KAI F. WEHMEIER

CONSISTENT FRAGMENTS OF *GRUNDGESETZE* AND THE EXISTENCE OF NON-LOGICAL OBJECTS

ABSTRACT. In this paper, I consider two curious subsystems of Frege's *Grundgesetze der Arithmetik*: Richard Heck's predicative fragment H, consisting of *schema* V together with predicative second-order comprehension (in a language containing a syntactical abstraction operator), and a theory T_Δ in monadic second-order logic, consisting of *axiom* V and Δ^1_1-comprehension (in a language containing an abstraction function). I provide a consistency proof for the latter theory, thereby refuting a version of a conjecture by Heck. It is shown that both H and T_Δ prove the existence of infinitely many non-logical objects (T_Δ deriving, moreover, the nonexistence of the value-range concept). Some implications concerning the interpretation of Frege's proof of referentiality and the possibility of classifying any of these subsystems as logicist are discussed. Finally, I explore the relation of T_Δ to Cantor's theorem which is somewhat surprising.

1. INTRODUCTION

The aim of this paper is twofold: First, to prove the consistency of Δ^1_1-comprehension with Frege's *Grundgesetz* V, thereby refuting (a version of) a conjecture by Richard Heck in (Heck, 1996). The concern for proving subsystems of Frege's *Grundgesetze der Arithmetik* consistent originated with Terence Parsons' (1987) proof of Peter Schroeder-Heister's conjecture (1987, 78) that schema V is consistent with first-order logic. Much along the lines of Parsons' proof, Heck (1996) shows that comprehension restricted to formulas without second-order quantifiers (predicative formulas) is consistent with schema V. Both Parsons and Heck treat the abstractor as a syntactical operator building terms from formulas, which explains their use of a schematic version of *Grundgesetz* V. In our system T_Δ, the abstractor is a proper function symbol and we consider the single (quantified) axiom V. This appears to be quite natural in two respects: First, it is not entirely clear how the notion of, say, Σ^1_1-formula should be defined in a language where singular terms can also embody formula complexity. Second, such an account is perhaps closer in spirit to Frege's conception of the abstractor as a second-level *function*.

The second aim of this essay is to discuss some curious features exhibited

Synthese **121:** 309–328, 1999.

by Heck's predicative system H and our theory T_Δ: Both systems prove (in different ways) the existence of non-value-ranges, i.e. the sentence $\exists x \forall F (x \neq \hat{z}.Fz)$. Moreover, T_Δ refutes the existence of the concept of being a value-range:

$$\mathsf{T}_\Delta \vdash \neg\exists H \forall x (Hx \leftrightarrow \exists F (x = \hat{z}.Fz)).$$

Frege himself seems to have identified the notions of logical object and value-range, respectively: witness his letter to Russell of 7-28-1902 (Frege, 1976, item XXXVI/7, p. 223; the translation is mine):

> Es handelt sich dabei um die Frage: wie fassen wir logische Gegenstände? und ich habe keine andere Antwort darauf gefunden, als die: wir fassen sie als Umfänge von Begriffen, oder allgemeiner als Werthverläufe von Functionen.

> The question is this: how do we apprehend logical objects? and I have found no answer other than this: we apprehend them as extensions of concepts, or more generally as value-ranges of functions.

Under this identification, one may put our results as follows: The consistent subtheories H and T_Δ of *Grundgesetze* both prove the existence of infinitely many non-logical objects, and the latter theory even refutes the existence of the concept of a logical object.[1]

These phenomena are, I believe, remarkable in at least three respects: First, the results seem – at least at first sight – to be at odds with one aspect of Heck's diligent work (Heck, 1997a) concerning sections 29–32 of *Grundgesetze*, viz. his claim that nothing was wrong in Frege's restricting the domain of his theory to value-ranges only. Second, one might argue that proving the existence of non-logical objects constitutes a self-refutation of *Grundgesetze der Arithmetik* and these consistent subtheories as *logicist* systems even below the level of outright inconsistency. And third, we have the following curious situation concerning Cantor's theorem (saying that there is no bijection between the first- and the second-order entities): Full axiomatic second-order logic implies Cantor's theorem in the following sense: If the language of second-order logic is augmented by a single unary function symbol, to be attached to unary predicate variables and yielding terms of individual type, and if all formulas of this extended language are eligible as comprehension formulas, then it can be proved that the corresponding function cannot be injective, which may be paraphrased by saying that there are too few individuals. The same system with just Δ^1_1-comprehension is consistent with such a function being injective, but still proves Cantor's theorem by showing that no such injection can be *onto* the individuals: There are either too few (no injection) or too many (no surjection) individuals![2]

The paper is organized as follows: In Section 2, I discuss Heck's predicative fragment H and show how it derives the existence of infinitely many non-logical objects. Section 3 provides the consistency proof for Δ^1_1-comprehension with axiom V. The final section 4 is devoted to some corollaries and the discussion of our results.

2. HECK'S PREDICATIVE FRAGMENT

The theory H is formulated in a second-order language (including equality) containing a term-forming (and variable-binding) operator $\hat{x}$. The terms and formulas are defined simultaneously, the crucial clause being: If $A(x)$ is a formula, then $\hat{x}A(x)$ is a term. The logical axioms and rules of H are those of axiomatic second-order logic with the comprehension schema restricted to instances whose comprehension formulas contain no second-order quantifiers; the non-logical (?) axioms are the instances of schema V:

$$\hat{x}A(x) = \hat{x}B(x) \leftrightarrow \forall x(A(x) \leftrightarrow B(x))$$

for any formulas $A(x)$ and $B(x)$. For a fuller exposition of the system and a proof of its consistency, see Heck (1996).

The following result is already somewhat surprising:

THEOREM 2.1. In Heck's predicative fragment H, it is provable that there are non-value-ranges. That is, the sentence

$$\exists x \forall F(x \neq \hat{y}Fy)$$

is derivable in H. In fact, there is an H-proof of this sentence which makes no use of the predicative comprehension schema.

Proof. Argue informally within H. Let r abbreviate the term $\hat{x}(\exists\, G(x = \hat{y}Gy \wedge \neg Gx))$. Suppose $\mathsf{r} = \hat{y}Fy$. By the relevant instance of schema V,

$$\forall x(Fx \leftrightarrow \exists\, G(x = \hat{y}Gy \wedge \neg Gx)).$$

If $F\mathsf{r}$, then for some G, $\mathsf{r} = \hat{y}Gy \wedge \neg G\mathsf{r}$. But if $\hat{y}Fy = \mathsf{r} = \hat{y}Gy$, then (by V) $\forall x(Fx \leftrightarrow Gx)$, in particular $G\mathsf{r}$, contradicting $\neg G\mathsf{r}$. We conclude that $\neg F\mathsf{r}$. This implies $\forall\, G(\mathsf{r} = \hat{y}Gy \rightarrow G\mathsf{r})$; in particular, by $\mathsf{r} = \hat{y}Fy$, $F\mathsf{r}$, contradiction. (Note that this proof uses *no* instance of comprehension.) This shows that H proves $\forall F \neg(\mathsf{r} = \hat{y}Fy)$, and by existential generalization we obtain the desired conclusion $\exists x \forall F(x \neq \hat{y}Fy)$. □

But much more is true: H proves the existence of infinitely many non-value-ranges. This can be seen as follows: For each $n < \omega$, let $R_n(x)$ be the formula

$$\exists G_0\Big(x = \hat{y}(G_0(y)) \wedge \forall G_1, \ldots, G_n$$

$$\Big[\bigwedge_{i<n} G_{i+1}(\hat{z}(G_i(z))) \to \neg G_0(\hat{z}(G_n(z)))\Big]\Big).$$

Just as r was modelled on the Russell class $\{x : x \notin x\}$, the value range of R_n corresponds to the class $\{x : \neg\exists z_1, \ldots, z_n(x \in z_1 \wedge z_1 \in z_2 \wedge \ldots \wedge z_n \in x\}$.

