Amending Frege's *Grundgesetze der Arithmetik* [draft]

To the memory of Nhê (1925-2001)

Fernando Ferreira
Departamento de Matemática
Universidade de Lisboa
Rua Ernesto de Vasconcelos, C1-3
P-1749-016 Lisboa, Portugal
ferferr@cii.fc.ul.pt

September 25, 2002

Abstract. Frege's *Grundgesetze der Arithmetik* is formally inconsistent. This system is, except for minor differences, second-order logic together with an abstraction operator governed by Frege's Axiom V. A few years ago, Richard Heck showed that the ramified predicative second-order fragment of the *Grundgesetze* is consistent. In this paper, we show that the above fragment augmented with the axiom of reducibility for concepts true of only finitely many individuals is still consistent, and that elementary Peano arithmetic (and more) is interpretable in this extended system.

Keywords: Predicative definitions, finite reducibility, Frege, logicism.

1. Introduction

"How did the serpent of inconsistency enter Frege's paradise?" asks Michael Dummett in the opening of a section of his book *Frege: Philosophy of Mathematics* (1991). According to the traditional view, Frege's *Grundgesetze der Arithmetik* is formally inconsistent because of the presence of an abstraction operator (viz. the value-range operator), as regulated by the (in)famous Axiom V. The formal system of the *Grundgesetze* is, except for minor differences, second-order logic augmented by Frege's Axiom V, which we may take as the scheme

$$\text{(AxV)} \qquad \hat{x}.A(x) = \hat{x}.B(x) \leftrightarrow \forall x(A(x) \leftrightarrow B(x)),$$

where $A(x)$ and $B(x)$ are arbitrary formulas of the language. In the above, the *value-range* operator $\hat{}$ yields a first-order term $\hat{x}.A(x)$ when applied to a formula $A(x)$. Axiom V states a kind of extensionality principle: The value-range of the concept given by the formula $A(x)$ is the same as the value-range of the concept given by the formula $B(x)$ if, and only if, the two concepts hold of exactly the same individuals.

Russell's paradox is forthcoming in the above framework. Define $x \in y$ by the formula $\exists G(Gx \wedge y = \hat{z}.Gz)$ and consider the concept R

given by the formula $Rx :\leftrightarrow \neg(x \in x)$. The individual $\hat{x}.Rx$ is a member of itself if, and only if, it isn't: Russell *dixit*. Dummett blames the inconsistency primarily on Frege's insouciance concerning second-order quantification. He rightly adds that without second-order quantification the paradoxes would not be forthcoming,[1] although at the cost of paralysing the formal system (viz. the above membership relation cannot be defined). In his analysis of the situation, Dummett does *not* blame second-order quantification per se, but only the *impredicativity* underlying it in Frege's setting. In effect, the Russellian concept R is specified by saying that an individual x falls under it if, and only if, for all concepts G – including R itself, as it were – G is not true of x whenever the value-range of G is x. The Russellian concept R is specified in terms of a quantification that ranges over the concept being specified: this is the hallmark of an impredicative definition.

A few years ago, Richard Heck (1996) gave some weight to Dummett's position by proving that the simple predicative fragment of Frege's *Grundgesetze* is consistent. In other words: On a predicative reading, we may coherently assume that every concept has an extension. Heck's predicative fragment H is obtained from Frege's system by suitably restricting the comprehension scheme (a tacitly used instantiation rule in Frege's original setting) as follows:

$$\exists G \forall x (Gx \leftrightarrow A(x)),$$

where $A(x)$ is any formula not containing the variable G and *not containing bound second-order variables*. It is therefore pertinent to ask how much arithmetic (and analysis) can be developed within H. Heck proves that Raphael Robinson's theory of arithmetic Q is interpretable in the simple predicative fragment. This shows that the predicative fragment is not trivial, since Q is a well-known example of an essentially undecidable theory.[2] Very recent work has also shown that Q is not trivial in another respect, viz. in a more mathematical sense. The present author and António Fernandes showed in (2002) that Tarski's elementary theory of the real-closed ordered fields (informally: elementary algebra and analysis, including analytic geometry) is interpretable in Q, and hence in H. These facts not withstanding, the theory Q is proof-theoretically very weak, and we believe that the prospects for interpreting (say) primitive recursive arithmetic in the predicative fragment are slim.

Heck also shows that the *ramified* predicative fragment of Frege's *Grundgesetze* is consistent.[3] In the present paper, we show that this ramified fragment augmented with the axiom of reducibility for concepts true of only finitely many individuals is still consistent and able to interpret elementary Peano arithmetic. We suggest that the axiom of finite reducibility – as opposed to (full) reducibility in settings like that of Whitehead and Russell's *Principia Mathematica* (1925) – is an analytic truth, that is to say, true in virtue of the meaning of the notions involved. In the closing section, we briefly compare our system with Crispin Wright's system based on Hume's principle.

