THE JOURNAL OF SYMBOLIC LOGIC
Volume 60, Number 1, March 1995

TYPE REDUCING CORRESPONDENCES AND WELL-ORDERINGS: FREGE'S AND ZERMELO'S CONSTRUCTIONS RE-EXAMINED

J. L. BELL

A key idea in both Frege's development of arithmetic in the ***Grundlagen*** [7] and Zermelo's 1904 proof [10] of the well-ordering theorem is that of a "type reducing" correspondence between second-level and first-level entities. In Frege's construction, the correspondence obtains between *concept* and *number*, in Zermelo's (through the axiom of choice), between *set* and *member*. In this paper, a formulation is given and a detailed investigation undertaken of a system $\mathscr{F}$ of many-sorted first-order logic (first outlined in the Appendix to [6]) in which this notion of type reducing correspondence is accorded a central role and which enables Frege's and Zermelo's constructions to be presented in such a way as to reveal their essential similarity. By adapting Bourbaki's version of Zermelo's proof of the well-ordering theorem, we show that, within $\mathscr{F}$, any correspondence c between second-level entities (here called *concepts*) and first-level ones (here called *objects*) induces a well-ordering relation $W(c)$ in a canonical manner. We shall see that, when c is the "Fregean" correspondence between concepts and cardinal numbers, $W(c)$ is (the well-ordering of) the ordinal $\omega + 1$, and when c is a "Zermelian" choice function on concepts, $W(c)$ is a well-ordering of the universal concept embracing all objects.

In $\mathscr{F}$ an important role is played by the notion of *extension* of a concept. To each concept X we assume there is assigned an object $e(x)$ in such a way that, for any concepts X, Y satisfying a certain predicate E, we have $e(X) = e(Y)$ iff the same objects fall under X and Y. For concepts X satisfying E, $e(X)$ is called the *extension* of X. (E thus represents the property of possessing an extension: Russell's paradox implies that not every concept can satisfy E.) We show that the canonical well-ordering $W(e)$ induced by the correspondence e is precisely that of the *von Neumann ordinals*.

Since the concepts of set and membership are manifestly *not* primitive in $\mathscr{F}$ (although we do eventually introduce them by definition), it follows that the constructions analyzed anew here—Frege's development of the natural number system, Zermelo's proof of the well-ordering theorem, even the formation of the von Neumann ordinals—need not be regarded as "set-theoretical" constructions as such, but rather as natural outgrowths of the four "logical" notions of the system $\mathscr{F}$: *concept, object, predication, extension.*

Received November 3, 1993; revised July 12, 1994.
The author is grateful to William Demopoulos for many illuminating discussions on the subject matter of this paper. Assistance in its preparation was provided by NSERC of Canada.

0022-4812/95/6001-0012/$02.30

§1. **The system $\mathscr{F}$.** $\mathscr{F}$ is a system of many-sorted first-order logic possessing five sorts to which we assign the following names: I: *objects*, II: *basic* (*or first-level*) *concepts*, III: *relations*, IV: *second-level concepts*, V: *second-level relational concepts*.

We assume that $\mathscr{F}$ contains the following *variables and constants* of each sort:

Sort	Variable	Constant
I	$x, y, z, \ldots$	$a, b, c, \ldots$
II	$X, Y, Z, \ldots$	$A, B, C, \ldots$
III	$\underline{X}, \underline{Y}, \underline{Z}, \ldots$	$\underline{A}, \underline{B}, \underline{C}, \ldots$
IV	$\underline{\underline{X}}, \underline{\underline{Y}}, \underline{\underline{Z}}, \ldots$	$\underline{\underline{A}}, \underline{\underline{B}}, \underline{\underline{C}}, \ldots$
V	$X^*, Y^*, Z^*, \ldots$	$A^*, B^*, C^*, \ldots$

In addition to variables and constants, $\mathscr{F}$ and its extensions will contain various additional terms. Terms of sorts II, IV, or V will be called *concept terms*.

We assume the presence in $\mathscr{F}$ of an *identity symbol* $=$ yielding atomic statements of the form $\sigma = \tau$, where σ and τ are terms of the same sort.

We also assume the presence in $\mathscr{F}$ of a *predication symbol* η yielding atomic statements of the form $s\eta t$, $(s', t')\eta u$, where s is of sort I, II, III and t is of sort II, IV, V; and s', t' are both of sort I and u is of sort III. We read "$s\eta t$" as "*s falls under t.*"

The sole axiom scheme in $\mathscr{F}$ is the

COMPREHENSION SCHEME FOR CONCEPTS. *Corresponding to any formula $\phi(x)$, $\phi(x, y)$, $\phi(X)$ or $\phi(\underline{X})$ we are given a term s of sort* II, III, IV, *or* V, *respectively, for which we adopt as axioms the formulas*

$$\forall\omega[\omega\eta s \leftrightarrow \phi(\omega)],$$

where ω is x, X or $\underline{X}$, and

$$\forall x\forall y[(xy)\eta s \leftrightarrow \phi(x, y)].$$

We write $x^\frown\phi, (xy)^\frown\phi, X^\frown\phi, \underline{X}^\frown\phi$ for s, as the case may be. A term of the first, third, or fourth types is called the *concept term* determined by ϕ, and a term of the second type the *relation term* determined by ϕ.

We define the relation $\equiv$ of *extensional equality* of terms by

$$s \equiv t \Leftrightarrow_{\mathrm{df}} \forall\omega(\omega\eta s \leftrightarrow \omega\eta t),$$

where s, t are concept terms of the same sort (and ω a variable of the appropriate sort), and

$$\underline{X} \equiv \underline{Y} \Leftrightarrow_{\mathrm{df}} \forall x\forall y[(xy)\eta\underline{X} \leftrightarrow (xy)\eta\underline{Y}].$$

Clearly, concept terms are determined uniquely by formulas up to *extensional equality*.