LEMMA 2.2. For each $n < \omega$, $\mathsf{H} \vdash \neg\exists F(\hat{x}R_n(x) = \hat{x}F(x))$. In fact, these sentences can be proved in H without using any comprehension axioms.

Proof. Assume by way of contradiction that $\mathsf{r}_n := \hat{z}R_n(z) = \hat{z}F(z)$ for some concept F, so that $\forall x(R_n(x) \leftrightarrow Fx)$ by V. Now suppose that $R_n(\mathsf{r}_n)$, i.e. for some G_0, $\mathsf{r}_n = \hat{z}(G_0 z)$ and

$$\forall G_1, \ldots, G_n \left(\bigwedge_{i<n} G_{i+1}(\hat{z}(G_i z)) \to \neg G_0(\hat{z}(G_n z))\right).$$

Since $\mathsf{r}_n = \hat{z}(G_0 z) = \hat{z}F(z)$, by V, R_n, G_0 and F are coextensive. Instantiating every G_{i+1} by F, we obtain $\neg R_n(\mathsf{r}_n)$, since $F(\hat{z}Fz)$ was supposed. Cancelling this assumption, we conclude that $\neg R_n(\mathsf{r}_n)$. Hence, given that for F as G_0 we have $\mathsf{r}_n = \hat{z}(G_0 z)$ (and hence $\forall x(R_n(x) \leftrightarrow G_0(x))$ by V), there are $G_1, \ldots, G_n$ such that

(*) $$G_1(\mathsf{r}_n) \wedge G_2(\hat{z}G_1 z) \wedge \cdots \wedge G_n(\hat{z}G_{n-1}z) \wedge R_n(\hat{z}G_n z).$$

By the last conjunct, for some H_0 with $\hat{z}(G_n z) = \hat{z}(H_0 z)$ (so that G_n and H_0 coincide by V),

(**) $$\forall H_1, \ldots, H_n \left(\bigwedge_{i<n} H_{i+1}(\hat{z}H_i z) \to \neg G_n(\hat{z}H_n z)\right).$$

Instantiating, for $i < n$, each H_{i+1} by G_i, we obtain $\neg G_n(\hat{z}G_{n-1}z)$ from (*) and (**), contradicting the second-to-last conjunct in (*). □

LEMMA 2.3. For each $k > 0$, H proves $\exists x(R_n(x) \wedge \neg R_{n+k}(x))$, and hence $\hat{z}R_n(z) \neq \hat{z}R_{n+k}(z)$.

Proof. Argue within H. By predicative comprehension, there are concepts $H_0, \ldots, H_{n+k}$ with the following properties: $H_1(x) \leftrightarrow x = x$, $H_{i+1}(x) \leftrightarrow x = \hat{z}H_i(z)$ for $1 \leq i < n+k$, $H_0(x) \leftrightarrow x = \hat{z}H_{n+k}(z)$. Put $\mathrm{t} = \hat{z}H_0(z)$. Obviously, $\neg R_{n+k}(\mathrm{t})$, since $\bigwedge_{i<n+k} H_{i+1}(\hat{z}H_i(z)) \wedge H_0(\hat{z}H_n(z))$. To see that $R_n(\mathrm{t})$, we show that $H_0, \ldots, H_{n+k}$ are all distinct. By induction on $i>1$ up to $n+k$ one sees that $H_1, \ldots, H_{n+k}$ are distinct: H_2 and H_1 are not the same because the latter is true of more than one element. Suppose that $H_1, \ldots, H_i$ are distinct, and assume further that H_{i+1} is among them. Clearly, as in the induction basis, H_{i+1} is not coextensive with H_1. If H_{i+1} and H_j coincide, $1 < j \leq i$, then by definition of these concepts, $\hat{z}H_i(z) = \hat{z}H_{j-1}(z)$, and so H_i and H_{j-1} are coextensive, contradicting the induction hypothesis, and the induction is completed. Finally, H_0 is clearly not H_1, and if it were coextensive with H_j, $1 < j \leq n+k$, then by definition of H_0 and H_j, H_{n+k} and H_{j-1} would coincide, which is impossible, as we just saw.

It remains to show that $R_n(\mathrm{t})$, i.e.

$$\exists G_0\Bigg(\mathrm{t} = \hat{y}(G_0(y)) \wedge \forall G_1, \ldots, G_n$$

$$\left[\bigwedge_{i<n} G_{i+1}(\hat{z}(G_i(z))) \rightarrow \neg G_0(\hat{z}(G_n(z)))\right]\Bigg).$$

G_0 can be instantiated by H_0. Now let $G_1, \ldots, G_n$ be given such that $G_1(\mathrm{t}) \wedge \bigwedge_{1\leq i<n} G_{i+1}(\hat{z}G_i(z))$. Assume by way of contradiction that $H_0(\hat{z}G_n(z))$. By definition of H_0, $\hat{z}G_n(z) = \hat{z}H_{n+k}(z)$, so G_n and H_{n+k} coincide; hence $\hat{z}G_{n-1}(z) = \hat{z}H_{n+k-1}(z)$ and G_{n-1} and H_{n+k-1} coincide, ..., hence $\hat{z}(G_1(z)) = \hat{z}H_{1+k}(z)$ so G_1 and H_{1+k} coincide, hence $\hat{z}(H_0(z)) = \hat{z}H_k(z)$ and H_0 coincides with H_k, which is impossible for $k>0$, as proved above. □

COROLLARY 2.4. For each $n > 0$, H proves that there are n objects $x_1, \ldots, x_n$ which are not identical with $\hat{z}Fz$ for any concept F. □

These results would of course come as no surprise if H were inconsistent – but it isn't. Still, some readers may suspect that we have hit upon an 'accidental' feature of H, due to the presence of many names for classes we know very little about.[3] Be that as it may, Corollary 2.4 does not seem to be accidental, since it also holds good for the theory to be proved consistent in the next section.

3. AXIOM V AND Δ^1_1-COMPREHENSION.

Let L be the language of monadic second-order logic (including equality) given by the following non-logical vocabulary:

- a unary function symbol ϵ, to be attached to constants or variables of second-order type, yielding terms of individual type;
- an individual constant n for each $n \in \omega$;
- a predicate constant H for each finite or cofinite subset $H \subseteq \omega$ (where $H \subseteq \omega$ is cofinite if and only if $\omega \setminus H$ is finite).

The constant symbols only serve to render us independent of variable assignments in the model to be presented below; if the theory under consideration is consistent when formulated in L, it will be so when formulated in L with the constant symbols deleted, too.

Our base theory is the system T of axiomatic monadic second-order logic in the language L without the comprehension schema, but augmented by the single non-logical (?) axiom V:

$$\forall F \forall G(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx)).$$

An L-formula is called Σ^1_1 (Π^1_1) if it is of the form $\exists F\varphi$ ($\forall F\varphi$), where φ contains no second-order quantifier (but may contain free first- and second-order variables). Σ^1_1-comprehension or Σ^1_1-CA (Π^1_1-comprehension or Π^1_1-CA) is the schema $\exists H\forall x(Hx \leftrightarrow \varphi(x))$, where $\varphi(x)$ is any Σ^1_1-formula of L (Π^1_1-formula of L) not containing H free.