2. Finite reducibility

The language of pure ramified second-order logic with equality is that extension of the language of pure first-order logic with equality obtained by adding an infinite stock of second-order variables (concept variables) $F^0, G^0, \ldots,\ F^1, G^1, \ldots,$ etc., $F^n, G^n, \ldots$ and the like, for each (natural number) superscript n, the so-called *level* of the concept variable. Concept variables behave syntactically like unary predicades in ordinary first-order logic, with the conspicuous exception that we may quantify over them. We should further note that equality is a primitive symbol that infixes only between first-order variables. This being said, the formulas of pure ramified second-order language are built up in the usual way, as in (Heck, 1996). The comprehension axioms of second-order ramified predicative logic consist of all formulas of the form

(PComp) $$\exists G^n \forall x (G^n x \leftrightarrow A(x)),$$

where G^n is a n-th level concept variable not occuring in $A(x)$, with the proviso that the formula $A(x)$ contains no bound variables of levels greater than or equal to n, and no free variables of level greater than n (the latter variables, as well as the first-order free variables that might also appear in $A(x)$, are called the *parameters* of this particular instance of the comprehension scheme). Heck's ramified system also includes the value-range operator which, when applied to any formula $A(x)$, yields a first-order term $\hat{x}.A(x)$ in which x and the bound variables of $A(x)$ are bound. The value-range operator is regulated by axiom scheme (AxV). A proper treatment of these syntactical matters demands a definition of 'term' and 'formula' by simultaneous induction, but this is straightforward. By definition, the comprehension axioms (PComp) also apply to formulas of the extended Fregean language, with the proviso also including the free and bound variables of the value-range operators. The theory which concerns this paper is Heck's ramified theory augmented with the scheme of finite reducibility. We will call this theory HFR. The following paragraphs describe the scheme of finite reducibility.

We have followed Heck (1996) in setting up the Fregean language with concept variables of one argument-place only, although Frege himself also uses binary concept variables (binary relational variables) in the *Grundgesetze*. While the restriction to (unary) concept variables simplifies notation enormously, it does not, in the context of Heck's ramified predicative fragment, represent any loss of generality. We may introduce ordered pairs via the well-known Wiener-Kuratowski definition:

$$\langle u, v\rangle := \hat{x}.(x = \hat{y}.(y = u) \vee x = \hat{y}.(y = u \vee y = v)).$$

In the presence of (AxV), one easily proves

$$\forall x \forall y \forall u \forall v(\langle x, y\rangle = \langle u, v\rangle \leftrightarrow (x = u \wedge y = v)).$$

We can therefore speak of binary relational variables simply by explicating them in terms of (unary) concept variables true of the pertinent ordered pairs. We use the infix notation xRy for saying that the (unary) concept R is true of the pair $\langle x, y\rangle$. We may also say that concept H is linearly ordered by a relation R, and write $Lin(R, H)$. By this, we mean that the elements of R are pairs of the form $\langle x, y\rangle$, with H true of x and y, and that these pairs form an anti-reflexive, transitive and trichotomic relation. Notice that the quantifications that occur in the formula $Lin(R, H)$ are all first-order. The same is the case with the formulas $Min(x, R)$ and $Max(y, R)$ that naturally represent the properties that x is the least and (respectively) y is the greatest element of the linear order R.

In 1907, Paul Stäckel defined a finite set as one which can be *doubly well-ordered*, that is, one for which there is a (necessarily linear) ordering with respect to which every non-empty subset has both a least and a greatest element.[4] Our formal definition of the notion of a zeroth level concept H^0 being true of only finitely many individuals is an adaptation of Stäckel's definition. Let $Dwo(R^0, H^0)$ be the conjunction of $Lin(R^0, H^0)$ with

$$\forall G^0(\emptyset \neq G^0 \subseteq H^0 \rightarrow \exists x\, Min(x, R^0|_{G^0}) \wedge \exists y\, Max(y, R^0|_{G^0})),$$

where $\emptyset \neq G^0 \subseteq H^0$ abbreviates $\exists x G^0 x \wedge \forall x(G^0 x \rightarrow H^0 x)$ and $R^0|_{G^0}$ is the relation true of the pairs $\langle x, y\rangle$ in R^0 such that G^0 is true of both x and y. A *finite* zeroth level concept is formally defined as follows:

$$Fin^0(H^0) :\leftrightarrow \exists R^0 Dwo(R^0, H^0), \tag{1}$$

We are now ready to state the *axiom scheme of finite reducibility* (one single axiom for each level n):

$$\forall H^0 \forall F^n(Fin^0(H^0) \wedge F^n \subseteq H^0 \rightarrow \exists G^0 \forall x(F^n x \leftrightarrow G^0 x)), \tag{2}$$

where $F^n \subseteq H^0$ abbreviates $\forall x(F^n x \rightarrow H^0 x)$. This scheme formalizes the idea that a sub-concept of a finite concept must be finite and, therefore, must be co-extensional with a zeroth level concept (see the discussion in the next section). It is easy to check that this zeroth level concept is finite according to our formal definition (1). In a similar vein, we define finitude for non-zero level concepts H^n:

$$Fin^n(H^n) :\leftrightarrow \exists G^0(Fin^0(G^0) \wedge \forall x(H^n x \leftrightarrow G^0 x)), \qquad (3)$$

Suppose that $Fin^0(H^0)$ holds and that R^0 is a double well-ordering of H^0 in the sense of the definition of finiteness for zeroth level concepts, i.e., such that every non-empty zeroth level sub-concept of H^0 has an R^0-least element and an R^0-greatest element. Then, in fact, every non-empty sub-concept of H^0, zeroth level or not, has an R^0-least element and an R^0-greatest element. This follows from the axiom scheme of finite reducibility since the sub-concept in question is co-extensional with a zeroth level concept.