If $\mathscr{E}$ is $\mathscr{F}$ or any of its extensions (to be introduced presently), we write "$\mathscr{E} \vdash \phi$" for "*ϕ is first-order inferrable from the axioms of $\mathscr{E}$.*"

§2. **The Zermelo-Bourbaki lemma.** Our next task is to show within $\mathscr{F}$ that any assignment of objects to (basic) concepts yields a well-ordering relation in a canonical manner. The argument establishing this is an adaptation of that used by Bourbaki ([5, Chapitre 3, §2, Lemme 3]); the proof is itself a generalization of Zermelo's argument in [10] to prove Zermelo's well-ordering theorem.

We first make the following definitions in $\mathscr{F}$:

$$X \subseteq Y \Leftrightarrow_{\text{df}} \forall x(x\eta X \rightarrow x\eta Y),$$

$$\underline{X} \subseteq \underline{Y} \Leftrightarrow_{\text{df}} \forall xy[(xy)\eta \underline{X} \rightarrow (xy)\eta \underline{Y}],$$

$$\text{Field}(\underline{X}) =_{\text{df}} x\hat{}[\exists y[(xy)\eta \underline{X} \vee (yx)\eta \underline{X}]],$$

$$\begin{aligned}\text{WO}(\underline{X}) \Leftrightarrow_{\text{df}} & \forall x[x\eta\, \text{Field}(\underline{X}) \rightarrow (xx)\eta \underline{X}] \wedge \forall xy[(xy)\eta \underline{X} \wedge (yx)\eta \underline{X} \rightarrow x = y] \\ & \wedge \forall xyz[(xy)\eta \underline{X} \wedge (yz)\eta \underline{X} \rightarrow (xz)\eta \underline{X}] \\ & \wedge \forall Y[Y \subseteq \text{Field}(\underline{X}) \wedge \exists x.x\eta Y \rightarrow \exists x[x\eta Y \wedge \forall y[y\eta Y \rightarrow (xy)\eta \underline{X}]]]\end{aligned}$$

($\text{WO}(\underline{X})$ thus says that $\underline{X}$ is a *well-ordering*),

$$Y \textit{ is an initial segment of } \underline{X} \Leftrightarrow_{\text{df}} Y \subseteq \text{Field}(\underline{X}) \wedge \forall xy[y\eta Y \wedge (xy)\eta \underline{X} \rightarrow x\eta Y],$$

$$\underline{X}^x =_{\text{df}} y\hat{}.(yx)\eta \underline{X},$$

$$\underline{X}_x =_{\text{df}} y\hat{}[x \neq y \wedge (yx)\eta \underline{X}]$$

($\underline{X}^x$ and $\underline{X}_x$ are, respectively, the initial segment and the strict initial segment of $\underline{X}$ determined by x),

$$[x] =_{\text{df}} y\hat{}(x = y),$$

$$X - Y =_{\text{df}} x\hat{}[x\eta X \wedge \neg x\eta Y],$$

$$X \cup Y =_{\text{df}} x\hat{}(x\eta X \vee x\eta Y),$$

$$\underline{X}|Y =_{\text{df}} (xy)\hat{}[(xy)\eta \underline{X} \wedge x\eta Y \wedge y\eta Y],$$

$$\text{Stab}(\underline{\underline{X}}) \Leftrightarrow_{\text{df}} \forall XY(X \equiv Y \wedge X\eta \underline{\underline{X}} \rightarrow Y\eta \underline{\underline{X}}).$$

A term $\underline{\underline{S}}$ of sort IV is said to be *stable* if $\mathscr{F} \vdash \text{Stab}(\underline{\underline{S}})$.

Now suppose given a constant $\underline{\underline{S}}$ of sort IV and a term t such that $t(\sigma)$ is of sort I for any term σ of sort II. Let $\mathscr{F}(\underline{\underline{S}}, t)$ be obtained from $\mathscr{F}$ by adding as axioms the statements $\text{Stab}(\underline{\underline{S}})$ and $\forall XY[X \equiv Y \wedge X\eta \underline{\underline{S}} \rightarrow t(X) = t(Y)]$. We now prove in $\mathscr{F}(\underline{\underline{S}}, t)$ the

2.1. Zermelo-Bourbaki Lemma. *We can construct a relation* $\underline{R}$ *from* $\underline{\underline{S}}$ *and* t *such that* $\mathscr{F}(\underline{\underline{S}}, t) \vdash \text{WO}(\underline{R})$ *and, writing* M *for* $\text{Field}(\underline{R})$,

(i) $\mathscr{F}(\underline{\underline{S}}, t) \vdash \forall x[x\eta M \rightarrow \underline{R}_x \eta \underline{\underline{S}} \wedge t(\underline{R}_x) = x]$,

(ii) $\mathscr{F}(\underline{\underline{S}}, t) \vdash M\eta \underline{\underline{S}} \rightarrow t(M)\eta M$.

Proof. Define

$$U^* =_{\text{df}} X\hat{}[\text{WO}(\underline{X}) \wedge \forall x[x\eta\, \text{Field}(\underline{X}) \rightarrow \underline{X}_x \eta \underline{\underline{S}} \wedge x = t(\underline{X}_x)]],$$

$$\underline{R} =_{\text{df}} (xy)\hat{}[\exists \underline{X}[\underline{X}\eta U^* \wedge (xy)\eta \underline{X}]].$$

Thus $\underline{R}$ is the "union" of the second level relational concept U^*.

We claim that $\mathscr{F}(\underline{\underline{S}}, t) \vdash \text{WO}(\underline{R})$. Since $\underline{R}$ is the "union" of well-orderings, for this to be the case it suffices—by the usual order-theoretic argument (which is easily formulable in $\mathscr{F}$)—to prove within $\mathscr{F}(\underline{\underline{S}}, t)$ that, for any $\underline{X}, \underline{Y}$ such that $\underline{X}\eta U^*, \underline{Y}\eta U^*$, one, $\underline{X}$ say, is included in the other and in that case $\text{Field}(\underline{X})$ is an initial segment of Field($\underline{Y}$). To prove this, arguing within $\mathscr{F}(\underline{\underline{S}}, t)$, write W for Field($\underline{X}$), Z for Field($\underline{Y}$), and define

$$P =_{\text{df}} x\,\hat{}\,[x\eta W \wedge x\eta Z \wedge \underline{X}_x \equiv \underline{Y}_x \wedge \underline{X}|\underline{X}_x \equiv \underline{Y}|\underline{X}_x].$$

Then P is an initial segment of $\underline{X}$ and $\underline{Y}$, and $\underline{X}|P \equiv \underline{Y}|P$. It thus suffices to show that $P \equiv W$ or $P \equiv Z$.