Letting $\mathsf{T}_\Sigma := \mathsf{T} + \Sigma^1_1 - \text{CA}$ and $\mathsf{T}_\Pi := \mathsf{T} + \Pi^1_1 - \text{CA}$, we see that both T_Σ and T_Π are inconsistent:

By appeal to the instances $\exists H\forall x(Hx \leftrightarrow \exists F(x = \epsilon F \wedge \neg Fx))$ or $\exists H\forall x(Hx \leftrightarrow \forall G(x = \epsilon G \rightarrow \neg Gx))$ of $\Sigma^1_1 - \text{CA}$ and $\Pi^1_1 - \text{CA}$ respectively, one can derive Russell's paradox in the usual way. We illustrate this for T_Σ, arguing informally.

Suppose $\forall x(Hx \leftrightarrow \exists F(x = \epsilon F \wedge \neg Fx))$. First assume $H(\epsilon H)$. By hypothesis, $\exists F(\epsilon H = \epsilon F \wedge \neg F(\epsilon H))$. For any such F, by V, $F(\epsilon H) \leftrightarrow H(\epsilon H)$ and so, by assumption, $F(\epsilon H)$ and $\neg F(\epsilon H)$, contradiction. So $\neg H(\epsilon H)$. Hence, by hypothesis, $\forall F(\epsilon H = \epsilon F \rightarrow F(\epsilon H))$, in particular, $H(\epsilon H)$, contradiction. Thus T_Σ (in fact, T alone) proves $\neg\exists H\forall x(Hx \leftrightarrow \exists F(x = \epsilon F \wedge \neg Fx))$, contradicting $\Sigma^1_1 - \text{CA}$.

This leads us to consider the 'intersection' T_Δ of T_Σ and T_Π:

Δ^1_1-comprehension or Δ^1_1-CA is the schema

$$\forall x[\varphi(x) \leftrightarrow \psi(x)] \rightarrow \exists H\forall x(Hx \leftrightarrow \varphi(x)),$$

where $\varphi(x)$ is a Σ^1_1-L-formula, $\psi(x)$ is a Π^1_1-L-formula, and H is not free in $\varphi(x)$. T_Δ is $\mathsf{T} + \Delta^1_1 - \mathrm{CA}$.

Why use the abstractor as a function symbol proper, and not, in the Parsons-Heck tradition, as a term-building operator? The reason is that there does not seem to be a completely obvious definition of formula complexity in the presence of singular terms built from compound formulas. Consider e.g. the formula

$$\hat{z}[\exists F(x = \hat{y}(Fy) \wedge \neg Fx)] = \hat{z}(\forall y(y = y)).$$

This looks like an atomic formula (if one ignores the internal structure of the value-range terms), but it is, under schema V, equivalent to

$$\exists F(x = \hat{y}(Fy) \wedge \neg Fx)$$

and so clearly cannot be allowed as a comprehension formula. Note, however, that the difference in setting up the abstractor makes it difficult to compare Heck's H and T_Δ in strength: It follows from Corollaries 3.4 and 3.5 that every Δ^1_1-set in the model described below is finite or cofinite, while there are first-order formulas of H which define infinite coinfinite sets in any model, e.g. $\exists z(x = \hat{y}(y = z))$.

We now specify an L-structure $\mathfrak{A}$ which will turn out to be a model for T_Δ.

- The first-order domain of $\mathfrak{A}$ is the set ω of natural numbers.
- The constant n is interpreted by the number $n \in \omega$.
- The second-order domain of $\mathfrak{A}$ is the collection $\mathfrak{N}$ of all finite and cofinite subsets of ω.
- The predicate constant H is interpreted by the set $H \subseteq \omega$.
- The function symbol ϵ is interpreted by the function $\varepsilon : \mathfrak{N} \to \omega$ such that, for any $x_1, \ldots, x_r \in \omega$ with $x_1 < \cdots < x_r$, $\varepsilon(\{x_1, \ldots, x_r\}) = \langle 0, x_1, \ldots x_r\rangle$ and $\varepsilon(\omega - \{x_1, \ldots, x_r\}) = \langle 1, x_1, \ldots, x_r\rangle$, where $\langle, \ldots, \rangle$ is the standard coding function by prime numbers.

Concerning the function ε, we have that $x_{i_j} < \langle x_{i_1}, \ldots, x_{i_r}\rangle$ for each $j \in \{1, \ldots, r\}$, and so for each $H \in \mathfrak{N}$, $\mathfrak{A}$ satisfies $\mathsf{H}(\epsilon\mathsf{H})$ if and only if H is cofinite. Note that the range of this function, $\mathsf{ran}(\varepsilon)$, is neither finite nor cofinite. Given any $x = \langle x_0, \ldots, x_{r-1}\rangle$, we let $(x)_0 = x_0$.

Clearly $\mathfrak{A}$ satisfies axiom V. It thus remains to show that $\mathfrak{A}$ is a model of $\Delta^1_1 - \mathrm{CA}$. Since every individual and every second-order entity of $\mathfrak{A}$ has a name in L, we may assume henceforth that the instances of $\Delta^1_1 - \mathrm{CA}$ actually contain no free variables. By the following observation, we may also assume that the instances of $\Delta^1_1 - \mathrm{CA}$ contain no predicate constants:

In a given instance, every subformula $\mathsf{H}t$ may be replaced by $\bigvee_{i<r} t = \mathsf{k}_i$ or by $\bigwedge_{i<r} t \neq \mathsf{k}_i$, accordings as H is $\{k_1, \ldots, k_r\}$ or $\omega - \{k_1, \ldots, k_r\}$, and every occurrence of a term $\epsilon\mathsf{H}$ may be replaced by an occurrence of n, where n is the value of ε at H.

We say that an L-formula $\varphi(x)$ in the sole free variable x defines the set $\{k : \mathfrak{A} \models \varphi(\mathsf{k})\}$. What we must show is that every set which has a Σ^1_1- as well as a Π^1_1-definition is either finite or cofinite. Since the Π^1_1-sets are precisely the complements of the Σ^1_1-sets, we turn first to the Σ^1_1-sets of $\mathfrak{A}$. To this end, we establish a normal form theorem for the matrices of Σ^1_1-formulas.

LEMMA 3.1. Any L-formula containing no second-order quantifiers or constants and containing at most one free predicate variable F (but possibly any finite number of free individual variables and individual constants) is equivalent in $\mathfrak{A}$ to a Boolean combination of atomic formulas and formulas of the forms $\exists^{=n} yFy$, $\exists^{=n} y\neg Fy$, in at most the same free variables.

Proof. If $\mathfrak{A}$ models $\varphi(x) \leftrightarrow \psi(x)$, then it also satisfies $\varphi(\epsilon F) \leftrightarrow \psi(\epsilon F)$, so we may assume that ϵF does not occur in the formula, and similarly for individual constants.