3. On the notion of finiteness

The axiom of reducibility states that *every* concept is co-extensional with a predicative (i.e., zeroth level) concept. In our Fregean setting, it is not difficult to see that (full) reducibility engenders a Russell type inconsistency. On the other hand, the axiom of reducibility restricted to concepts true of only finitely many individuals is arguably an analytic truth, that is to say, true in virtue of the meaning of the notions involved. Let us spell out the argument. If a concept is true of only finitely many individuals then, on the intended interpretation, it is true of only the items of a finite list of pairwise distinct individuals, say $a_1, a_2, \ldots, a_k$. Thus, the concept has a zeroth level co-extensional definition, viz. $x = a_1 \vee x = a_2 \vee \ldots \vee x = a_k$.[5,6]

In our formal setting, the principle of finite reducibility is given by the mutual contributions of two definitions and an axiom scheme: (1) the definition of finiteness for zeroth level concepts; (2) the scheme of finite reducibility *strictu sensu*; and (3) the definition of finiteness for non-zeroth level concepts. As argued, if one is justified in accepting (1) as a proper characterization of finiteness (for zeroth level concepts), then one is justified in accepting (2) and (3).

Whereas the adaptation of Stäckel's definition of a finite set (see previous section) to the framework of second-order logic does not raise any particular questions, its use in ramified predicative second-order logic poses a problem. Take our definition (1) in isolation: Even though every non-empty zeroth level sub-concept of G^0 has an R^0-least element and an R^0-greatest element, it is conceivable that this is no longer true of higher level sub-concepts. For sure, we have argued in the end of the previous section that this very possibility is precluded. However, our argument rested on the acceptance of (2), a principle whose accepting grounds lie precisely on the acceptance of (1) as a correct rendering of the notion of finiteness. We are in a circle. A virtuous circle, it may well be argued, but a circle nonetheless.

The circle can be broken. Let us grant that the notion of a concept being true of only finitely many individuals is meaningfully available for use in any (interpreted) language based on first-order logic with equality. We claim that definition (1) in isolation faithfully captures that notion (for zeroth level concepts) *provided* that in interpreting our language we require that the class of zeroth level concepts be closed under the operator "there are finitely many."[7] Let us see why. Consider an interpretion of our language and let us use the quantifier notation $\mathsf{F}xHx$ for expressing the ordinary notion that concept H is true of only finitely many individuals. Suppose that $Fin^0(H^0)$ holds. We must argue that $\mathsf{F}xH^0x$ also holds (the converse statement is straightforward: see endnote 6 and the discussion which it affixes). By hypothesis, let R^0 be a double well-ordering of H^0 witnessing (1). Consider the concept:

$$Gx :\leftrightarrow H^0x \wedge \mathsf{F}z(zR^0x).$$

We have mandated that this concept G be available at the zeroth level. Moreover, if H^0 is true of something then G is also true of something, namely of the R^0-least element of H^0. Therefore, G has an R^0-greatest element. It is easy to argue that this maximum element is, in fact, the R^0-greatest element of H^0. This entails that $\mathsf{F}xH^0x$, as wanted.[8]

Let us pause for a moment. We showed that if finiteness is a warranted notion – meaningfully available for use in any interpreted language based on first-order logic with equality – then it has a purely logical definition in ramified predicative second-order logic, viz. the one given by (1) and, for higher level concepts, by (3). It would take us far afield to pin down the idea of a *warranted* notion. At the moment, I am only prepared to say that a warranted notion must not engender contradictions. This very weak injunction may, nevertheless, baffle some readers. After all, if a notion is at all formalizable within a language (in our case, within the language of pure ramified predicative second-order logic) then it should be so by means of a *definition*, and definitions cannot engender contradictions. Plainly, a mere abbreviation like (1) cannot engender a contradiction. However, if (1) – as argued – is indeed a proper formalization of the ordinary notion of a concept being true of only finitely many individuals, then on reflection upon that ordinary notion and upon our uses of language we are bound to accept the scheme of finite reducibility (2). We may say, using a catchy phrase,[9] that the axiom scheme of finite reducibility *makes explicit* within language certain uses of language. Namely: We do use the fact that concepts true of only finitely many individuals are co-extensional with concepts expressed by suitable disjunctions. When the individuals falling under these concepts are listed explicitly, such disjunctions may be given forthwith; otherwise, they are usually indicated by means of the meta-linguistic device of inserting ellipsis points. *A fortiori*, these disjunctions yield concepts belonging to the zeroth level fragment of the language. It is *this* aspect of finitude (together with the feature that a sub-concept of a concept true of only finitely many individuals is true of only finitely many individuals) that the axiom of finite reducibility (2) makes explicit *within* language.

In saying that a warranted notion – formalized via a definition – must not engender contradictions, we are using the word 'definition' in a wider (and more vague) sense than that of a mere abbreviation.[10] If we concede that contextual definitions (in the sense of Frege) are definitions in this wider sense,[11] then the ominous example of an unwarranted notion is that of the extension of a concept, as regulated by Axiom V, in the framework of second-order logic.

Reassuringly, we have the following:

THEOREM 1. *The theory* HFR *is consistent.*

Proof. Our proof is a simple adaptation of Heck's proof of the consistency of the predicative (and ramified predicative) fragment of Frege's *Grundegesetze*. In the sequel, we assume familiarity with Heck's argument as exposed in his (1996).

The consistency of the theory HFR is proved by presenting a model of it. Therefore, we must give a suitable domain for the individual variables, give suitable domains for the concept variables, one for each level, consisting of subsets of the individual domain, and give a suitable interpretation of the value-range operator. In fact, our structure is the reduct of a structure for a slightly expanded language, viz. the one obtained from the language of pure ramified second-order logic by adding a constant symbol n for each natural number n and binary function symbols $+$ and $\times$.

