the introduction of new logical objects in a methodologically satisfactory manner. Thus, his assumption concerning extensions of concepts overshadows, at a crucial point, his reductionist enterprise as outlined in the *Foundations.*

In the *Foundations,* Frege intended to make the crucial step in his logicist programme by defining cardinal numbers as extensions of concepts. In the *Basic Laws,* where the logicist programme was to be carried out formally, Frege adheres mostly to this definition. He stresses, moreover, that all numbers are to be defined as extensions of concepts (BLA, 44; GGA I, 14). The domain of objects of objects of his formal system, consisting initially only of the two truth-values, has accordingly to be enlarged to include extensions of concepts. In § 3 of the *Basic Laws,* Frege introduces courses-of-values of functions which comprise extensions of concepts and of relations as special cases. Cardinal numbers are now defined as extensions of *first*-level concepts. Since, in contrast to the *Foundations,* it is the extension of a concept that is construed as the "bearer" of number, the numerical operator appears as a *first*-level function-name.

Frege introduces courses-of-values (respectively a name of the form "the course-of-values of the function $\Phi(\xi)$") in the context of a proposition, namely by stating a criterion of identity for them: "I use the words 'the function $\Phi(\xi)$ has the same course-of-values as the function $\Psi(\xi)$' generally to denote the same as the words 'the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument'". This informal stipulation corresponds to the ill-fated Axiom V of the *Basic Laws:*

$$(\dot{\varepsilon}f(\varepsilon) = \dot{\alpha}g(\alpha)) = (\neg\!\overset{\mathfrak{a}}{\smile}\!- f(\mathfrak{a}) = g(\mathfrak{a}))$$

When Frege comes to introduce courses-of-values in the manner just described, he encounters a variation of his old indeterminacy problem from the *Foundations*. At the outset of § 10 of the *Basic Laws,* he says of his informal stipulation that it "by no means determines completely the reference of a name like '$\dot{\varepsilon}\Phi(\varepsilon)$'". The criterion of identity for courses-of-values embodied in Basic Law V takes care of the truth-conditions of equations in which both related names are of the form "$\dot{\varepsilon}\Phi(\varepsilon)$". Yet, the criterion fails to determine the truth-value of "$\dot{\varepsilon}\Phi(\varepsilon) = q$" if "$q$" is not of the form "$\dot{\varepsilon}\Psi(\varepsilon)$". Frege's "permutation argument" in § 10 is intended to clarify this. I present it here in modern terms.

(P1) Suppose φ is the intended assignment of objects to course-of-values names satisfying Axiom V. Let h be a non-trivial permutation (of all objects), and consider the assignment φ' of objects to course-of-values names which is related to φ as follows: If Δ is assigned by φ to a given course-of-values name and $\Gamma = h(\Delta)$, then Γ is assigned by φ' to that course-of-values name. It follows that φ' is an assignment of objects to course-of-values names distinct from φ, but such that it satisfies Axiom V if φ does.

Frege proposes to resolve the referential indeterminacy of a course-of-values name "by determining for every function, as it is introduced, what values it takes on for courses-of-values as arguments, just as for all other arguments". As to the *other* arguments, he confines himself to the two truth-values. Up to § 10, Frege has introduced three primitive functions: the horizontal function $-\xi$, negation $\top\xi$, and the equality function $\xi = \zeta$. The horizontal function is a first-level concept under which falls the True alone; negation is a first-level concept under which falls every object with the sole exception of the True. The determination of the values of negation can be neglected. The horizontal function, however, turns out to be reducible to the identity relation: $\xi = (\xi = \xi)$ and $—\xi$ are obviously co-extensive concepts. Consequently, it remains merely to determine what value $\xi = \zeta$ takes on if we insert into one of the two argument-places of "$\xi = \zeta$" a course-of-values name and into the other a name of a truth-value which has not the form of a course-of-values name. Yet, Axiom V is powerless to decide whether or not either truth-value is a course-of-values.