The proof proceeds by induction on the inductive generation of L-formulas. Clearly the only interesting case is that of first-order universal quantification. Consider $\forall x A(x, \bar{z}, F)$, where $A(x, \bar{z}, F)$ contains no second-order quantifier or constant, contains as free second-order variable at most F and has all its free individual variables among $x, \bar{z}$. By the induction hypothesis and conjunctive normal form, we may write $\forall x A(x, \bar{z}, F)$ as $\forall x \bigwedge_{a<b} \bigvee_{c<d_a} P_{ac}(x, \bar{z}, F)$, where each $P_{ac}(x, \bar{z}, F)$ is atomic, negated atomic or of one of the forms $\exists^{=n} yFy$, $\neg\exists^{=n} yFy$, $\exists^{=n} y\neg Fy$, $\neg\exists^{=n} y\neg Fy$. The universal quantifier distributes over conjunctions, so that our formula is equivalent to $\bigwedge_{a<b} \forall x \bigvee_{c<d_a} P_{ac}(x, \bar{z}, F)$. Further, every disjunct $P_{ac}(x, \bar{z}, F)$ not containing x free may be pulled outside the universal quantifier; this concerns, in particular, every P_{ac} of one of the forms $(\neg)\exists^{=n} y(\neg)Fy$. It now suffices to consider formulas of the form $\forall x \bigvee_{c<d} P_c(x, \bar{z}, F)$, where each $P_c(x, \bar{z}, F)$ is atomic or negated atomic and actually contains x free. Such formulas may be written as

$$\forall x \left[\bigvee_{\alpha=1,\ldots,\beta} (x = z_{i_\alpha}) \vee \bigvee_{\gamma=1,\ldots,\eta} (x \neq z_{j_\gamma}) \vee (Fx)^{\delta_1} \vee (\neg Fx)^{\delta_2} \right],$$

where $\beta, \eta \geq 0$, $\delta_i \in \{0, 1\}$ and ψ^0 is $\perp$, ψ^1 is ψ. Clearly we may assume that the δ_i are not both 1, otherwise our formula is equivalent to $\top$. If $\eta > 0$, we may rewrite the formula as

$$\forall x \Bigg[x = z_{j_1} \to \bigvee_{\alpha=1,\ldots,\beta} (x = z_{i_\alpha}) \vee \bigvee_{\gamma=2,\ldots,\eta} (x \neq z_{j_\gamma}) \vee (Fx)^{\delta_1} \vee (\neg Fx)^{\delta_2} \Bigg],$$

which is equivalent to

$$\bigvee_{\alpha=1,\ldots,\beta} (z_{j_1} = z_{i_\alpha}) \vee \bigvee_{\gamma=2,\ldots,\eta} (z_{j_1} \neq z_{j_\gamma}) \vee (Fz_{j_1})^{\delta_1} \vee (\neg Fz_{j_1})^{\delta_2},$$

and we are done. We thus assume $\eta = 0$. If both δ_i are 0, we are also done since $\forall x \bigvee_{\alpha=1,\ldots,\beta}(x = z_{i_\alpha})$ is false in $\mathfrak{A}$ (which is infinite). If δ_1 is 1, the formula is equivalent to $\forall x(\neg Fx \to \bigvee_{\alpha=1,\ldots,\beta}[x = z_{i_\alpha}])$ (or, more suggestively, $\omega - F \subseteq \{z_{i_1}, \ldots, z_{i_\beta}\}$) which is equivalent to

$$\bigvee_{A \subseteq \{i_1,\ldots,i_\beta\}} \Bigg[\exists^{=|A|} y \neg Fy \wedge \bigwedge_{j \in A} \neg Fz_j \wedge \bigwedge_{i,j \in A, i \neq j} z_i \neq z_j \Bigg],$$

and we are done. The same argument applies, *mutatis mutandis*, if $\delta_2 = 1$ (interchange $\neg F$ and F). □

COROLLARY 3.2. Every L-formula containing no second-order variables is equivalent, in $\mathfrak{A}$, to a quantifier-free one; in particular, each such formula in one free individual variable defines in $\mathfrak{A}$ a finite or cofinite set.

Proof. The first part follows immediately from the lemma. The second part is a consequence of the closure of $\mathfrak{N}$ under finitary Boolean operations. □

These preliminaries out of the way, we can now turn to the characterization of the Σ^1_1-sets of $\mathfrak{A}$.

THEOREM 3.3. Every Σ^1_1-set of $\mathfrak{A}$ is of the form $\bigcup_{a<b}[A_a \cap B_a]$, where $b \in \omega$, each A_a is finite or cofinite, and each B_a is either finite, cofinite or a subset of $\mathsf{ran}(\varepsilon)$, the range of ε.

Before taking on the proof of Theorem 3.3, we shall record some corollaries.

COROLLARY 3.4. Every Σ^1_1-set of $\mathfrak{A}$ is cofinite or contains only finitely many elements not in the range of ε.

Proof. By induction on $b \in \omega$ we show that every set $\bigcup_{a<b}[A_a \cap B_a]$ with each $A_a \in \mathfrak{N}$ and each B_a either in $\mathfrak{N}$ or a subset of $\mathsf{ran}(\varepsilon)$ has the desired property. If $b = 0$, we obtain the finite set $\emptyset$, containing no k in $\omega \setminus \mathsf{ran}(\varepsilon)$. $\bigcup_{a=1,\dots,b+1}[A_a \cap B_a] = (\bigcup_{a=1,\dots,b}[A_a \cap B_a]) \cup [A_{b+1} \cap B_{b+1}]$. By induction hypothesis, $\bigcup_{a=1,\dots,b}[A_a \cap B_a]$ is cofinite or contains only finitely many k not in the range of ε. If A_{b+1} or B_{b+1} is finite, so is their intersection. If B_{b+1} contains only elements k from $\mathsf{ran}(\varepsilon)$, so does $A_{b+1} \cap B_{b+1}$. In all these cases we are done. If both A_{b+1} and B_{b+1} are cofinite, so is their intersection, and the claim follows. □

COROLLARY 3.5. Every Π^1_1-set of $\mathfrak{A}$ is either finite or contains all but finitely many k with $k \notin \mathsf{ran}(\varepsilon)$.

Proof. The Π^1_1-sets are precisely the complements of the Σ^1_1-sets. □

COROLLARY 3.6. $\mathfrak{A}$ is a model of Δ^1_1-CA.

Proof. By Corollaries 3.4 and 3.5, every set which is Σ^1_1 as well as Π^1_1 must be finite or cofinite. □

COROLLARY 3.7. T_Δ is consistent.

Proof. $\mathfrak{A}$ is a model of T_Δ. □

Proof of Theorem 3.3. As observed earlier, we may assume that the Σ^1_1-formulas have only one free individual variable, no free second-order variable and no predicate constant.

By Lemma 3.1 and disjunctive normal form, every Σ^1_1-formula may be written as $\exists F \bigvee_{a=1,\dots,b} \bigwedge_{c=1,\dots,d_a} P_{ac}(x, F)$, where each $P_{ac}(x, F)$ is atomic, negated atomic or of one of the forms $(\neg)\exists^{=n} y(\neg)Fy$. The existential quantifier distributing over disjunctions, we may rewrite this formula as

$$\bigvee_{a=1,\dots,b} \exists F \bigwedge_{c=1,\dots,d_a} P_{ac}(x, F).$$

Thus each Σ^1_1-set is a finite union of sets defined by formulas of the form $\exists F \bigwedge_{c=1,\dots,d} P_c(x, F)$, each P_c atomic, negated atomic or of type $(\neg)\exists^{=n} y(\neg)Fy$. Every $P_c(x, F)$ in which F does not occur may be pulled across the existential quantifier, resulting in a formula

$$\bigwedge_{e=1,\dots,f} Q_e(x) \wedge \exists F \bigwedge_{g=1,\dots,h} R_g(x, F),$$

where each $R_g(x, F)$ contains F, is atomic or negated atomic or of type $(\neg)\exists^{=n}y(\neg)Fy$, and each $Q_e(x)$ is atomic or negated atomic. By Lemma 3.1, $\bigwedge_{e=1,\ldots,f} Q_e(x)$ defines a finite or cofinite set. Thus every Σ^1_1-set is a finite union of binary intersections of (co-)finite sets with sets defined by formulas $\exists F \bigwedge_{g=1,\ldots,h} R_g(x, F)$ as above; it remains to see that each of these latter sets is either finite, cofinite or a subset of $\mathsf{ran}(\varepsilon)$.