Following Heck, the individual domain of our structure is the set of natural numbers. The extra constants and function symbols are interpreted in the usual arithmetical manner. The next step consists in interpreting the value-range terms in which concept variables are wholly absent. Following Terence Parsons (1987b), Heck defined such interpretations satisfying Frege's axiom V, with the further property that the range of these value-range terms form a co-infinite subset of the natural numbers. We let the domain of the zeroth level concept variables be constituted by the definable subsets of the natural numbers obtained by means of formulas $A(x)$ which contain no free variables other than 'x' and which contain no concept variables at all (note that these formulas may contain value-range terms which do not contain themselves concept variables). It is now easy to show that this interpretation determines the interpretations of the value-range terms in the zeroth level fragment of the ramified language in which concept quantification does not occur (concept parameters may appear). The scheme (PComp) can now be evaluated for $n = 0$, and it is easy to see that it turns out to be true. Finally, Heck shows how to extend the interpretation to all the value range terms of the zeroth level fragment of the language (i.e., also to those value-range terms which have quantified zeroth level concept variables). The extended interpretation also satisfies Frege's axiom V and is such that the range of these value-range terms is still co-infinite.[12]

Thus far, we have fully interpreted the zeroth level fragment of the ramified language. The extension of this interpretation to the higher reaches of ramification can be obtained by a simple iteration of the above procedure. This iteration is thoroughly explained in (Heck, 1996). As argued by Heck, we end up with a model of the ramified predicative fragment of Frege's *Grundgesetze*. It now remains to see that this model also satisfies the axiom scheme of finite reducibility (2).

The sole difference between the above structure and Heck's structure is the inclusion of arithmetical function symbols $+$ and $\times$ in our starting language, interpreted ordinarily. Therefore, the arithmetical sets (i.e., the sets of natural numbers definable by formulas of the language of first-order Peano arithmetic) are included in the domain of variation of the zeroth level concept variables. Pick your favorite arithmetical coding procedure in order to code finite sets by natural numbers. More specifically: Take arithmetical formulas $Code(z)$ and $Elem(x, z)$, with their free variables as displayed, such that for each finite set F of natural numbers there is a natural number c so that '$Code(\mathsf{c})$' is true and, for all natural numbers n, $n \in F$ iff the sentence '$Elem(\mathsf{n}, \mathsf{c})$' is true. Conversely, if c is a natural number such that the sentence '$Code(\mathsf{c})$' is true, then the set of natural numbers n for which the sentence '$Elem(\mathsf{n}, \mathsf{c})$' holds true is finite. It is now easy to see that the domain of variation of the zeroth level concept variables is closed under the application of the operator "there are finitely many." In effect, the claim that there are finitely many individuals x such that $A(x)$ holds is formally expressible by the following first-order clause:

$$\exists z(Code(z) \wedge \forall x(A(x) \leftrightarrow Elem(x,z))).$$

The argument in the fourth paragraph of this section shows that the above closure property entails that zeroth level concepts H^0 such that $Fin^0(H^0)$ holds in the structure are true of only finitely many individuals. As we saw, this implies that the axiom scheme of finite reducibility (2) holds in our structure. □

We shall see in the next section that HFR is able to interpret elementary Peano arithmetic. Technically, the strength of our formal definition (in the wide sense of the word 'definition') of finiteness lies squarely with the scheme of finite reducibility: If a contradiction can be engendered by our formal notion of finiteness, then it is engendered by the axiom scheme (2). Even though we have just proved that no such contradiction is forthcoming, our proof is merely relative to a (say) suitable fragment of set theory. We know that we cannot aspire to more on this regard. Such is life after Gödel.

4. Arithmetic in the amended system

Although Frege's *Grundgesetze* is inconsistent, this does not (by itself) render worthless the *way* Frege derived the axioms of second-order arithmetic from it (together with appropriate definitions of the arithmetical notions). It was first noted by Charles Parsons in his (1965) that Frege's proof of the axioms of second-order arithmetic proceeds in two independent steps. Firstly, Frege uses the value-range operator (as regulated by Axiom V) to define a cardinality operator # that yields first-order terms $\#x.A(x)$ when applied to formulas $A(x)$, and that satisfies Hume's Principle,

(HP) $$\#x.A(x) = \#x.B(x) \leftrightarrow A \approx_x B,$$

where $A \approx_x B$ is a formula saying that the As are in one-one correspondence with the Bs. Secondly, Frege essentially shows that the axioms of second-order arithmetic (suitably interpreted) follow from second-order logic together with Hume's Principle. The point is that the value-range operator and Axiom V quietly drop out of sight in this second stage. The theory consisting of second-order logic augmented with a *primitive* operator # satisfying Hume's Principle (HP) is called Frege's Arithmetic FA. The second step described above is now known as *Frege's theorem.*[13]

The proof of Frege's theorem uses impredicative definitions through and through. There are two quite different reasons for this. On the one hand, Frege's theorem shows that second-order arithmetic – a blatantly impredicative theory with full comprehension – is interpretable in FA. No wonder that impredicative definitions are essential in this regard. On the other hand, Frege's definition of the natural numbers uses his well-known notion of the *ancestral* of a given relation, a prima facie impredicative notion. This wouldn't be decisive were it not for the fact that in the proof of Frege's theorem one needs to use the notion of the ancestral of the predecessor relation as a *concept* for reasons *other than* proving comprehension. More specifically, one needs the *concept* of the ancestral of the predecessor relation in order to prove the induction axioms (even for proving arithmetical induction) and the statement that every natural number has a successor. The formation of this concept amounts to a suitable use of (prima facie impredicative) comprehension. In this section, we make two observations: (a) If the operator "there are finitely many" is predicatively justified, then the notion of the ancestral of a relation is predicative; (b) If the notion of the ancestral of a relation is predicative, then it is possible to prove a restricted form of Frege's theorem, viz. the one in which second-order arithmetic is replaced by *predicative* second-order arithmetic (more on this later). To show (a), we argue that the first level fragment of HFR is able to deal effectively with the concept of ancestrality. To show (b), we consider George Boolos' pointed reconstruction of Frege's theorem as expounded in the appendix of his (1998b) and observe that his reconstruction pulls through, almost unmodified, within the first level fragment of HFR for predicative second-order arithmetic.