2. Critical remarks on Benacerraf

In his stimulating article "Frege: The Last Logicist" (1981), Benacerraf emphasizes the kinship between Frege's attempted contextual definition of "$N_x F(x)$" and his introduction of courses-of-values by means of Axiom V. (In the latter case, Benacerraf speaks somewhat incorrectly of a contextual definition.) I cannot quite agree with what he says concerning the kinship that I just mentioned. It is inaccurate if not false to claim, as Benacerraf does, that not only the problem raised in § 10 of the *Basic Laws* but also its solution has exactly the same form as in the case of numbers and extensions of concepts in the *Foundations* (cf. Benacerraf 1981, 31). It is true that problem (*a*) "Axiom V does not determine completely the reference of a course-of-values term '$\dot{\varepsilon}\Phi(\varepsilon)$'" is basically the same as problem (*b*) "The proposed contextual definition (I) in the *Foundations* does not fix uniquely the reference of a numerical term '$N_x F(x)$'". We ought to bear in mind, however, that the exposition of (*a*) appears to be embedded in a different framework. Unlike that of (*b*), it relates expressly to an axiomatic system. In the *Basic Laws,* Frege makes his semantic stipulations within a fully worked out theory of reference, relying on a clear distinction between the object-language (i.e., his Begriffsschrift) and the metalanguage (i.e., German). It is in the latter that the syntactic and semantic construction of the formal system is carried out. Notice that the definitions, unlike the elucidations of the primitive function-names and the remaining non-definitional stipulations, are framed within the formal language. In the *Foundations,* we encounter nothing similar.

Now, despite the close affinity between (*a*) and (*b*), the solutions that Frege offers in the two cases are strikingly different. We have seen that Frege endeavours to solve problem (*b*) by setting up the explicit definition (III) of the term "$N_xF(x)$". Since in the system of second-order logic of the *Basic Laws* courses-of-values are indefinable —the course-of-values function $\grave{\varepsilon}\varphi(\varepsilon)$ is one of the eight primitive functions of that system— Frege must try another way than in the *Foundations* to overcome the referential indeterminacy of a name "$\grave{\varepsilon}\varphi(\varepsilon)$". He attempts to accomplish this by constructing a variant of his initial permutation argument and by making a stipulation on the basis of it. For brevity, I shall present his argument in the same fashion as before.

(P2) As in (P1), let φ be an assignment of objects to course-of-values names satisfying Axiom V. Let $f(\xi)$ and $g(\xi)$ be two particular, extensionally non-equivalent functions. Let φ assign a to "$\grave{\varepsilon}f(\varepsilon)$" and b to "$\grave{\alpha}g(\alpha)$". Let h be a function such that

(i) h(a) is the True,

(ii) h(the True) is a,

(iii) h(b) is the False,

(iv) h(the False) is b, and,

(v) for every argument x distinct from these, $h(x) = x$.

Finally, let φ' be an assignment of objects to course-of-values names related to φ as in (P1), with respect to the particular permutation h just specified. Then, as in (P1), φ' will satisfy Axiom V.

By appealing to (P2), Frege concludes that we are free to stipulate, without contradicting Axiom V, that the True shall be identical with a course-of-values of an arbitrary monadic first-level function, and the False with a course-of-values of any other extensionally non-equivalent function of the same type. Thus, he proposes to remove the referential indeterminacy of course-of-values names by making precisely such a stipulation: he identifies the True and the False with their own unit classes. The truth-values figure now as values of the course-of-values function $\grave{\varepsilon}\varphi(\varepsilon)$ for certain arguments and, hence, as objects satisfying Axiom V.

Benacerraf writes: "[Frege] then picks a particular course-of-values and stipulates that *it* is to be the True and another the False. [...] If we call the one he picked 'George,' then 'George = the True' lacked a truth-value before he did the picking, and acquired the True as its value from the pick. But had Frege not picked George but something else instead, 'George = the True' would have been false. Since George then figures in every course-of-values, he figures in the extension of every (non-empty) concept. Had he not been the lucky one chosen, the extension of every concept would have been different" (31).

Firstly: I think it is reasonable to suppose that for Frege (*p*) "$\dot{\varepsilon}(-\varepsilon) = (\neg\overset{\mathfrak{a}}{\smile}\!-\alpha = \alpha)$" had a determinate truth-value (probably the False) before he identified the True with the extension of the concept $-\xi$ and independently of our means of ascertaining that truth-value. The fact that Axiom V is powerless to decide the truth-value of (*p*), does not, in Frege's view, deprive (*p*) of its truth-value. In sharp contrast to this supposed realist attitude, Frege manipulates, as it were, the truth-values of (*p*) and (*q*) "$\dot{\varepsilon}(\varepsilon = (\neg\overset{\mathfrak{a}}{\smile}\!-\alpha = \alpha)) = (\neg\overset{\mathfrak{a}}{\smile}\!-\alpha = \alpha)$" by making his stipulation in § 10. Benacerraf seems to be aware of this conflict when he says that a straightforwardly "realist" construal of Frege's intentions will not do justice to his practice (31).