Each $R_g(x, F)$ has one of the following forms: $(\neg)x = \epsilon F$, $(\neg)\epsilon F = \epsilon F$, $(\neg)\mathsf{k} = \epsilon F$, $(\neg)Fx$, $(\neg)F\mathsf{k}$, $(\neg)F(\epsilon F)$, $(\neg)\exists^{=n}yFy$, $(\neg)\exists^{=n}y\neg Fy$. Without loss of generality, we may assume the following:

(i) $\epsilon F = \epsilon F$ does not occur among the R_g.

(ii) $\neg\epsilon F = \epsilon F$ does not occur among the R_g (otherwise we obtain the finite set $\emptyset$).

(iii) If R_g and R_j are of the form $\exists^{=n}yFy$, then $g = j$ (conjuncts ascribing different cardinalities to F lead to $\emptyset$, others may be contracted).

(iv) If R_g and R_j are of the form $\exists^{=n}y\neg Fy$, then $g = j$.

(v) If R_g and R_j are both $F\mathsf{k}$, then $g = j$.

(vi) If R_g and R_j are both $\neg F\mathsf{k}$, then $g = j$.

(vii) No R_g has the form $\mathsf{k} = \epsilon F$:
Otherwise we may write $\exists F \bigwedge_{g=1,\ldots,h} R_g(x, F)$ as $\exists F(\mathsf{k} = \epsilon F \wedge \bigwedge R_g(x, F))$ which defines $\emptyset$ if $k \notin \mathsf{ran}(\varepsilon)$ and is equivalent to $\bigwedge R_g(x, \mathsf{H})$ if $\varepsilon(H) = k$. By the elimination procedure for predicate constants described above, this is equivalent to a first-order-formula, thus defining a finite or cofinite set by Lemma 3.1.

(viii) No R_g is the negation of some R_j (otherwise we obtain $\emptyset$).

(ix) If some R_g has the form $\exists^{=n}yFy$, then no R_j is of the form $\exists^{=m}y\neg Fy$ and vice versa (otherwise we obtain $\emptyset$ since no set is both finite and cofinite).

(x) If $F(\epsilon F)$ occurs among the R_g, then no formula of the form $\exists^{=n}yFy$ does (since $\mathsf{H}(\epsilon\mathsf{H})$ holds in $\mathfrak{A}$ iff H is cofinite, and so we would obtain $\emptyset$).

(xi) If $\neg F(\epsilon F)$ occurs among the R_g, then no formula of the form $\exists^{=n}y\neg Fy$ does.

(xii) If some formula $\exists^{=n}yFy$ occurs among the R_g, then $n>0$ (otherwise our formula is equivalent to a first-order one as above).

(xiii) If some formula $\exists^{=n}y\neg Fy$ occurs among the R_g, then $n>0$.

(xiv) $x = \epsilon F$ is not among the R_g (otherwise the set defined will be a subset of $\mathsf{ran}(\varepsilon)$ anyway).

(xv) If some formula $\exists^{=n}yFy$ is among the R_g, then strictly fewer than n formulas of the form $F\mathsf{k}$ are among the R_g (if there are strictly more than n, by (v) the formula defines $\emptyset$; if there are precisely n, it is equivalent to a first-order formula).

(xvi) If some $\exists^{=n} y \neg Fy$ is among the R_g, then strictly fewer than n formulas of the form $\neg F\mathsf{k}$ are among the R_g.

Those forms of $\exists F \bigwedge_{g=1,\ldots,h} R_g(x, F)$ that remain to be considered may be summarized by the following schema, where it is understood that conditions (i)-(xvi) are in force, in particular that they govern the possible combinations of values for the $\delta_i \in \{0, 1\}$ and $\eta, \xi, P, N, L \geq 0$:

$$\exists F((x \neq \epsilon F)^{\delta_0} \wedge (Fx)^{\delta_1} \wedge (\neg Fx)^{\delta_2} \wedge (F(\epsilon F))^{\delta_3} \wedge (\neg F(\epsilon F))^{\delta_4} \wedge (\exists^{=v} yFy)^{\delta_5} \wedge (\exists^{=s} y \neg Fy)^{\delta_6} \wedge \bigwedge_{\gamma=1,\ldots,\eta} \neg \exists^{=r_\gamma} yFy \wedge \bigwedge_{\xi=1,\ldots,\zeta} \neg \exists^{=t_\xi} y \neg Fy \wedge \bigwedge_{l=1,\ldots,P} F\mathsf{n}_l \wedge \bigwedge_{k=1,\ldots,N} \neg F\mathsf{m}_k \wedge \bigwedge_{i=1,\ldots,L} \mathsf{k}_i \neq \epsilon F).$$

We shall show that the instances of this schema all define cofinite sets. To this end, it suffices to consider instances with a maximal number of conjuncts, since the sets defined by instances with fewer conjuncts will only be larger, and supersets of cofinite sets are cofinite. We thus always require δ_0 to be 1.

For ease of notation, we often write $\{\bar{m}\}$ instead of $\{m_1, \ldots, m_N\}$, $\{\bar{m}, x\}$ instead of $\{m_1, \ldots, m_N, x\}$ and $\{\bar{m}, \bar{n}\}$ for $\{m_1, \ldots, m_N, n_1, \ldots, n_P\}$, etc.

CASE I. $\delta_1 = 1$, and hence $\delta_2 = 0$.

First suppose $\delta_3 = 1$. By (i)–(xvi) and the facts that $F(\epsilon F)$ implies any $\neg \exists^{=r_\gamma} yFy$ and that $\exists^{=s} y \neg Fy$ implies any $\neg \exists^{=t_\xi} y \neg Fy$ we must consider the formula

$$\exists F\Bigg(x \neq \epsilon F \wedge Fx \wedge F(\epsilon F) \wedge \exists^{=s} y \neg Fy \wedge \bigwedge_{l=1,\ldots,P} F\mathsf{n}_l \wedge \bigwedge_{k=1,\ldots,N} \neg F\mathsf{m}_k \wedge \bigwedge_{i=1,\ldots,L} \mathsf{k}_i \neq \epsilon F\Bigg).$$

Now take $m_1, \ldots, m_N$. By (xvi), $N < s$; there are infinitely many sets $\{\bar{a}\}$ of cardinality $s - N$ disjoint from $\{n_1, \ldots, n_P\}$. Take one of them such that $\varepsilon(\omega - \{\bar{m}, \bar{a}\}) \notin \{k_1, \ldots, k_L\}$. Then clearly $\omega - \{\bar{m}, \bar{a}, \varepsilon(\omega - \{\bar{m}, \bar{a}\})\}$ is a cofinite subset of the set defined by our formula which is therefore cofinite itself.