For ease of notation, in the remainder of this section all unsuperscripted concept variables are first level concept variables of the language of HFR. Given a (first level) concept R, we define its *ancestral* relation R^* as follows:

DEFINITION 1. $R^*(x, y)$ *iff*

$$\forall F(\forall a \forall b(((a = x \vee Fa) \wedge aRb) \to Fb) \to Fy).$$

The relation R^* is prima facie of second level. The following proposition is crucial. The idea of its proof is straightforward: $R^*(x, y)$ can be expressed by the *first* level concept saying that there is a way of going from x to y in finitely many steps according to the given relation R.

PROPOSITION 1. *The theory* HFR *proves the sentence:*

$$\forall R \exists Q \forall x \forall y\, (R^*(x, y) \leftrightarrow xQy).$$

Proof. Let $A(L^0, H^0, R)$ abbreviate the conjunction of $Dwo(L^0, H^0)$ with $\forall a \forall b(Pred(a, b, L^0) \to aRb)$, where $Pred(a, b, L^0)$ says that a is an immediate L^0-predecessor of b, i.e., $aL^0b \wedge \neg\exists c(aL^0c \wedge cL^0b)$.[14] We claim that $R^*(x, y)$ is equivalent to

$$\exists H^0 \exists L^0(A(L^0, H^0, R) \wedge Min(x, L^0) \wedge \forall z(Max(z, L^0) \to zRy)).$$

Notice that, if the above equivalence obtains, then by (PComp) we can form a first level concept for R^*. Assume $R^*(x, y)$. Consider the (first level) concept F defined by

$$Fw :\leftrightarrow \exists H^0 \exists L^0(A(L^0, H^0, R) \wedge Min(x, L^0) \wedge \forall z(Max(z, L^0) \to zRw)).$$

In order to show Fy, suppose $a = x \vee Fa$, aRb. It is enough to show Fb. If $a = x$, let H^0 and L^0 be (respectively) the concepts $[w : w = x]$ and $[w : w \neq w]$.[15] Clearly, Fb. If Fa and $a \neq x$, then there are concepts H^0 and L^0 such that $A(L^0, H^0, R)$, $Min(x, L^0)$, and zRa for z the L^0-greatest element of H^0. There are two cases to consider. In the first case, H^0 is not true of a. Let $\check{H}^0$ and $\check{L}^0$ be the concepts $[w : H^0w \vee w = a]$ and $[\langle u, v\rangle : uL^0v \vee (H^0u \wedge v = a)]$ (respectively). Fb follows. In the second case, H^0 is true of a. Let $\check{H}^0$ and $\check{L}^0$ be the concepts $[w : wL^0a \vee w = a]$ and $[\langle u, v\rangle : uL^0v \wedge (vL^0a \vee v = a)]$ (respectively). Fb follows.

Conversely, assume $A(L^0, H^0, R)$, $Min(x, L^0)$, and zRy for the L^0-greatest element z of H^0. Let F be an arbitrary first level concept and suppose $\forall a \forall b(((a = x \vee Fa) \wedge aRb) \to Fb)$. We must show Fy. If $z = x$, then Fy follows by letting $a = x$, $b = y$. If $z \neq x$, consider the first level sub-concept $[w : Fw \wedge H^0w]$ of the finite concept H^0. By finite reducibility, this sub-concept is co-extensional with a zeroth level concept and, moreover, it is true of something, viz. of the immediate L^0-successor of x. Therefore, it has an L^0-greatest element, say u. If u isn't z, take v such that $Pred(u, v, L^0)$. We thus have Fu and uRv. Therefore, Fv. Contra the choice of u. So, u is z. Hence Fz. Fy follows by letting $a = z$, $b = y$. □

Frege's theorem is based upon Hume's Principle (HP). We use Weak Hume's Principle (WHP)[16] instead. In a nutshell, we restrict (HP) to finite concepts. Following Frege, given a finite concept H^0 we define:

$$\#H^0 := \hat{z}.\exists G^0(H^0 \approx G^0 \wedge z = \hat{x}.G^0x),$$

where $H^0 \approx G^0$ abbreviates the formula which says that there is a zeroth level one-one correspondence between the H^0s and the G^0s. With this definition, we clearly have:

(WHP) $$Fin^0(H^0) \wedge Fin^0(G^0) \to (\#H^0 = \#G^0 \leftrightarrow H^0 \approx G^0).$$

We modify Frege's definition of the predecessor relation accordingly:

DEFINITION 2. *xPy iff*

$$\exists H^0 \exists z(Fin^0(H^0) \wedge H^0 z \wedge \#H^0 = y \wedge \#[w : H^0 w \wedge w \neq z] = x).$$