Let us take a closer look at this important issue. To have shown that the truth-values can be identified with their unit classes without falling prey to an inconsistency in the face of Axiom V is, in effect, Frege's central achievement in § 10 of the *Basic Laws*. In an intricate footnote to this Section, however, he appears to be calling into question the legitimacy of his additional stipulation. There he considers the possibility of generalizing it so that all objects whatever, including those already referred to by course-of-values terms, are identified with their unit classes. The suggestion goes awry. Frege rejects it on the grounds that it may contradict the criterion of identity for courses-of-values embodied in Axiom V, if the object to be identified with its unit class is already given as a course-of-values. At the same time, he jettisons the intuitively appealing proposal of identifying all and only those objects, which are not given as courses-of-values, with their unit classes. He does so on the grounds that the mode of presentation or designation of an object must not be regarded as an invariant property of it, since the same object can be given in different ways. Does this line of argument carry conviction? Does it square with Frege's additional stipulation in § 10 of the *Basic Laws*?

The question whether or not the tentative stipulation $\dot{\varepsilon}(\Delta = \varepsilon) = \Delta$ can be extended consistently to objects already designated by course-of-values terms proves to be irrelevant for any envisaged solution of the indeterminacy problem concerning course-of-values terms. The question whether, e.g., $\dot{\varepsilon}(\dot{\varepsilon}(\varepsilon = (\varepsilon = \varepsilon)) = \varepsilon)$ coincides with Julius Caesar poses the same problem as the question whether $\dot{\varepsilon}(\varepsilon = (\varepsilon = \varepsilon))$ is identical with the Roman general who crossed the Rubicon. Furthermore, even if such an extension were possible, it would not by itself license the identification of an object Δ not given as a course-of-values with its unit class. This holds especially from the perspective of Frege's platonism. If Frege's tenet that the way in which an object is given is not an invariant property of it is sound —and I believe it is— then it undermines not only the idea of generalizing the stipulation concerning the truth-values, but also that very stipulation itself. For an object not designated by a course-of-values term such as $\neg\!\overset{\mathfrak{a}}{\smile}\!- \mathfrak{a} = \mathfrak{a}$ may yet *be* a course-of-values and, in particular, a course-of-values distinct from its unit class. We must conclude, by Frege's own standards I am afraid, that the identification of the True and the False with their unit classes is a *faux pas*. So much the worse for his foundational project.

Secondly: The last two claims of Benacerraf in the quotation above are implausible. They seem to rest on the erroneous assumption that for Frege the course-of-values of a function f is a class or collection of ordered pairs of arguments and function-values: $\{\langle x,y\rangle : f(x) = y\}$. Yet, even if Benacerraf's interpretation were faithful to Frege's conception, which it is not, it would not be the case that $\dot{\varepsilon}(-\varepsilon)$ (i.e., the True) figures in every course-of-values; in fact, $\dot{\varepsilon}(-\varepsilon)$ would figure only in the extension of every non-empty concept or relation.

3. Critical remarks on Hodes

It is unfortunate that Hodes' basically interesting comparison between Frege's treatment of numerical terms in the *Foundations* and of course-of-values terms in the *Basic Laws* in his article 'Logicism and the Ontological Commitment of Arithmetic' is marred by a number of inaccuracies and misleading remarks. In what follows, I shall deal with some of them. For simplicity, I use my own instead of Hodes' notation.