Suppose that neither $P \equiv W$ nor $P \equiv Z$. Then $\exists x.x\eta(W - P)$ and $\exists x.x\eta(Z - P)$. Let a, b be the "least elements" of $W - P, Z - P$ with respect to the well-orderings $\underline{X}, \underline{Y}$ respectively. Then $P \equiv \underline{X}_a$ and $P \equiv \underline{Y}_b$. But since $\underline{X}\eta U^*$, we have $\underline{X}_a \eta \underline{\underline{S}}$, so, since $\underline{\underline{S}}$ is stable, $P\eta\underline{\underline{S}}$. Also $a = t(\underline{X}_a) = t(P)$, $b = t(\underline{Y}_b) = t(P)$. From the definition of P it now follows that $a\eta P$; since $P \equiv \underline{X}_a$ we conclude that $a\eta\underline{X}_a$, and we have a contradiction.

The claim is, accordingly, proved. Then $M =_{\text{df}}$ Field($\underline{R}$) satisfies (i). To establish (ii), again arguing within $\mathscr{F}(\underline{\underline{S}}, t)$, suppose that $M\eta\underline{\underline{S}}$. Define

$$\underline{O} =_{\text{df}} (xy)\,\hat{}\,[(xy)\eta\underline{R} \vee x = y = t(M) \vee [x\eta M \wedge y = t(M)]].$$

($\underline{O}$ is the well-ordering obtained from $\underline{R}$ by "tacking on" $t(M)$ as last "element".) Writing m for $t(M)$, we have $\underline{O}_m \equiv M$, whence $\underline{O}_m\eta\underline{\underline{S}}$ and $t(\underline{O}_m) = t(M) = m$. Hence $\underline{O}\eta U^*$, so Field($\underline{O}$) $\subseteq M$. Since $m\eta$ Field($\underline{O}$), it follows that $m\eta M$, proving (ii). ■

As a consequence we obtain a version of *Zermelo's well-ordering theorem*. Let $\mathscr{F}(\varepsilon)$ be obtained by adding to $\mathscr{F}$ a term ε such that $\varepsilon(\sigma)$ is of sort I for any term σ of sort II, together with the axioms

(a) $\forall X(\exists x.x\eta X \rightarrow \varepsilon(X)\eta X)$

(b) $\exists XY(X \equiv Y \rightarrow \varepsilon(X) = \varepsilon(Y))$.

Axiom (a) is, essentially, *Hilbert's epsilon axiom*, and (b) is *Ackermann's extensionality principle* (cf. [1]). A formulation and discussion of a principle similar to axiom (a) within a framework resembling the present one may be found in [4].

Define the *universal concept* $V =_{\text{df}} x\,\hat{}\,.x = x$. Then we have

2.2. COROLLARY (the well-ordering principle). *We can construct a relation $\underline{R}$ from ε such that*

$$\mathscr{F}(\varepsilon) \vdash \text{WO}(\underline{R}) \wedge \text{Field}(\underline{R}) = V.$$

PROOF. We apply the Zermelo-Bourbaki lemma in $\mathscr{F}(\varepsilon)$ with $\underline{\underline{S}}$ the second-level concept $X\,\hat{}\,(\exists x.\neg x\eta X)$ and t the term given by $t(X) =_{\text{df}} \varepsilon(x\,\hat{}\,.\neg x\eta X)$. Arguing in $\mathscr{F}(\varepsilon)$, we obtain a well-ordering $\underline{R}$ with field M such that, writing m for $t(M)$,

$$(*) \qquad \exists x.\neg x\eta M \rightarrow m\eta M.$$

But from condition (a) above we have, writing M' for $x\,\hat{}\,.\neg x\eta M$ and noting that $\varepsilon(M') = m$,

$$\exists x.\neg x\eta M \rightarrow \exists x.x\eta M' \rightarrow \varepsilon(M')\eta M' \rightarrow \neg m\eta M.$$

We conclude from this and $(*)$ that $\forall x.x\eta M$. Thus $M = V$, and the result is proved. ■

§3. **Extensions of concepts and Frege's theorem.** We form the system $\mathscr{F}_e$ by adding to $\mathscr{F}$

- a term e such that $e(\tau)$ is well-formed and of sort I for any concept term τ, and
- a predicate symbol E such that $E(\tau)$ is well-formed for any concept term τ,

and the *axioms*

$$(\mathrm{Ext}_1) \qquad \forall \upsilon \forall \omega [E(\upsilon) \wedge E(\omega) \rightarrow [e(\upsilon) = e(\omega) \leftrightarrow \upsilon \equiv \omega]],$$

$$(\mathrm{Ext}_2) \qquad \forall \upsilon \forall \omega [E(\upsilon) \wedge \upsilon \equiv \omega \rightarrow E(\omega)],$$

where υ, ω are concept variables of the same sort.

If we think of $e(\tau)$ as an object *representing* τ, (Ext_1) expresses the idea that *extensional equality* of concepts satisfying E is equivalent to *identity* of their representing objects. That is, for any concept τ satisfying E, $e(\tau)$ may be regarded as the *extension* of τ. And the predicate E represents the property of *possessing an extension.* As for (Ext_2), it states the reasonable requirement that any concept extensionally equivalent to a concept possessing an extension itself possesses one (that is, $\equiv$ is a congruence relation with respect to E). If $\mathscr{F}_e \vdash \neg E(\tau)$, we shall say that the concept τ *does not possess an extension.*

Now a short Russell-type argument (which we leave to the reader) in $\mathscr{F}_e$ shows that *the non-self-predication concept*

$$x^\frown[\exists X(E(X) \wedge e(X) = x \wedge \neg x\eta X)]$$

does not possess an extension. And nor, indeed, does the *well-foundedness concept*

$$\mathrm{WF} =_{\mathrm{df}} x^\frown[\forall X[\forall y[y \sqsubseteq X \rightarrow y\eta X] \rightarrow x\eta X]],$$

where $y \sqsubseteq X$ stands for $\exists Y[E(Y) \wedge e(Y) = y \wedge Y \subseteq X]$. Since the argument in this case is more intricate, here it is.