By (PComp), this relation is indeed given by a first level concept. Boolos' above mentioned reconstruction of Frege's theorem adapts to our framework. The following is accomplished: There is a first level concept $\mathbb{N}$ (the concept 'Finite' in Boolos' reconstruction) and an individual 0 such that the theory HFR proves,

1. $\mathbb{N}0$
2. $\mathbb{N}x \wedge xPy \rightarrow \mathbb{N}y$
3. $\forall x \forall y \forall z(\mathbb{N}x \wedge xPy \wedge xPz \rightarrow y = z)$
4. $\forall x \forall y \forall z(\mathbb{N}x \wedge \mathbb{N}y \wedge xPz \wedge yPz \rightarrow x = y)$
5. $\neg \exists x(\mathbb{N}x \wedge xP0)$
6. $\forall x(\mathbb{N}x \rightarrow \exists y\, xPy)$
7. $\forall F(F0 \wedge \forall x \forall y(Fx \wedge xPy \rightarrow Fy) \rightarrow \forall x(\mathbb{N}x \rightarrow Fx))$

In the framework of second-order logic, the above constitutes a Dedekind-Peano axiomatization of second-order arithmetic, a system with full comprehension. *That* much comprehension is simply a by-product of the *logic* in question. In our ramified predicative framework HFR, we obtain comprehension for arithmetical formulas with (individual and concept) parameters. The second-order theory of arithmetic ACA_0 (an acronym for *arithmetical comprehension axioms*) is the second-order extension of elementary Peano arithmetic PA which has comprehension for arithmetical formulas (with first and second order parameters), and the *axiom* of induction as in point 7 above. Summing up: We have interpreted ACA_0 in HFR.[17,18]

The theory ACA_0 has attracted some attention in recent years because it seems capable of formalizing most (if not all) of the mathematics that is necessary for scientific applications. I refer the reader to Stephen Simpson's book (1999) – specially to its introduction – for a discussion of this and other subsystems of second-order arithmetic, and to Solomon Feferman's articles (1998a), (1998b) and (1998c) for very incisive discussions on the philosophical importance of systems related to ACA_0 and on the scope of the indispensability arguments.

5. Brief comparisons

The inconsistency of a formal system is a global property of its axiom system. In general, no single axiom is the culprit. If in Frege's *Grundgesetze der Arithmetik* we trade the value range operator, as governed by Axiom V, for the cardinality operator, as governed by Hume's Principle, we get Frege's Arithmetic FA – a consistent system. If, on the other hand, we keep the value range operator and Axiom V, but replace second-order logic by ramified predicative second-order logic together with the axiom scheme of finite reducibility, we get HFR – again a consistent theory. These consistency results show that there are several choices – able to develop significant parts of arithmetic – for amending Frege's inconsistent system. These amendments are important insofar as they may contribute to chart the prospects for vindicating one form or other of logicism.

The neo-Fregeans – with Crispin Wright at the forefront – claim that Frege's Arithmetic is a partial vindication of logicism. Critics have advanced several problems concerning this claim. One of these criticisms concerns whether second-order logic really is *logic*. Quine famously quipped that second-order logic is set theory in sheep's clothing.[19] The question whether predicative (or ramified predicative) second-order logic is *logic* can be (and has been) put as well. However, I believe that a positive answer to this latter question faces much better prospects than that of the case of second-order logic.

A vindication of Frege's Arithmetic as a form of logicism must argue that Hume's principle is analytic in some sense or other. Hume's principle is an abstraction principle, i.e., a principle of the form

(Abst) $$\forall G \forall H (\%G = \%H \leftrightarrow Eq(G, H)),$$

where Eq is an equivalence relation between concepts. In the framework of second-order logic, abstraction principles can lead to inconsistency, as Frege's value-range operator shows. *A fortiori*, there are abstraction principles that are not analytic. The neo-Fregean must explain why certain abstraction principles are analytic (viz. Hume's principle) while others aren't. The easy answer that merely draws the line at inconsistency won't do because George Boolos showed in (1998b) that there are pairs of abstraction principles, each one of which is consistent in isolation, but which are inconsistent when put together. This is Boolos' "bad company" objection. Let us show that predicative second-order logic is immune to Boolos' objection.

In the framework of predicative second-order logic, the value-range operator is a *universal* operator in the following sense: Any abstraction operator on concepts can be defined from it. Let us define $\%G$ by

$\hat{z}.\exists H(Eq(G,H) \wedge z = \hat{x}.Hx)$. With this definition, Heck's predicative system H proves (Abst). Therefore, the bad company objection does not apply to the framework of predicative second-order logic.[20] Similar sorts of considerations also indicate that the bad company objection does not apply to the framework of ramified predicative second-order logic.

We choose to focus on criticisms of neo-Fregeanism that do not (or do not fully) apply to HFR. However, we want to mention a criticism that applies to both systems FA and HFR: The fact that both systems imply the existence of infinitely many individuals. We ourselves are won over by Boolos' view expressed in (1998a) that *no logic, no logicism*, and that logic doesn't entail the existence of infinitely many individuals. Hence, we are not proposing HFR as a vindication of logicism. We are rather putting the above criticism in between parentheses while studying and discussing the merits and weaknesses of amendments to Frege's *Grundgesetze*. Simply put: We want to *discuss the matter*.[21]

6. Acknowledgements

Cordial thanks to the organizers of the Bielefeld meeting *Philosophy of Mathematics* for their interest in my contribution. Earlier drafts of this paper were also presented to audiences at the universities of Santiago de Compostela (Spain), Aarhus (Denmark), Athens (Greece), Lisbon and Coimbra (Portugal). I wish to thank António Zilhão, Ulrich Kohlenbach, Constantinos Dimitracopoulos, Olga Pombo and Henrique Jales Ribeiro for making these presentations possible. Special thanks to my former student Sancho Silva for some helpful and enjoyable discussions, and to Roy Dyckhoff for suggesting some improvements upon reading a late draft of the paper. Finally, thanks to CMAF (Fundação de Ciência e Tecnologia) for the partial financial support of this research, and to António Sá Fonseca for the research environment he helped create at Complexo Interdisciplinar da Universidade de Lisboa.