Hodes correctly claims that Frege considered instances of Hume's principle (T) as expressing the same thought (or judgeable content). Hodes goes on to say that in § 66 of the *Foundations* Frege suggests that the instances of (T) so conceived determine a unique assignment of senses to all terms of the form "$N_x F(x)$". Since sense determines reference, these instances determine the *standard numberer* to be that numberer thereby associated with "N". What Hodes calls a *standard numberer* is a second-level cardinality function assigning to a first-level concept the number of objects falling under that concept; for instance, $N_x\varphi(x)$ assigns the number 9 to the concept *planet*. Now, the first claim concerning § 66 is plainly false. Frege's chief concern in that Section is to explain why (T) (or the proposed contextual definition of "$N_xF(x)$") *fails* to fix uniquely the reference of "$N_xF(x)$". (Here I transfer again to the numerical operator what Frege says about the direction-operator.) Surprisingly, the ensuing two passages in Hodes' exposition suggest that he is perfectly aware of Frege's critical assessment of the tentative contextual definition (I). So Hodes might simply have confounded § 66 with § 65.

However this may be, another point remains to be criticized. Frege is, first and above all, concerned to determine the *reference* of "$N_xF(x)$". Furthermore, in the *Foundations,* we encounter his later distinction between *sense* and *reference* at most in a very crude form, lacking any terminological rigour. The truth is that here Frege is still indulging in a freewheeling use of both terms, though he uses them perhaps not always interchangeably. Be it mere accident or for some hidden reason, at least in the course of expounding his context principle he applies the term "sense" only to sentences and reserves the term "meaning" ("Bedeutung") for words. By way of contrast, he employs the term "content" with respect to both sentences and words. In his letter to Husserl of 24.5.1891, Frege informs us that in § 97 of the *Foundations* he would now prefer to speak of "having a reference [Bedeutung]" instead of "having a sense". Accordingly, in §§ 100, 101, 102, he would now replace "sense" by "reference". At any rate, to operate in the context of the *Foundations* with Frege's later doctrine that sense determines reference, as Hodes does, is certainly unjustified.

Hodes writes: "In *The Basic Laws of Arithmetic,* Frege avoids the set-theoretic analog of the problems posed by (iv) [i.e., by a sentence like "The number of Jupiter's moons = England"] by restricting himself to a language in which all singular terms are course-of-value abstracts. He then retells a familiar story: all instances of this schematic equation are true:

the thought that $(\dot{\varepsilon}.f(\varepsilon) = \dot{\varepsilon}.g(\varepsilon))$ = the thought that $(\forall\varepsilon)\,(f(\varepsilon) = g(\varepsilon))$

these simultaneous identities uniquely determine the senses of all course-of-value abstraction terms. So the basic act is still that of § 66 of *The Foundations,* replayed on a wider stage [...] Here Frege's positive account ends" (1984, 136f.).

These remarks are scarcely enlightening. They provoke a number of criticisms.

Firstly: In the *Basic Laws,* Frege by no means confines himself to a formal language in which all singular terms are course-of-values names. Rather, his stipulations allow the construction of singular terms of various other kinds: truth-value names, definite descriptions and numerical terms. If Hodes wants to convey that, owing to Frege's stipulation in § 10 of the *Basic Laws,* every well-formed object-name of the formal language refers to a course-of-values, his way of putting it would be quite misleading. Admittedly, the formal language of the *Basic Laws* prohibits the formation of a sentence like "$\dot{\varepsilon}(-\!\!-\varepsilon)$ = England". This does not guarantee, however, that no course-of-values name may denote England. It is, moreover, undeniable that regardless of the different settings which the *Foundations* and the *Basic Laws* provide, the problem raised by a sentence like "$\dot{\varepsilon}(-\!\!-\varepsilon) = \neg\overset{\mathfrak{a}}{\smile}\!-\alpha$" in § 10 is intimately related to that posed by a sentence like "The number of planets = Julius Caesar" in § 66 of the *Foundations.* The former problem, as a matter of fact, appears as a variation of the latter in a formal dress.

Secondly: In the *Basic Laws,* Frege stipulates that both sides of Axiom V shall have the same reference, though he does not explicitly deny that they may have the same sense. Now if Frege really believed, as Hodes maintains, that both halves of Axiom V express the same thought, it would strike me as inscrutable why in the Preface to the *Basic Laws* he voices misgivings about that axiom. On the face of it, Axiom V, in view of its close formal kinship with the tentatively proposed definition (I) in the *Foundations,* was likely to arouse suspicion anyway. But neither in the Preface nor in the Appendix to the *Basic Laws* Frege was worried about that, because he believed to be able to solve the Julius Caesar problem for courses-of-values by means of his stipulation in § 10. What did cast serious doubt upon Basic Law V was rather its lack of the self-evidence which his other axioms possess. In the Appendix, Frege confesses that he had never concealed from himself this weakness. And we know, of course, that he held Axiom V responsible for the inconsistency of his logical system. Plainly, if Frege tacitly assumed that both sides of

Axiom V have the same sense, he could hardly have believed that it lacked self-evidence and caused the contradiction in his system.