Suppose that, arguing in $\mathscr{F}_e$, $E(\mathrm{WF})$; let $a = e(\mathrm{WF})$. It is then easy to see that $a\eta\,\mathrm{WF}$. Now let $U =_{\mathrm{df}} x^\frown.x \neq a$. If $y \sqsubseteq U$, then, for some Y, $E(Y), e(Y) = y$ and $Y \subseteq U$, i.e. $\neg a\eta Y$. If $y = a$, then $e(\mathrm{WF}) = e(Y)$, whence $\mathrm{WF} \equiv Y$, so since $a\eta\,\mathrm{WF}$ we would have $a\eta Y$, a contradiction. So we have shown that $y \sqsubseteq U \rightarrow y \neq a \rightarrow y\eta U$. Therefore $\mathrm{WF} \subseteq U$; since $\neg a\eta U$ it follows that $\neg a\eta\,\mathrm{WF}$. This contradiction shows that WF does not possess an extension.

Since we have shown that not all concepts can possess extensions, it is natural to ask what concepts we need require to possess extensions in order to be able to carry out essential mathematical constructions in $\mathscr{F}_e$: in particular, the construction of the *natural number system.* It was Frege's remarkable discovery that for this it suffices just to assume that extensions are possessed by the members of a certain class of simple and natural *second-level* concepts, those that, following Boolos [3], we shall term *numerical*.

Numerical concepts are defined as follows. First we define the relations Bij and $\approx$ of *bijectivity* and *equinumerosity* as usual:

$$\begin{aligned}\mathrm{Bij}(\underline{Z}, X, Y) \Leftrightarrow_{\mathrm{df}} & \forall xy[(xy)\eta\underline{Z} \to x\eta X \wedge y\eta Y] \\ & \wedge \forall xyz[(xy)\eta\underline{Z} \wedge (xz)\eta\underline{Z} \to y = z] \\ & \wedge \forall x[x\eta X \to \exists y.(xy)\eta\underline{Z}] \wedge \forall y[y\eta Y \to \exists x.(xy)\eta\underline{Z}], \\ X \approx Y \Leftrightarrow_{\mathrm{df}} & \exists\underline{Z}\,\mathrm{Bij}(\underline{Z}, X, Y).\end{aligned}$$

Note that $\mathscr{F} \vdash X \equiv Y \to X \approx Y$. With any basic concept X we associate the second-level concept

$$\|X\| =_{\mathrm{df}} Y\hat{}.X \approx Y.$$

Concepts of the form $\|X\|$ are called *numerical*.

If we assume that every numerical concept possesses an extension, i.e. $\forall XE(\|X\|)$, then the extension $|X| =_{\mathrm{df}} e(\|X\|)$ is called the *cardinal number* of X. Objects of the form $|X|$ are called *cardinal numbers*. Under these conditions it is easy to derive what Boolos [3] calls *Hume's principle*, viz.

$$\forall XY[X \approx Y \leftrightarrow |X| = |Y|].$$

We make the following definitions.

$$\begin{aligned}X \text{ is (Dedekind) } \textit{infinite} & \Leftrightarrow_{\mathrm{df}} \exists Y[Y \subseteq X \wedge \neg(Y \equiv X) \wedge X \approx Y], \\ x \text{ is an } \textit{infinite number} & \Leftrightarrow_{\mathrm{df}} \exists X[X \text{ is } \textit{infinite} \wedge x = |X|], \\ \underline{X} \text{ is } \textit{infinite} & \Leftrightarrow_{\mathrm{df}} \mathrm{Field}(\underline{X}) \text{ is } \textit{infinite}.\end{aligned}$$

Clearly,

$$\mathscr{F} \vdash X \textit{infinite} \wedge X \approx Y \to Y \textit{infinite}.$$

Let $\mathscr{F}^*$ be the system obtained from $\mathscr{F}_e$ by adding the axiom $\forall XE(\|X\|)$. As a corollary to the Zermelo-Bourbaki lemma we derive

3.1. Frege's Theorem. *In $\mathscr{F}^*$ we can prove that there exists both an infinite well-ordering and an infinite number.*

Proof. Arguing in $\mathscr{F}^*$, we apply the Zermelo-Bourbaki lemma with $\underline{S}$ the universal second-level concept $X\hat{}.X = X$ and $t(X) =_{\mathrm{df}} |X|$. We obtain a well-ordering $\underline{F}$ such that, writing M for Field($\underline{F}$), and m for $|M|$,

$$\forall x(x\eta M \to |\underline{F}_x| = x) \quad \text{and} \quad m\eta M.$$

Then $|M| = m = |\underline{F}_m|$, whence $\underline{F}_m \approx M$. Since $\neg m\eta\underline{F}_m$ by definition, it follows that M, and hence $\underline{F}$ and m, are infinite. ■

In the ***Grundlagen*** Frege outlines a proof (reconstructed in detail in [3]), from principles similar to those laid down in $\mathscr{F}^*$, of the existence of the *natural number system*. We now show how this fact can be derived within $\mathscr{F}^*$ from our version of Frege's theorem. We show, moreover, that the infinite well-ordering $\underline{F}$ constructed in the proof of Frege's theorem is of order type $\omega + 1$.