Now suppose that $\delta_3 = 0$ and $\delta_4 = 1$. By (i)–(xvi) and the facts that $\neg F(\epsilon F) \wedge Fx$ implies $x \neq \epsilon F$, that $\neg F(\epsilon F)$ implies any $\neg \exists^{=t_\xi} y \neg Fy$ and that $\exists^{=v} yFy$ implies any $\neg \exists^{=r_\gamma} yFy$, we must consider the formula

$$\exists F\Bigg(Fx \wedge \neg F(\epsilon F) \wedge \exists^{=v} yFy \wedge$$

$$\bigwedge_{l=1,\dots,P} F\mathsf{n}_l \wedge \bigwedge_{k=1,\dots,N} \neg F\mathsf{m}_k \wedge \bigwedge_{i=1,\dots,L} \mathsf{k}_i \neq \epsilon F\Bigg),$$

where we assume that $P = v-1$. Let H be the set defined by this formula. We claim that $H = \omega - (\{\bar{m}\} \cup \{r : r \notin \{\bar{n}\} \wedge \varepsilon(\{\bar{n}, r\}) \in \bar{k}\})$. For the left-to right inclusion, let $r \in H$. Clearly r is no m_k. Suppose $r \notin \{\bar{n}\}$. Since $r \in H$, r is in some v-element set containing $\{\bar{n}\}$ whose value under ε is not in $\{\bar{k}\}$; but $\{\bar{n}, r\}$ is the only possibility, so $\varepsilon(\{\bar{n}, r\}) \notin \{\bar{k}\}$. For the other direction, suppose $r \notin \{\bar{m}\}$ and first assume $r \notin \{\bar{n}\}$, $\varepsilon(\{\bar{n}, r\}) \notin \{\bar{k}\}$. Since $P = v-1$ and the n_l are pairwise distinct, r is in some v-element set disjoint from $\{\bar{m}\}$ but containing $\{\bar{n}\}$ whose value under ε is no k_i, so $r \in H$. But any n_l is also in H since there are infinitely many $k \notin \{\bar{m}\}$.

In particular, now, H is cofinite.

CASE II. $\delta_1 = 0$ and $\delta_2 = 1$. First suppose $\delta_3 = 1$. By (i)–(xvi) and the facts that $F(\epsilon F) \wedge \neg Fx$ implies $x \neq \epsilon F$, that $\exists^{=s} y \neg Fy$ implies any $\neg\exists^{=t_\xi} y \neg Fy$ and that $F(\epsilon F)$ implies any $\neg\exists^{=r_\gamma} yFy$ we must consider the formula

$$\exists F(\neg Fx \wedge F(\epsilon F) \wedge \exists^{=s} y \neg Fy$$

$$\wedge \bigwedge_{l=1,\dots,P} F\mathsf{n}_l \wedge \bigwedge_{k=1,\dots,N} \neg F\mathsf{m}_k \wedge \bigwedge_{i=1,\dots,L} \mathsf{k}_i \neq \epsilon F),$$

where we assume $N = s-1$. Let the set defined by this formula be H. We claim that $H = \omega - (\{\bar{n}\} \cup \{r : r \notin \{\bar{m}\} \wedge \varepsilon(\omega - \{\bar{m}, r\}) \in \{\bar{k}\}\})$. For the left-to right inclusion, let $r \in H$. Clearly r is no n_l. If $r \notin \{\bar{m}\}$, by $r \in H$ there is some cofinite set F whose complement has s elements, contains $\{\bar{m}, r\}$ and is not mapped into $\{\bar{k}\}$ by ε. But $\omega - \{\bar{m}, r\}$ is the only possibility, so $\epsilon(\omega - \{\bar{m}, r\}) \notin \{\bar{k}\}$. For the other inclusion, let $r \notin \{\bar{n}\}$. If $r \notin \{\bar{m}\}$ and $\varepsilon(\omega - \{\bar{m}, r\}) \notin \{\bar{k}\}$, r satisfies the instance of our formula for $F := \omega - \{\bar{m}, r\}$. But every m_k is also in H since there are infinitely many $l \notin \{\bar{n}, \bar{m}\}$. In particular, H is cofinite.

Now suppose that $\delta_3 = 0$ and $\delta_4 = 1$. We must consider the formula

$$\exists F(x \neq \epsilon F \wedge \neg Fx \wedge \neg F(\epsilon F) \wedge \exists^{=v} yFy$$

$$\wedge \bigwedge_{l=1,\dots,P} F\mathsf{n}_l \wedge \bigwedge_{k=1,\dots,N} \neg F\mathsf{m}_k \wedge \bigwedge_{i=1,\dots,L} \mathsf{k}_i \neq \epsilon F).$$

There are infinitely many $(v - P)$-element sets $\{\bar{a}\}$ disjoint from $\{\bar{m}\}$ such that the cardinality of $\{\bar{n}, \bar{a}\}$ is v; take one such that $\varepsilon(\{\bar{n}, \bar{a}\}) \notin \{\bar{k}\}$. Then $\omega - \{\bar{n}, \bar{a}, \varepsilon(\{\bar{n}, \bar{a}\})\}$ is a cofinite set contained in the set defined by our formula which, therefore, is also cofinite. $\square$

4. DISCUSSION

Like Heck's H, T_Δ also proves the existence of non-logical objects. In fact, more is true:

THEOREM 4.1. T_Δ derives $\neg\exists H \forall x(Hx \leftrightarrow \exists F(x = \epsilon F))$. That is, T_Δ refutes the existence of the concept of being a value-range.

Proof. Argue in T_Δ. Assume $\forall x(Hx \leftrightarrow \exists F(x = \epsilon F))$. Then we have the following:

$$\forall x(\exists F(x = \epsilon F \wedge \neg Fx) \leftrightarrow (Hx \wedge \forall G(x = \epsilon G \rightarrow \neg Gx))).$$

For suppose first that $x = \epsilon F \wedge \neg Fx$. Clearly, by our assumption, Hx. If $x = \epsilon G$, we have by axiom V that $\forall y(Fy \leftrightarrow Gy)$, and thus, by $\neg Fx$, also $\neg Gx$. Now suppose $Hx \wedge \forall G(x = \epsilon G \rightarrow \neg Gx)$. By Hx, for some F, $x = \epsilon F$. But then also $\neg Fx$.

Since $Hx \wedge \forall G(x = \epsilon G \rightarrow \neg Gx)$ is equivalent to the Π^1_1-formula $\forall G(Hx \wedge (x = \epsilon G \rightarrow \neg Gx))$, we may invoke Δ^1_1-CA to obtain a K with $\forall x(Kx \leftrightarrow \exists F(x = \epsilon F \wedge \neg Fx))$, from which a contradiction follows via the usual Russell-argument. $\square$

COROLLARY 4.2. T_Δ derives $\exists x \forall F(x \neq \epsilon F)$.

Proof. $\mathsf{T}_\Delta + \forall x \exists F(x = \epsilon F)$ proves $\exists H \forall x(Hx \leftrightarrow \exists F(x = \epsilon F))$, since the H with $\forall x(Hx \leftrightarrow x = x)$, existing by (trivial) Δ^1_1-CA, does the job. By the theorem, $\mathsf{T}_\Delta + \forall x \exists F(x = \epsilon F)$ is inconsistent; hence T_Δ proves $\neg\forall x \exists F(x = \epsilon F)$. $\square$

The theorem in fact implies that there must be infinitely many non-logical objects for T_Δ: Argue in T_Δ. Let $a_1, \ldots, a_k$ be any objects and suppose $\forall x(\bigvee_{i=1,\ldots,k} x = a_i \rightarrow \neg\exists F(x = \epsilon F))$, that is, no a_i is a value-range. Now suppose by way of contradiction that $\forall x(\neg\exists F(x = \epsilon F) \rightarrow \bigvee_{i=1,\ldots,k} x = a_i)$, that is, any non-logical object is one of the a_i. We then have that $\forall x(\exists F(x = \epsilon F) \leftrightarrow \bigwedge_{i=1,\ldots,k} x \neq a_i)$. But Δ^1_1-CA guarantees the existence of an H with $\forall x(Hx \leftrightarrow \bigwedge_{i=1,\ldots,k} x \neq a_i)$, and for this H we have $\forall x(Hx \leftrightarrow \exists F(x = \epsilon F))$. The negation of this sentence we know to be provable in T_Δ, and so we conclude that T_Δ proves

$\exists x(\neg\exists F(x = \epsilon F) \wedge \bigwedge_{i=1,\ldots,k} x \neq a_i)$. In other words, if all of the a_i are non-value-ranges, then there must be another object x, distinct from all the a_i, which is also a non-logical object.