Furthermore, everybody endorsing Hodes' view owes us a plausible explanation as to how Axiom V may have escaped the threat of epistemic triviality; for if its both sides were to express the same thought, it could be converted into a statement of the form "a = a". I hasten to add that, in Frege's opinion, an axiom, be it a constituent of a geometrical or a logical theory, must contain real knowledge (cf. CP, 274, 277; KS, 236, 265). This applies notably to the traditional, Euclidean conception of axioms which Frege adopts expressly as his own. It is quite true that in the debate with his antagonist Hilbert on the axiomatic method Frege is chiefly concerned with the methodological status of the axioms of (Euclidean) geometry. Yet, he seems to be suggesting that logical axioms, too, ought to contain real knowledge or possess epistemic value. One might be tempted to object that, if this really were Frege's position, it would be difficult to see on what grounds he could have construed a logical axiom like $\begin{array}{l}\vdash\!\!\top\, a\\ \;\;\llcorner\, a\end{array}$ as possessing epistemic value. (It is one of the two axioms of he *Basic Laws* which Frege lists under I.) I cannot go into this issue here, but refer the reader to the Chapter on Axiom V in my 1995.

Thirdly: Frege does not, of course, contend that his Basic Law V uniquely determines the senses of all course-of-values names. Nor does his "positive account" concerning course-of-values terms end with his stipulation in § 3 of the *Basic Laws,* as Hodes (1984, 137) suggests. On the contrary, Frege proposes to overcome the referential indeterminacy of course-of-values terms, to which this stipulation gives rise, by pursuing the method I described earlier.

Hodes goes on to criticize Frege as follows. To call the transformation of a generalized equality between function-values into an identity between courses-of-values a "fundamental law of logic" does nothing to explain our success at referring to such objects. This is undoubtedly correct. Yet, even from a cursory reading of the quotation taken from "Function and Concept" (TF, 26; KS, 130) and the appertaining context it is likewise clear that Frege does not intend to furnish any such explanation. Hodes then quotes a "revealing footnote", in which Frege allegedly "appeals to an ethereal sort of ostention": "In general, we must not regard the stipulations in Vol. i, with regard to primitive signs, as definitions. Only what is logically complex can be defined; what is simple can only be pointed to" (TF, 180; GGA II, 148). Hodes comments as follows: "This sounds like an appeal to Kantian pure intuition—a desperate move, given Frege's emphasis elsewhere on the difference between laws of logic and what intuition offers us. In any case, if such ostention were possible for courses-of-values, it should also be available directly for the cardinal numbers themselves. I suspect that this explains why Russell's paradox seemed so devastating, not just to Frege's set-theoretic approach in *The Basic Laws,* but to the very thesis that cardinal numbers are objects" (137).

To begin with, it is not clear to me what precisely is supposed to explain the disastrous impact which Russell's discovery had on Frege's set-theoretic approach, nor in what exact sense Russell's paradox threatened Frege's conception of numbers as objects. Besides creating this uncertainty, Hodes misinterprets Frege's remarks in the footnote completely. When Frege says that what is simple cannot be defined but only pointed to, he is explicitly referring to the elucidations which he provides for the primitive, logically simple function-names of his system, with the notable exception of the course-of-values operator. These elucidations merely stipulate what values a given primitive function takes on for suitable arguments. In making these semantic stipulations, Frege in no way seeks refuge in ostention or a Kantian pure intuition, and there is no need for him to do so. In particular, intuition is not allowed to creep in when logical objects are to be introduced such as courses-of-values or cardinal numbers as special courses-of-values. To reproach Frege with a "desperate move" is, therefore, certainly misplaced here.