We first establish in $\mathscr{F}$ a special case of the *Schröder-Bernstein theorem* (in point of fact equivalent to it), viz., the

3.2. EQUINUMEROSITY THEOREM.

$$\mathscr{F} \vdash \forall XYZ[X \subseteq Y \wedge Y \subseteq Z \wedge X \approx Z \rightarrow Y \approx Z].$$

PROOF. The assertion here is the formulation within $\mathscr{F}$ of Lemma 5.1.1 of [9]. To prove it within $\mathscr{F}$ we simply adapt and condense the proof given in [9]. Thus suppose that $X \subseteq Y \subseteq Z$ and $\mathrm{Bij}(\underline{R}, Z, X)$. Define

$$\underline{R}[U] =_{\mathrm{df}} y\hat{}[\exists x.x\eta U \wedge (xy)\eta\underline{R}],$$
$$\mathrm{Cl}_{\underline{R}}[U] =_{\mathrm{df}} x\hat{}.\forall Y[[U \subseteq Y \wedge \underline{R}[Y] \subseteq Y] \rightarrow x\eta Y]].$$

Then if we define

$$\underline{S} =_{\mathrm{df}} (xy)\hat{}[[(xy)\eta\underline{R} \wedge x\eta\,\mathrm{Cl}_{\underline{R}}[Z - Y]] \vee [x = y \wedge x\eta(Z - \mathrm{Cl}_{\underline{R}}[Z - Y])]]$$

it can be shown without difficulty that $\mathrm{Bij}(\underline{S}, Z, Y)$, whence $Y \approx Z$. ■

3.3. COROLLARY. $\mathscr{F} \vdash X\,\textit{infinite} \wedge x\eta X \rightarrow X \approx X - [x]$.

PROOF. Arguing in $\mathscr{F}$, if X is infinite then there is $Y \subseteq X$ such that $\neg(X \equiv Y)$ and $Y \approx X$. Suppose now that $x\eta X$. We consider two cases. 1) $\neg x\eta Y$. In this case $Y \subseteq X - [x] \subseteq X$ and $X - [x] \approx X$ by the equinumerosity theorem. 2) $x\eta Y$. In this case choose $y\eta(X - Y)$. Then $(Y - [x]) \cup [y] \subseteq X - [x] \subseteq X$ and it follows easily from $Y \approx X$ that also $(Y - [x]) \cup [y] \approx X$. The equinumerosity theorem again yields $X \approx X - [x]$. ■

Now write M for the field of the infinite well-ordering $\underline{F}$ obtained in the proof of Frege's theorem, and m for $|M|$. We recall that $\mathscr{F}^* \vdash m\eta M$. We shall write $x \leq y$ for $(xy)\eta\underline{F}$ and $x < y$ for $x \leq y \wedge x \neq y$.

3.4. PROPOSITION. $\mathscr{F}^* \vdash$ *m is the $\leq$-largest "element" of M.*

PROOF. Arguing in $\mathscr{F}^*$, if $m \leq x$, then $\underline{F}_m \subseteq \underline{F}_x \subseteq M$. But $m = |\underline{F}_m| = |M|$, so that $\underline{F}_m \approx M$. It follows from the equinumerosity theorem that $\underline{F}_x \approx M$, whence $x = |\underline{F}_x| = |M| = m$. ■

Now define $0 =_{\mathrm{df}} |x\hat{}.x \neq x|$. Then clearly we can prove in $\mathscr{F}^*$ that 0 is the *least element* of M. We also define $x^+ =_{\mathrm{df}} |\underline{F}^x|$. If x with $x\eta M$ has an $\underline{F}$-successor, then clearly x^+ is its *immediate successor*. This suggests that we define

$$X \text{ is } \textit{inductive} \Leftrightarrow_{\mathrm{df}} X \subseteq M \wedge 0\eta X \wedge \forall x(x\eta X \rightarrow x^+\eta X).$$

3.5. PROPOSITION. (i) $\mathscr{F}^* \vdash x < m \rightarrow x^+ < m$.

(ii) $\mathscr{F}^* \vdash m^+ = m$.

(iii) $\mathscr{F}^* \vdash x^+ = y^+ \rightarrow x = y$.

PROOF. Throughout, we argue in $\mathscr{F}^*$.

(i) Suppose $x^+ = m$. Then $|M| = |\underline{F}^x|$, so that $\underline{F}^x$ is infinite. It now follows from 3.3 that $\underline{F}_x \equiv \underline{F}^x - [x] \approx \underline{F}^x$, whence $x = |\underline{F}_x| = |\underline{F}^x| = |M| = m$.

(ii) We have $|M| = m = |\underline{F}_m|$, so, since $\underline{F}_m \subseteq \underline{F}^m \subseteq M$, the equinumerosity theorem implies that $\underline{F}^m \approx M$, whence $m^+ = |M| = m$.

(iii) If $x, y < m$, then (iii) follows from the fact that in this case x^+ (y^+) is the immediate successor of x (y). On the other hand if, say, $y = m$, then $x^+ = m^+$ and (ii) gives $m = m^+ = x^+$. Unless $x = m$ we would violate (i). ■

3.6. PROPOSITION. $\mathscr{F}^* \vdash \underline{E}_m$ *is inductive.*

PROOF. We have (arguing in $\mathscr{F}^*$) $0\eta\underline{E}_m$, since if $m = 0$ then $|M| = |\underline{R}_0|$, but $\underline{R}_0$ cannot be infinite since it is empty. The rest follows immediately from 3.5(i). ■

Define the *natural number concept N* by

$$N =_{\text{df}} x\hat{\ }\forall X(X \text{ inductive} \rightarrow x\eta X).$$

Clearly N is the least inductive concept. It follows immediately from the definitions of 0 and N and from 3.5(iii) that N *satisfies Peano's axioms*, i.e.

$$\mathscr{F}^* \vdash \forall xy[x\eta N \wedge y\eta N \rightarrow (0 \neq x^+) \wedge (x^+ = y^+ \rightarrow x = y)] \wedge \forall X[X \text{ inductive} \rightarrow N \subseteq X].$$

3.7. PROPOSITION. $\mathscr{F}^* \vdash N$ *is an initial segment of M.*

PROOF. For this it suffices to show that, arguing in $\mathscr{F}^*$,

$$A =_{\text{df}} x\hat{\ }\forall y(y \leq x \rightarrow y\eta N)$$

is inductive. Clearly, $0\eta A$. If $x\eta A$, then $x\eta N$, and so $x^+\eta N$. Now if $y \leq x^+$, either $y = x^+$ or $y \leq x$, and in either case $y\eta N$. So $x^+\eta A$, and A is inductive. ■