Notes

[1] This is a result of Terence Parsons in (1987b).

[2] See, for instance, (Boolos and Jeffrey, 1990).

[3] On a different direction, we showed – in collaboration with Kai Wehmeier – that adding the principle of Δ^1_1-comprehension to the predicative fragment is consistent: see our (2002). In this paper, we conjecture that neither this fragment nor the ramified fragment are able to carry much more mathematics than the predicative

fragment. Notes 1 and 2 of (Burgess and Hazen, 1998) have some pertinent remarks on related matters.

[4] See (Parsons, 1987a) for references.

[5] Observe that the justification of axiom scheme (2) also depends on the feature that a sub-concept of a concept true of only finitely many individuals is still true of only finitely many individuals.

[6] For later reference, notice that the double well-ordering relation true of precisely the pairs $\langle a_i, a_j \rangle$, for $1 \leq i < j \leq k$, is predicative. This relation can be presented by a disjunction of $k(k-1)/2$ suitable disjuncts.

[7] Given our starting assumption, this is a meaningful requirement. We require it by *fiat*: were it unfulfilled, we would enlarge the class of the zeroth level concepts by closing it under applications of the operator "there are finitely many." To put it in other words: In our *intended* interpretations, the zeroth level concepts are closed under the application of the operator "there are finitely many."

[8] For the record, the above argument only uses three properties of the ordinary notion of finiteness. That $\neg\exists x Hx \rightarrow \mathsf{F}x Hx$, that $\mathsf{F}x Hx \rightarrow \forall y \mathsf{F}x(Hx \vee x = y)$, and that $\forall x(Hx \leftrightarrow Gx) \rightarrow (\mathsf{F}x Hx \leftrightarrow \mathsf{F}x Gx)$.

[9] *Making It Explicit* is a well-known book of Robert Brandom.

[10] Patrick Suppes' book *Introduction to Logic* presents two criteria about the character of definitions: (i) a defined symbol should always be eliminable from any formula of the theory; (ii) a new definition does not permit the proof of relationships among the old symbols which were previously unprovable. We bid good riddance to the second criterium. Our use here of the word 'definition' is *creative*, in Suppes' terminology. These two criteria are due to the Polish logician S. Leśniewski.

[11] Under this concession, we are bidding good riddance to Leśniewski's first criterion as well (see the previous footnote).

[12] The insistence on always having infinitely many natural numbers which are not interpretations of any of the value-range terms so far considered is done in order to make room for the interpretations of the value-range terms yet to be considered.

[13] In *Frege's Conception of Numbers as Objects*, Crispin Wright drew attention to Hume's principle, and conjectured that FA is consistent. A number of people (George Boolos, John Burgess, Allen Hazen, and Harold Hodes) proved the conjecture soon afterwards. A nice introduction to the story and philosophical significance of Frege's theorem can be found in Heck's paper 'Frege's Theorem: an Introduction.'

[14] If L^0 doubly well-orders H^0, it is an easy matter to show that, except for the L^0-least element of H^0, every other element has a unique immediate predecessor. *Mutatis mutandis* for the non L^0-greatest elements of H^0.

[15] The notation $[w : C(w)]$ is not primitive. It is to be understood in context, be it in the context of a proof, be it in the context of a formula. As an example of the latter context, $\#[w : C(w)] = x$ abbreviates $\exists H^0(\#H^0 = x \wedge \forall w(H^0 w \leftrightarrow C(w)))$.

[16] The terminology is Heck's, in the final pages of (1997).

[17] At this juncture, the reader might be worried about the absence of the definitions of the arithmetical operations of sum and product. As it is well-known, Dedekind in his epoch-making *Was sind und was sollen die Zahlen* showed that (in modern terms) the primitive recursive functions can be defined in second-order arithmetic, with zero and the successor operation as the sole non-logical primitives. However, Dedekind's proof readily adapts to predicative frameworks like HFR, in which quantifying over finite concepts is predicatively justified – see (Feferman and Hellman, 1995) and (Feferman and Hellman, 1998) for a modern treatment of this issue – and where it

is possible to prove that, for each natural number n, the collection of numbers not exceeding n forms a finite concept (a fact that can be proved in HFR by induction).

[18] The theory ACA is the extension of ACA_0 in which the induction axiom is replaced by the *scheme* $A(0) \wedge \forall x \forall y(A(x) \wedge xPy \rightarrow A(y)) \rightarrow \forall x A(x)$, for $A(x)$ *any* formula of the language of second-order arithmetic. Under the discussed interpretation, this scheme is also provable in HFR. In effect, if $A(x)$ is a formula of the first level fragment of our ramified language which is true of 0 but false of a natural number n, then the *second* level concept $[w : A(w) \wedge w \leq n]$ is true of 0 and false of n. This concept is a sub-concept of a finite concept (see the previous note). Hence, by finite reducibility (2), it is co-extensional with a zeroth level concept. This readily entails a counterexample to the inductiveness of $A(x)$.