I come to my final objection to Hodes' exposition. He believes to have shown that the inconsistency of the system of the *Basic Laws* was only a minor flaw in Frege's logicist programme. "Its fundamental flaw was its inability to account for the way in which the senses of number terms are determined" (139). It seems to me that here Hodes is reversing the true order of things. There can be no serious doubt that the emergence of the contradiction in the *Basic Laws* turned out to be a fatal flaw in Frege's project. It is equally true, however, that no contradiction could have arisen in Frege's system, if in § 31 of the *Basic Laws* he had succeeded in carrying out a valid proof of referentiality for all well-formed expressions of his formal language including course-of-values terms. Frege's first reaction to Russell's startling discovery suggests that he had more than an inkling of this interconnection: "It seems accordingly that the transformation of the generality of an equality into an equality of courses-of-values (§ 9 of my *Basic Laws)* is not always permissible, that my law V (§ 20, p. 36) is false, and that my reasonings in § 31 do not suffice to secure a reference for my combinations of signs in all cases" (PMC, 132; WB, 213). In fact, Frege's proof of referentiality founders irremediably. In particular, his attempt to prove the primitive name "$\dot{\varepsilon}\varphi(\varepsilon)$" to be referential miscarries (see Bartlett 1961, Thiel 1975, Resnik 1986, Dummett 1991, Schirn 1995).

Universität München

REFERENCES

I use the following abbreviations for references to Frege's works:

BLA: *The Basic Laws of Arithmetic. Exposition of the System,* trans. and ed. M. Furth, University of California Press, Berkeley and Los Angeles, 1964.

CP: *Collected Papers on Mathematics, Logic, and Philosophy,* ed. B. McGuinness, trans. M. Black et al., Oxford: Basil Blackwell, 1984.

FA: *The Foundations of Arithmetic. A logico-mathematical enquiry into the concept of number,* trans. J. L. Austin, Oxford: Basil Blackwell, 1950, 1953.

GGA: *Grundgesetze der Arithmetik, Begriffsschriftlich abgeleitet,* Vol. I, H. Pohle, Jena, 1893, Vol. II, Jena: H. Pohle, 1903, reprinted Darmstadt and Hildesheim, 1962.

GLA: *Die Grundlagen der Arithmetik. Eine logisch mathematische Untersuchung über den Begriff der Zahl,* Breslau: W. Koebner, 1884, reprinted Darmstadt and Hildesheim, 1961.

KS: *Kleine Schriften,* ed. I. Angelelli, Hildesheim: G. Olms, 1967.

NS: *Nachgelassene Schriften,* ed. H. Hermes, F. Kambartel and F. Kaulbach, Hamburg: Felix Meiner, 1969.

PMC: *Philosophical and Mathematical Correspondence,* ed. B. McGuinness, trans. H. Kaal, Oxford: Basil Blackwell, 1980.

TF: *Translations from the Philosophical Writings of Gottlob Frege,* ed. and trans. P. Geach and M. Black, Oxford: Basil Blackwell, 1952, 1960.

WB: *Wissenschaftlicher Briefwechsel,* ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel and A. Veraart, Hamburg: Felix Meiner, 1976.

I refer by author and year of publication to the following works:

Bartlett, James, M. 1961: *Funktion und Gegenstand. Eine Untersuchung in der Logik von Gottlob Frege*, Phil. Diss. München.

Benacerraf, Paul 1981: "Frege: The Last Logicist", Midwest Studies in Philosophy VI, eds. P. French, T. Uehling, and H. Wettstein, Minneapolis: University of Minnesota Press, pp. 17–35.

Dummett, Michael 1991: *Frege: Philosophy of Mathematics*, London: Duckworth.

Hodes, Harold 1984: "Logicism and the Ontological Commitments of Arithmetic", *The Journal of Philosophy*, **81**: 123–149.

Resnik, Michael 1986: "Frege's Proof of Referentiality", in Haaparanta, Leila and Hintikka, Jaakko (eds.), *Frege Synthesized. Essays on the Philosophical and Foundational Work of Gottlob Frege*, Dordrecht/Boston: D. Reidel.

Schirn, Matthias 1995: *Gottlob Frege. Grundlagen der Mathematik*, forthcoming, Berlin/New York: Walter de Gruyter.

Thiel, Christian 1975: "Zur Inkonsistenz der Fregeschen Mengenlehre", in C. Thiel (ed.), *Frege und die moderne Grundlagenforschung. Symposium gehalten in Bad Homburg im Dezember 1973*, Meisenheim am Glan: Anton Hain, pp. 134–159.