3.8. PROPOSITION. $\mathscr{F}^* \vdash N \equiv \underline{E}_m \wedge m = |N|$.

PROOF. Arguing in $\mathscr{F}^*$, clearly $N \subseteq \underline{E}_m$ since the latter is inductive (3.6). Now let a be the <-least "element" of $M - N$ (which must exist since $\neg m\eta N$). Then, since N is an initial segment of M (3.7), we have $\underline{E}_a \equiv N$. Therefore $\text{Bij}(\underline{R}, \underline{F}^a, \underline{E}_a)$, where

$$\underline{R} =_{\text{df}} (xy)\hat{\ }[(x = a \wedge y = 0) \vee (x\eta N \wedge y = x^+)]$$

(i.e. $\underline{R}$ is "the graph of the map $a \rightsquigarrow 0, x \rightsquigarrow x^+$"). So $\underline{F}^a \approx \underline{E}_a$, whence $a^+ = a$. If $a < m$, then $a \neq a^+$, since a^+ is in this case the immediate successor of a. Accordingly $a = m$ and $N \equiv \underline{E}_a \equiv \underline{E}_m$. ■

From 3.8 and 3.4 and the fact that N satisfies Peano's axioms we immediately obtain

3.9. THEOREM. $\mathscr{F}^* \vdash M$ *has order type* $\omega + 1$. ■

REMARK. Since, as is well known, Frege's system as originally presented in the ***Grundgesetze*** was inconsistent, we should assure ourselves that $\mathscr{F}^*$ is consistent. The easiest way to see this is by noting that the following set-theoretic interpretations yield a model of the axioms of $\mathscr{F}^*$. To wit, interpret sort I as $\omega + 1$, sort II as $P(\omega + 1)$, sort III as $P((\omega + 1) \times (\omega + 1))$, sort IV as $PP(\omega + 1)$, sort V as $PP((\omega + 1) \times (\omega + 1))$, E as the subset of $PP(\omega + 1)$ consisting of all elements of the form $\|X\| =_{\text{df}} \{Y \in P(\omega + 1)$: X and Y have the same cardinality$\}$, η as the membership relation, and e as the map $PP(\omega + 1) \rightarrow \omega + 1$ which sends each $\|X\|$ with $X \in P(\omega + 1)$ to the cardinality $|X|$ of X, and everything else to $\varnothing$. Thus the axioms of $\mathscr{F}^*$ may be regarded as a consistent fragment of Frege's original system.

§4. **Ordinals and the axioms of set theory in $\mathscr{F}_e$.** It is natural to apply the Zermelo-Bourbaki lemma within $\mathscr{F}_e$ when t is the *extension term* e and $\underline{\underline{S}}$ the second level concept $X^\frown.E(X)$ corresponding to the *possession of an extension*. We obtain a well-ordering $\underline{W}$ with field W for which we can prove in $\mathscr{F}_e$ that

$(\underline{W}_1)$ $\quad \forall x[x\eta W \to E(\underline{W}_x) \wedge e(\underline{W}_x) = x],$

$(\underline{W}_2)$ $\quad E(W) \to e(W)\eta W.$

In a natural sense, W turns out to be the concept in $\mathscr{F}_e$ corresponding to the (*von Neumann*) *ordinals*. To see this, we need to make the following definitions in $\mathscr{F}_e$ (into which we introduce some obvious notational abbreviations):

$$\text{set}(x) \Leftrightarrow_{\text{df}} \exists X[E(X) \wedge e(X) = x],$$
$$x \in y \Leftrightarrow_{\text{df}} \exists Y[E(Y) \wedge e(Y) = y \wedge x\eta Y],$$
$$x \notin y \Leftrightarrow_{\text{df}} \neg x \in y,$$
$$\{x : \phi(x)\} =_{\text{df}} e(x^\frown \phi(x)),$$
$$X \subseteq x \Leftrightarrow_{\text{df}} X \subseteq y^\frown.y \in x,$$
$$\text{trans}(x) \Leftrightarrow_{\text{df}} \text{set}(x) \wedge \forall y \in x.\text{set}(y) \wedge \forall y \in x \forall z \in y.z \in x,$$
$$\in\text{WO}(X) \Leftrightarrow_{\text{df}} \forall y \in x.y \notin y \wedge \forall wyz \in x[w \in y \wedge y \in z \to w \in z]$$
$$\wedge \forall yz \in x[y \in z \vee y = z \vee z \in y]$$
$$\wedge \forall X[X \subseteq x \wedge \exists y.y\eta X \to \exists y[y\eta X \wedge \forall z[z\eta X \to y \in z]]],$$
$$\text{Ord}(x) \Leftrightarrow_{\text{df}} \text{trans}(x) \wedge \in\text{WO}(x).$$

Here set(x) asserts that x is a *set*, i.e. an object arising as the extension of a concept possessing one, $\in$ is the *membership relation*, trans(x) asserts that x is a *transitive sets of sets*, $\in$WO(X) asserts that the membership relation on x is a (strict) *well-ordering*, and Ord(x) asserts that x is an *ordinal*.

We shall need the following fact.

4.1 LEMMA. $\mathscr{F}_e \vdash \forall x[\text{set}(x) \leftrightarrow E(y^\frown.y \in x) \wedge x = \{y : y \in x\}]$.

PROOF. Argue in $\mathscr{F}_e$. Clearly the right-hand side of the equivalence implies the left-hand side. Conversely, assuming set(x), let X be such that $E(X)$ and $e(X) = X$. Then

$$y \in x \leftrightarrow \exists Y[E(Y) \wedge e(Y) = x \wedge y\eta Y] \leftrightarrow y\eta X.$$

So $y^\frown.y \in X \equiv X$, whence $E(y^\frown.y \in x)$ and $x = e(X) = e(y^\frown.y \in x) = \{y : y \in x\}$. ■

In the statement and proof of the theorem that follows, we write $x \prec y$ for $x\eta\underline{W}_y$ (equivalently, $(xy)\eta\underline{W} \wedge x \neq y$).