T_Δ being consistent and T_Σ being inconsistent, we knew beforehand that T_Δ would refute certain instances of Σ^1_1-CA. One such is rather obvious, and we already checked it in proving the inconsistency of T_Σ; it is the instance $\exists H\forall x(Hx \leftrightarrow \exists F(x = \epsilon F \wedge \neg Fx))$. The observation of the last paragraph is slightly more surprising: In fact the simplest possible instance of Σ^1_1-CA is refuted in T_Δ, viz. $\exists H\forall x(Hx \leftrightarrow \exists F(x = \epsilon F))$. This instance can hardly be said to involve any kind of self-reference; it simply states that the image of the function ϵ is a concept. It does not seem to be impossible that the real culprit of Frege's system is the postulation of the concept of *being a value-range*, rather than a vicious circle phenomenon. In fact, there seems to be a certain analogy to the situation in axiomatic set theory, where the Russell paradox shows that there is no set of all sets. The collection of all objects governed by an axiom of extensionality (such as axiom V) is as problematic here as it is in T_Δ, where we have that there is no concept – *a fortiori*, no value-range (set) – of all value-ranges (sets) (whereas the concept of *everything*, i.e. an H such that $\forall x(Hx \leftrightarrow x = x)$ exists and gives rise to a value-range (set)). The set-theoretic universe V thus corresponds, under this analogy, to the non-concept of being a value-range $\exists F(x = \epsilon F)$. Similarly, it is unclear up to this day whether Quine's theory NF, postulating a universal set coinciding with the set of all sets, is consistent, whereas the (weak) theory NFU allowing urelemente is known to be consistent. Thus, the contradiction-free theory NFU might well become inconsistent by subjecting every object to extensionality.

Let us note that the non-logical objects – or urelemente, if you prefer – proved to exist by H and T_Δ are rather mysterious things: All we can say about them is that they are there. In particular, no solution to the Caesar problem (to specify truth conditions for $x = \epsilon F$ for variable x) seems to be forthcoming: Just as in the case of Hume's principle, the only apparent solution $x = \epsilon F \leftrightarrow \exists\, G(x = \epsilon G \wedge \forall x(Fx \leftrightarrow Gx))$ is circular (for more on the Caesar problem, see Heck (1997b).

4.1. *Frege's Proof of Referentiality*

This takes us to a problematic point in Frege's proof (sections 29–32 of Frege, 1893) that every name of his system has a denotation. In section 31, he writes (I am quoting from Heck (1997a, 457); the brackets are Heck's):

> The question is whether '$\xi = \hat{y}\Phi(y)$' is a denoting name of a first-level function of one argument, and to that end it is to be asked in turn whether all proper names, that result from our substituting in the argument-place either a name of a truth-value or a fair value-

range name, denote. By our stipulation, that '$\hat{y}\Phi(y) = \hat{y}\Psi(y)$' is always to have the same denotation as '$\forall x(\Phi(x) \leftrightarrow \Psi(x))$', that [the True is identical with its own unit class], and that [the False is identical with its own unit class], a denotation is thus secured in every case for a proper name of the form '$\Gamma = \Delta$' ...

Heck proposes that 'we can take Frege tacitly restricting the domain of the theory to truth-values and value-ranges' (1997a, 458), where by the stipulations of Frege's section 10 the truth values are value-ranges themselves, and suggests that '[if] so, his considering only the two sorts of instances he does is not a flaw in his argument' (ibid. 460).[4]

Now there are really a number of points here that need to be discussed. In an obvious sense, something *is* wrong with Frege's restriction (if that is what he did), because the validity of *Grundgesetz* V (in the context of either Heck's H or T_Δ) already implies the existence of infinitely many non-value-ranges. Heck only seems to claim that making such a restriction did not constitute an *obvious* error at the relevant point of the proof. But it is not even clear to me whether *this* form of Heck's claim can be maintained: Granted, the contradiction arising from *Grundgesetz* V plus $\forall x \exists F(x = \hat{z}Fz)$ in H and T_Δ cannot possibly have been obvious to Frege. Still, not verifying the crucial assumption that *Grundgesetz* V holds in the domain consisting of truth-values and value-ranges only would seem to be an obvious error.

Another question is, of course, whether Frege did indeed intend to restrict the domain to value-ranges only. There are at least two ways to understand this question, or rather to interpret the notion of restriction here:[5] First, to tacitly add a new axiom $\forall x \exists F(x = \hat{y}Fy)$ to the theory, or, second, to put all and only those things that really are value-ranges into the domain, and then interpret the theory in this domain.

As a proponent of the first alternative we may take Edward Martin (1982), who has claimed the following:

(...) two principles Frege holds true: every function has a course-of-values [i.e. a value-range, K.W.], and every object of the *Grundgesetze* theory is a course-of-values (...) (Martin, 1982, 160)

This seems untenable for the following reasons. The extensive footnote to section 10 (Frege, 1893, 18) shows that Frege did consider the chances of forcing everything to be a value-range without reaching a final answer. Is it pausible to assume that he naively accepted such a principle only a few sections later? No, and in fact he did not. In section 34, p. 53, Frege introduces his version of the $\in$-relation, explicitly considering the case that in a formula $a \in u$, u is *not* a value-range; in section 35, p. 54, this point is taken up again:

Wenn als Argument der Funktion $[2 \in \xi]$ ein Gegenstand genommen wird, der kein Werthverlauf ist, so haben wir kein entsprechendes Argument der Function zweiter Stufe $\varphi(2)$ und die gegenseitige Vertretbarkeit der beiden Functionen erster und zweiter Stufe hört auf.

If as argument for the function $[2 \in \xi]$ an object is taken which is not a value-range, then we have no corresponding argument for the function of second order $\varphi(2)$ and the mutual replaceability of the two functions of first and second order terminates. [My translation.]

Finally, the last paragraph of section 36 also treats of the case, concerning membership in double-value-ranges, in which the object in question is not a value-range. It seems highly unlikely that Frege would have bothered with such questions had he assumed that these *waste cases* do not occur at all. It is worth noting that, in the sections quoted, we are still in the semantical part of *Grundgesetze* where Frege's main concern is to argue that every primitive, complex, or defined name of his system has a denotation. Only here do the *waste cases* play a role at all; they are not made part of the axioms of the formal theory. Thus the only axiomatic stipulation concerning the description operator \ included in the theory is *Grundgesetz* VI: $a = \backslash\hat{x}(a = x)$; there is no mention of the default definition given at the end of section 11, viz. $\neg\exists y(a = \hat{x}(x = y)) \rightarrow \backslash a = a$, in the later sections, and Frege nowhere claims that it should be a *theorem* of *Grundgesetze* (although, via the contradiction, it of course *is* a theorem).

Heck, on the other hand, seems to be proposing the second alternative suggested above. According to him, Frege is simply assuming that non-value-ranges like Caesar just aren't in the domain. There are, I believe, serious problems for this interpretation as well: Since $\hat{y}(y = \text{Caesar})$ *really* is a value-range, it would have to be in the domain; but then $\backslash\hat{y}(y = \text{Caesar})$ which, according to *Grundgesetz* VI, just is Caesar, would be in the domain as well. We might try to remedy this situation by taking the domain to consist of all and only those *real* value-ranges which can be named in the language of *Grundgesetze*. But $\hat{y}(\forall F(y \neq \hat{z}Fz))$ really *is* such a value-range – it is the extension of a concept under which, among many other things, Julius Caesar falls: We must put it into our domain. Now in the real world we have $\hat{y}(\forall F(y \neq \hat{z}Fz)) \neq \hat{y}(y \neq y)$. In our model, there are two possibilities:

(i) The sentence '$\hat{y}(\forall F(y \neq \hat{z}Fz)) = \hat{y}(y \neq y)$' is true in the model. First of all, it would be obvious that absoluteness between the real world and the model fails. Worse still, the validity of this sentence would entail that of '$\forall x\exists Fx = \hat{y}(Fy)$', and we are back to the problems of the first approach.