[19] "Set theory in sheep's clothing" is the title of a section of W. O. Quine's *Philosophy of Logic*.

[20] Whereas it makes no difference in second-order logic whether one formulates an abstraction principle for concepts – as above – or for *formulas*, there is a difference in the predicative setting. The above argument does not apply to abstraction principles for formulas, i.e., to principles of the form $\%x.A(x) = \%x.B(x) \leftrightarrow Eq_x(A, B)$, where $A(x)$ and $B(x)$ are formulas and $Eq_x(,)$ is a formula *skeleton* yielding a schematic equivalence relation. (The argument fails to prove the left to right implication for lack of suitable comprehension.) However, even in this case, it is possible to see that the bad company objection does not apply. Here is why. Embed H into Heck's ramified predicative fragment by mapping concepts into zeroth level concepts. Define $\%x.A(x)$ by $\hat{z}.\exists G^1(Eq_x(G^1, A) \wedge z = \hat{x}.G^1 x)$. It can be seen that the equivalence $\%x.A(x) = \%x.B(x) \leftrightarrow Eq_x(A, B)$ holds in Heck's model of the ramified predicative fragment of Frege's *Grundgesetze*, as expounded in (Heck, 1996). Hence, no contradiction can arise from abstraction principles for formulas.

[21] These are Heck's words in (1997).

References

Boolos, G.: 1998a, 'Is Hume's Principle Analytic?'. In: *Logic, Logic, and Logic*. Harvard University Press, pp. 301–314. First published in 1997.

Boolos, G.: 1998b, 'The Standard of Equality of Numbers'. In: *Logic, Logic, and Logic*. Harvard University Press, pp. 202–219. First published in 1990.

Boolos, G. and R. Jeffrey: 1990, *Computability and Logic*. Cambridge University Press.

Brandom, R.: 1994, *Making It Explicit*. Harvard University Press.

Burgess, J. and A. Hazen: 1998, 'Predicative Logic and Formal Arithmetic'. *Notre Dame Journal of Formal Logic* **39**, 1–17.

Dedekind, R.: 1963, *The Nature and Meaning of Numbers*, Essays on the Theory of Numbers. Dover. First published in German in 1888 under the title *Was sind und was sollen die Zahlen*.

Dummett, M.: 1991, *Frege: Philosophy of Mathematics*. Harvard University Press.

Feferman, S.: 1998a, 'Infinity in Mathematics: Is Cantor Necessary? (Conclusion)'. In: *In the Light of Logic*. Oxford University Press, pp. 229–248.

Feferman, S.: 1998b, 'Weyl Vindicated: *Das Kontinuum* Seventy Years Later'. In: *In the Light of Logic*. Oxford University Press, pp. 249–283.

Feferman, S.: 1998c, 'Why a Little Bit Goes a Long Way: Logical Foundations of Scientifically Applicable Mathematics'. In: *In the Light of Logic.* Oxford University Press, pp. 284–298.

Feferman, S. and G. Hellman: 1995, 'Predicative Foundations of Arithmetic'. *Journal of Philosophical Logic* **24**(1), 1–17.

Feferman, S. and G. Hellman: 1998, 'Challenges to Predicative Foundations of Arithmetic'. In: G. Sher and R. Tieszen (eds.): *Between Logic and Intuition: Essays in honor of Charles Parsons.* Cambridge University Press.

Fernandes, A. M. and F. Ferreira: 2002, 'Groundwork for Weak Analysis'. *The Journal of Symbolic Logic* **67**, 557–578.

Ferreira, F. and K. Wehmeier: 2002, 'On the Consistency of the Δ^1_1-CA Fragment of Frege's *Grundgesetze*'. *Journal of Philosophical Logic* **31**, 301–311.

Frege, G.: 1967, *Basic Laws of Arithmetic.* University of California Press. Translation of §1-§52 of *Grundgesetze der Arithmetik* by M. Furth.

Heck, R.: 1996, 'The Consistency of Predicative Fragments of Frege's *Grundgesetze der Arithmetik*'. *History and Philosophy of Logic* **17**, 209–220.

Heck, R.: 1997, 'Finitude and Hume's Principle'. *Journal of Philosophical Logic* **26**(6), 589–617.

Heck, R.: 1999, 'Frege's Theorem: An Introduction'. *The Harvard Review of Philosophy* **7**, 56–73.

Parsons, C.: 1965, 'Frege's Theory of Number'. In: *Philosophy in America.* George Allen & Unwin, pp. 180–203. Reprinted with a post-script in *Frege's Philosophy of Mathematics*, Harvard University Press (1995).

Parsons, C.: 1987a, 'Developing Arithmetic in Set Theory without Infinity: Some Historical Remarks'. *History and Philosophy of Logic* **8**, 201–213.

Parsons, T.: 1987b, 'On the Consistency of the First-order Portion of Frege's Logical System'. *Notre Dame Journal of Formal Logic* **28**, 161–168.

Quine, W. O.: 1970, *Philosophy of Logic.* Prentice-Hall.

Simpson, S.: 1999, *Subsystems of Second-Order Arithmetic*, Perspectives in Mathematical Logic. Springer-Verlag.

Suppes, P.: 1957, *Introduction to Logic.* D. Van Nostrand Company.

Whitehead, A. N. and B. Russell: 1925, *Principia Mathematica*, Vol. 1. Cambridge University Press, 2nd edition.

Wright, C.: 1983, *Frege's Conception of Numbers as Objects.* Aberdeen University Press.