THEOREM 4.2. *In* $\mathscr{F}_e$ *the following are provable*:

(i) $\neg E(W)$,

(ii) $\prec$ is the membership relation on W, i.e.

$$\forall xy[x\eta W \wedge y\eta W \to [x \prec y \leftrightarrow x \in y]],$$

(iii) $W \equiv x^\frown.\text{Ord}(x)$.

PROOF. Throughout, we argue in $\mathscr{F}_e$.

(i) Assume $E(W)$. Then, writing p for $e(W)$, we have $p\eta W$ by (W_2). Hence $E(\underline{W}_p)$ and $e(\underline{W}_p) = p = e(W)$ by (W_1). It follows that $\underline{W}_p \equiv W$, so $p\eta W$ gives $p\eta \underline{W}_p$, a contradiction.

(ii) Suppose $x\eta W$ and $y\eta W$. If $x \prec y$, then, since $E(\underline{W}_y)$ and $e(\underline{W}_y) = y$ (by (W_1)), it follows that $x \in y$. Conversely, if $x \in y$, then for some Y we have $E(Y), e(Y) = y$ and $x\eta Y$. But since $y\eta W$, we have $y = e(\underline{W}_y) = e(Y)$, whence $\underline{W}_y \equiv Y$. Since $x\eta Y$, it follows that $x\eta \underline{W}_y$.

(iii) We have to show that, for any a,

$$(*) \qquad a\eta W \leftrightarrow \mathrm{Ord}(a).$$

Suppose first that $a\eta W$. We need to show that (a)trans(a), and (b) $\in$WO(a). We observe that (a) follows from $(\alpha), (\gamma), (\delta)$ below.

(α) set(a) by (W_1).

(β) $\forall x \in a.x\eta W$, since $x \in a \to x\eta \underline{W}_a \to x\eta W$.

(γ) $\forall x \in a.\mathrm{set}(x)$, since $x \in a \to x\eta W \to \mathrm{set}(x)$ by (α) and (β).

(δ) $\forall x \in a \forall y \in x.y \in a$. For if $x \in a$ and $y \in x$, then $x\eta W$ and $y\eta W$ by (β). Hence $x \prec a$ and $y \prec x$ by (ii), so $y \prec a$ since $\prec$ (i.e. $\underline{W}$) is an ordering. Thus $y \in a$ by (ii). This proves (δ), and so we have (a).

As for (b), we observe first that, if $x \in a$, then $x\eta W$ by (β) above; since $\neg(x \prec x)$ it follows that $x \notin e(\underline{W}_x) = x$. The remaining conjuncts in the definition of $\in$WO(a) follow from the fact that $\prec$ is a well-ordering which, when restricted to $x\hat{}.x \in a$, coincides with the relation induced by $\in$ there.

Before proving the reverse implication in $(*)$, note that, by the proof of the Zermelo-Bourbaki lemma,

$(**)$ *W includes the field of every well-ordering $\underline{X}$ such that*

$$\forall x[x\eta\, \mathrm{Field}(\underline{X}) \to E(\underline{X}_x) \wedge x = e(\underline{X}_x)].$$

Now suppose that Ord(a). Define

$$\underline{S} =_{\mathrm{df}} (xy)\hat{}[[x \in a \wedge y \in a \wedge [x \in y \vee x = y]] \vee [x \in a \wedge y = a] \vee x = y = a]].$$

Then $\underline{S}$ is "the $\in$-relation on $x\hat{}(x \in a \vee x = a)$", and from $\in$WO(a) it follows easily that WO($\underline{S}$). Also, if $x\eta\, \mathrm{Field}(\underline{S})$, then $x \in a$ or $x = a$, so set(x) in either case; since $\underline{S}_x = y\hat{}.y \in x$ it follows from 4.1 that $E(\underline{S}_x)$ and $e(\underline{S}_x) = \{y : y \in x\} = x$. Therefore, by $(**)$, Field$(\underline{S}) \subseteq W$. But clearly $a\eta\, \mathrm{Field}(\underline{S})$, so $a\eta W$. This proves (iii). ■

From (iii) of this theorem, we see that, as claimed, W is (extensionally equivalent to) the ordinal concept. However, it is consistent to assume, even in $\mathscr{F}^*$, that W is the *empty* concept. To see this, observe that it is easily proved in $\mathscr{F}_e$ that W is nonempty iff the empty concept $x\hat{}.x \neq x$ has an extension. But in the model of $\mathscr{F}^*$ presented at the end of §3, it is clear that the interpretation of no concept possessing an extension is empty. So "W is nonempty" fails in this model.

Accordingly, to make W nonempty we need to postulate the *empty set axiom*, which in $\mathscr{F}_e$ takes the form

$$E(x\hat{}.x \neq x).$$

Next, how can we ensure that W is *unbounded*? Of course, W is already "unbounded" in the sense that it does not possess an extension, but this, as we have seen, is consistent with W being empty. There are two more natural senses in which W could be deemed unbounded (assuming it is nonempty): (1) W has no largest "element", or (2) no subset of W is cofinal in W, i.e., for any subset a of W, there is an "element" of W larger than every element of a.

Using the fact (4.2(ii)) that the well-ordering on W is $\in$, it is easily shown that the derivability of (1) in $\mathscr{F}_e$ is ensured by postulating the *axiom of successor sets*, which here takes the form

$$\forall x[\text{set}(x) \rightarrow E(y\hat{}(y \in x \vee y = x))],$$

and that of (2) is ensured by postulating the *axiom of unions*, viz.

$$\forall x[\text{set}(x) \wedge \forall y \in x.\text{set}(y) \rightarrow E(z\hat{}(\exists y \in x.z \in y))].$$

Writing $\varnothing$ for $e(x\hat{}.x \neq x)$ and $s(x)$ for $e(y\hat{}(y \in x \vee y = x))$, it is now easily proved in $\mathscr{F}_e$ from the axioms of empty set and successor sets that the "correspondence" $x \rightsquigarrow s(x)$ establishes a bijection of W with a subconcept of $W - [\varnothing]$, so that W is infinite. This enables the natural numbers to be defined in $\mathscr{F}_e$ as *ordinals* in the customary way, viz.,

$$N^* =_{\text{df}} x\hat{}[\forall X[X \subseteq W \wedge \varnothing\eta X \wedge \forall y(y\eta X \rightarrow s(y)\eta X)] \rightarrow x\eta X].$$

Since N^* may not possess an extension, to ensure that it does we have no alternative but simply to postulate the *axiom of infinity*, viz. $E(N^*)$.