(ii) '$\hat{y}(\forall F(y \neq \hat{z}Fz)) \neq \hat{y}(y \neq y)$' is true in the model. It follows by *Grundgesetz* V that $\exists x \forall Fx \neq \hat{y}(Fy)$ holds in the model, and in this case, Frege's considering only the two sorts of instances he does *is* a flaw in his argument, as we pointed out before.

4.2. *Logicism?*

Of course *Grundgesetze* is disqualified as a logicist system by its inconsistency. But let us ignore this cruel fact for a moment. It has been argued by some – notably George Boolos, see e.g. (Boolos, 1987) – that the provability of the existence of an infinity of (logical?) objects would have disqualified *Grundgesetze* from being *logic*ist anyway. This argument is so anachronistic that it seems quite unsatisfactory to me: Evidently Frege *wanted* his theory to prove the existence of infinitely many objects and still conceived of it as logical. And what if there really *are* infinitely many logical objects – why should logic not prove their existence?

Be that as it may, one might argue that the provability of the existence of infinitely many objects other than logical ones is a *reductio ad absurdum* of a logicist system.[6] This seems rather convincing to me. Of course, our proofs of $\exists x \forall F(x \neq \epsilon F)$ made essential use of the reasoning leading to Russell's paradox, so that the proposition 'Had *Grundgesetze* been consistent, it would have been a failure anyway since it would still have proved $\exists x \forall F(x \neq \epsilon F)$' may not make much sense. For consistent subtheories such as H or T_Δ, however, the argument remains intact. That is, regardless of how much arithmetic is interpretable in such fragments, it does not seem possible to claim that these fragments supply a logicist foundation for those parts of arithmetic.

4.3. *Cantor's Theorem*

I take Cantor's theorem to be the assertion that there can be no bijection between the individuals and the second-order entities of a given domain. By the reasoning of Russell's paradox we have seen that Σ^1_1-CA in the language L of T as defined at the beginning of Section 3 above (and *a fortiori* full axiomatic second-order logic in that language) proves a version of Cantor's theorem:

$$\Sigma^1_1\text{-CA} \vdash \forall FG(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx)) \rightarrow \bot.$$

We may say that Σ^1_1-CA proves the (third-order, if you wish) assertion that any function from the second- to the first-order entities is non-injective. As we have seen, this is not the case for Δ^1_1-CA since T_Δ is consistent. One might think that Cantor's theorem must then be independent of Δ^1_1-CA. Not so! As we have seen, Δ^1_1-CA proves

$$\forall FG(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx)) \rightarrow \exists x \forall F(x \neq \epsilon F),$$

that is: any injective function from the second-order entities into the individuals *is not onto*, and hence there is no bijection!

Curiously, thus, Σ^1_1-CA implies that there are too many second-order entities, hence Cantor's theorem, while Δ^1_1-CA implies that, if there are not too many second-order entities, then there are too few second-order entities, hence Cantor's theorem. Put slightly differently: While T_Σ is inconsistent due to a lack of individuals, its consistent subtheory T_Δ proves that there might be more than enough.[7]

NOTES

[1] See also Heck (1997c), p. 12. Of course, the truth values are also logical objects. In *Grundgesetze*, section 10, it is claimed that these may be identified with certain value-ranges. Even if this were not so, there must be, according to both H and T_Δ, objects other than the two truth-values and the value-ranges. This is so because these theories prove the existence of *infinitely many* non-value-ranges. Also, it does not seem that this is a peculiarity of formalizing the theory as one of second-order logic, rather than a term logic like Frege's original system: The truth-values, in that system, do not play any systematic role *qua* truth-values. That something Γ is the True is denoted by $\vdash \Gamma$, not by $\vdash \Gamma = \top$, where $\top$ is some name of the True.

[2] Nino Cocchiarella's work is highly relevant to this point. See, for instance, Cocchiarella, 1985; 1986; 1992. Note in particular that his theories λHST^* and HST^*_λ actually *refute* Cantor's theorem.

[3] In particular, the counterexamples r_n, while not denoting value-ranges, *are* value-range terms.

[4] In this regard, see also the careful discussion by Matthias Schirn (1996, 2–13).

[5] Thanks to Richard Heck for pointing out this distinction to me.

[6] Cf. Heck, 1997c, p. 12n: '[C]onsider '$\exists x \forall F.x \neq \hat{x}Fx$', which asserts that some object is not a value-range. (...) But the question whether there are non-logical objects is none in the province of *logic*.'

[7] Discussions with Lev Beklemishev, Justus Diller and Michael Möllerfeld made me realize which structure could serve as a model for the theory T_Δ. Andrea Cantini, Justus Diller and Gottfried Gabriel read an early draft of this paper and supplied questions and remarks that helped to improve upon the text. I am grateful to an anonymous *Synthese* referee for indicating the need for conceptual clarification at some points. Special thanks to Richard Heck for extensive comments on an earlier version.

REFERENCES

Boolos, G.: 1987, 'The Consistency of Frege's *Foundations of Arithmetic*', in J. J. Thomson (ed.), *On Being and Saying: Essays for Richard Cartwright*, MIT Press, Cambridge, MA, pp. 3–20.

Cocchiarella, N. B.: 1985, 'Frege's Double Correlation Thesis and Quine's Set Theories NF and ML', *Journal of Philosophical Logic* **14**, 1–39.

Cocchiarella, N. B.: 1986, *Logical Investigations of Predication Theory and the Problem of Universals*, Bibliopolis, Naples.

Cocchiarella, N. B.: 1992, 'Cantor's Power-Set Theorem versus Frege's Double-Correlation Thesis', *History and Philosophy of Logic* **13**, 179–201.

Frege, G.: 1893, *Grundgesetze der Arithmetik I*, Hermann Pohle, Jena.

Frege, G.: 1976, *Wissenschaftlicher Briefwechsel*, G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, and A. Veraart (eds.), Felix Meiner, Hamburg.

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

Heck, R.: 1997a, 'Grundgesetze der Arithmetik I §§ 29–32', *Notre Dame Journal of Formal Logic* **38**, 437–474.

Heck, R.: 1997b, 'The Julius Caesar Objection', in R. Heck (ed.), *Language, Thought, and Logic*, Oxford University Press, Oxford, pp. 273–308.

Heck, R.: 1997c, 'Frege and Semantics', in T. Ricketts (ed.), forthcoming, *The Cambridge Companion to Frege*, preprint, 39 pp.

Martin, E.: 1982, 'Referentiality in Frege's Grundgesetze', *History and Philosophy of Logic* **3**, 151–164.

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

Schirn, M.: 1996, 'Introduction: Frege on the Foundations of Arithmetic and Geometry', in M. Schirn (ed.), *Frege: Importance and Legacy*, de Gruyter, Berlin/New York, pp. 1–42.

Schroeder-Heister, P.: 1987, 'A Model-Theoretic Reconstruction of Frege's Permutation Argument', *Notre Dame Journal of Formal Logic* **28**, 69–79.

Philosophical Institute
Rijksuniversiteit Leiden
Matthias de Vrieshof 4
Postbus 9515
2300 RA Leiden
The Netherlands
E-mail: wehmeier@rullet.leidenuniv.nl