The three "set-theoretic" axioms—empty set, successor sets, and infinity—together guarantee in $\mathscr{F}_e$ that all ordinal numbers $< \omega + \omega$ exist. This is to be contrasted with the fact that the *single* additional hypothesis in $\mathscr{F}^*$—that numerical concepts possess extensions—guarantees the existence of all numbers $\leq \omega + 1$. As far as providing a foundation for *arithmetic* is concerned, assuming the latter hypothesis seems a more elegant and less *ad hoc* expedient than the postulation of piecemeal set-theoretic axioms.

§5. **Defining the extension predicate.** It is natural to ask whether there is any reasonable way of *defining* the extension predicate E within $\mathscr{F}$. One possibility is suggested by *von Neumann's maximum principle* (cf. [8, p. 288], (1)) which in $\mathscr{F}_e$ would amount to the following axiom:

(VN) $$\forall X[E(X) \leftrightarrow \neg(X \approx V)].$$

(A discussion of a similar principle may be found in [2].) As observed in [8], (VN) is an extraordinarily powerful principle. From it one immediately derives in $\mathscr{F}_e$:

- the *well-ordering principle* (2.2): V can be well-ordered. For $\neg E(W)$ and so $W \approx V$; since W is well-ordered, V can be well-ordered.

- the *axiom of separation*: $X \subseteq Y \wedge E(Y) \rightarrow E(X)$. For if $\neg E(X)$ then $X \approx V$, so if also $X \subseteq Y$ then $Y \approx V$ by the equinumerosity theorem, so $\neg E(Y)$.
- the (*weak*) *axiom of replacement*: $E(X) \wedge X \approx Y \rightarrow E(Y)$.

(In fact, note that, using the well-ordering principle, it can be shown that the counterpart of the full axiom of replacement can be derived from (VN).)

Observe that (VN), together with the "set-theoretic" axioms of the previous section, may be consistently added to $\mathscr{F}^*$. To see this, note that the following set-theoretic interpretations yield a model of the resulting theory. Let κ be a strongly inaccessible cardinal, and for each ordinal α write $R(\alpha)$ for the set of sets of rank $< \alpha$. (Then $R(\kappa+1)$ is a model of Morse-Kelley set theory.) Now interpret sort I as $R(\kappa)$, sort II as $R(\kappa+1)$, sort III as $P(R(\kappa) \times R(\kappa))$, sort IV as $R(\kappa+2)$, sort V as $PP(R(\kappa) \times R(\kappa))$, E as $R(\kappa) \cup \{\|X\| : X \in R(\kappa+1) - \{\varnothing\}\}$, where $\|X\| = \{Y \in R(\kappa+1) : X \text{ and } Y \text{ have the same cardinality}\}$, η as the membership relation, and finally e as the map defined as follows: $e|R(\kappa) =$ identity, $e(\|X\|) =$ cardinality of X if $X \neq \varnothing$ and cardinality of $X < \kappa$, $e(\|X\|) =$ some fixed element $a \in R(\kappa)$ not a cardinal if cardinality of $X = \kappa$, and $e(U) = \varnothing$ for any U not of the previous forms.

In conclusion, we note that another possible way of defining E would be to stipulate:

$$(\dagger) \qquad \forall X[E(X) \leftrightarrow \exists Y[Y \subseteq M \wedge X \approx Y]],$$

where M is the (countable) well-ordered concept obtained in the proof of Frege's theorem. This stipulation makes all (basic) concepts possessing extensions *countable* and, like (VN), can be shown to be consistent with $\mathscr{F}^*$ and all the "set-theoretical" axioms we have introduced. Unlike (VN), however, (†) is not consistent with the counterpart of the *power set axiom* in $\mathscr{F}_e$, since, by Cantor's theorem (which is easily provable in $\mathscr{F}_e$), the "power concept" of M could not be equinumerous with a subconcept of M. Nonetheless, (†) would seem to be a perfectly sound principle on which to base a theory of the countable.

REFERENCES

[1] J. L. BELL, *Hilbert's epsilon-operator and classical logic*, ***Journal of Philosophical Logic***, vol. 22 (1993), pp. 1–18.

[2] G. BOOLOS, *Saving Frege from contradiction*, ***Proceedings of the Aristotelian Society***, vol. 87 (1987), pp. 137–151.

[3] ———, *The standard of equality of numbers*, ***Meaning and method: essays in honor of Hilary Putnam*** (G. Boolos, editor), University Press, Cambridge, 1990, pp. 261–277.

[4] ———, *The advantages of honest toil over theft*, ***Mathematics and mind*** (Alexander George, editor), Oxford University Press, New York, 1994, pp. 27–44.

[5] N. BOURBAKI, ***Théorie des ensembles,*** 2nd ed., Hermann, Paris, 1963.

[6] W. DEMOPOULOS and J. L. BELL, *Frege's theory of concepts and objects and the interpretation of second order logic*, ***Philosophia Mathematica,*** ser. 3, vol. 1 (1993), pp. 139–156.

[7] G. FREGE, ***Die Grundlagen der Arithmetik,*** Breslau,1884; English translation (by J. L. Austin), ***The foundations of arithmetic***, Northwestern University Press, Evanston, Illinois, 1980.

[8] M. F. HALLETT, ***Cantorian set theory and limitation of size***, Clarendon Press, Oxford, 1984.

[9] M. D. POTTER, ***Sets: an introduction***, Clarendon Press, Oxford, 1990.

[10] E. ZERMELO, *Beweis, dass jede Menge wohlgeordnet weden kann*, **Mathematische Annalen**, vol. 59, 1904, pp. 514–516; English translation, ***From Frege to Gödel, a source book in mathematical logic,*** (J. van Heijenoort, editor), Harvard University Press, Cambridge, Massachussetts, 1967, pp. 139–141.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO N6A 3K7, CANADA

E-mail: jbell@julian.uwo.ca