Frege's Theorem in a Constructive Setting[1]

John L. Bell

By *Frege's Theorem* is meant the result, implicit in Frege's *Grundlagen*, that, for any set E, if there exists a map ν from the power set of E to E satisfying the condition

$$\forall XY[\ \nu(X) = \nu(Y) \Leftrightarrow X \approx Y]\ ^{2},$$

then E has a subset which is the domain of a model of Peano's axioms for the natural numbers. (This result is proved explicitly, using classical reasoning, in section 3 of [1].) My purpose in this note is to strengthen this result in two directions: first, the premise will be weakened so as to require only that the map ν be defined on the family of (Kuratowski) *finite* subsets of the set E, and secondly, the argument will be *constructive,* i.e., will involve no use of the law of excluded middle. To be precise, we will prove, in constructive (or intuitionistic) set theory[3], the following

Theorem. Let ν be a map with domain a family of subsets of a set E to E satisfying the following conditions:

(i) $\varnothing \in \mathrm{dom}(\nu)$

(ii) $\forall U \in \mathrm{dom}(\nu)\ \forall x \in E - U\ \ U \cup \{x\} \in \mathrm{dom}(\nu)$

(iii) $\forall UV \in \mathrm{dom}(\nu)\ \ \nu(U) = \nu(V) \Leftrightarrow U \approx V.$

Then we can define a subset N of E which is the domain of a model of Peano's axioms.

Thus, for the system of natural numbers to be constuctively obtainable, it is enough that the domain of the "cardinality" map ν contain $\varnothing$ and be closed under union with (disjoint) singletons. This condition is satisfied, in particular, when $\mathrm{dom}(\nu)$ is the family of *Kuratowski finite* subsets of the given set E, that is, the smallest family **K** of subsets of E containing the empty set and all singletons, and closed under unions of pairs of its members.

We now turn to the proof of the Theorem. This breaks down into a sequence of lemmas: we observe that in establishing these lemmas no use of the law of excluded

---

[1] *Journal of Symbolic Logic* vol. 64, no. 2, 1999.

[2] We write $X \approx Y$ for *there exists a bijection between X and Y,* and, more generally, $f\colon X \approx Y$ for *f is a bijection between X and Y*.

[3] The Theorem and its proof can also be formulated within the intuitionistic version of the first-order system of [1].

middle is made.

For $X \in \mathrm{dom}(\nu)$ write X^+ for $X \cup \{\nu(X)\}$. Call a property Φ defined on the members of dom(ν) *inductive* if $\Phi(\varnothing)$ and, for any X, if $\Phi(X)$ and $\nu(X) \notin X$, then $\Phi(X^+)$. Call a subfamily ***A*** of dom(ν) *inductive* if the property of being a member of ***A*** is inductive. Then dom(ν) is inductive, as is the intersection ***N*** of the collection of all inductive families. From the fact that ***N*** is the least inductive family we infer immediately the

Principle of Induction for ***N***. For any property Φ defined on the members of ***N***, if Φ is inductive, then every member of ***N*** has Φ.

Lemma 1. For any $X \in \boldsymbol{N}$,

$$X = \varnothing \text{ or } X = Y^+ \text{ for some } Y \in \boldsymbol{N} \text{ such that } \nu(Y) \notin Y.$$

Proof. Write $\Phi(X)$ for this assertion. To establish the claim it is enough, by the principle of induction, to show that Φ is inductive. Clearly $\Phi(\varnothing)$. If $\Phi(X)$ and $\nu(X) \notin X$, then evidently $\Phi(X^+)$. So Φ is inductive. ■

Lemma 2. For any $X \in \boldsymbol{N}$ and any $x \in X$,

$$\text{there is } Y \in \boldsymbol{N} \text{ such that } Y \subseteq X \text{ and } x = \nu(Y).$$

Proof. Writing $\Phi(X)$ for this assertion, it suffices to show that Φ is inductive. Clearly $\Phi(\varnothing)$. Now assume $\Phi(X)$ and $x \in X^+$. Then either $x \in X$, in which case, since $\Phi(X)$ has been assumed, there is $Y \in \boldsymbol{N}$ for which $x = \nu(Y)$ and $Y \subseteq X$, *a fortiori* $Y \subseteq X^+$. Or $x = \nu(X)$, yielding the same conclusion with $Y = X$. So we obtain $\Phi(X^+)$, Φ is inductive, and the Lemma follows. ■

Lemma 3. If $X, Y \subseteq E$, $x \in E - X$, $y \in E - Y$, and $X \cup \{x\} \approx Y \cup \{y\}$, then $X \approx Y$.

Proof. Assume the premises and let $f: X \cup \{x\} \approx Y \cup \{y\}$. We produce a map $f': X \approx Y$. Let y' be the unique element of $Y \cup \{y\}$ for which $\langle x,y' \rangle \in f$. Then either $y' = y$, in which case we take f' to be the restriction of f to X, or $y' \in Y$, in which case the unique element $x' \in X \cup \{x\}$ for which $\langle x',y \rangle \in f$ satisfies $x' \in X$. (For if $x' = x$ then $\langle x,y \rangle \in f$ in which case $y' = y \notin Y$.) So in this case we define

$$f' = [f \cap (X \times Y)] \cup \{\langle x',y' \rangle\}.$$

In either case it is easily checked that $f': X \approx Y$. This proves the Lemma. ■

Lemma 4. For all X, Y in $\boldsymbol{N}$,

$$\nu(X) = \nu(Y) \Rightarrow X = Y.$$

Proof. Write $\Phi(X)$ for the assertion $X \in \boldsymbol{N}$ *and* $\forall Y \in \mathbf{N}[\nu(X) = \nu(Y) \Rightarrow X = Y]$. It suffices to show that Φ is inductive. $\Phi(\varnothing)$ holds because $\nu(\varnothing) = \nu(Y) \Rightarrow Y \approx \varnothing \Rightarrow \varnothing = Y$. Now assume that $\Phi(X)$ and $\nu(X) \notin X$; we derive $\Phi(X^+)$. Suppose that $Y \in \boldsymbol{N}$ and $\nu(X^+) = \nu(Y)$. Then $X^+ \approx Y$, and so in particular $Y \neq \varnothing$. By Lemma 1, there is $Z \in \boldsymbol{N}$ for which $\nu(Z) \notin Z$ and $Y = Z^+$, so that $X^+ \approx Z^+$. We deduce, using Lemma 3, that $X \approx Z$, so, since we have assumed $\Phi(X)$, $X = Z$. Hence $X^+ = Z^+ = Y$, and $\Phi(X^+)$ follows. So Φ is inductive and the Lemma proved. ■

Lemma 5. For any $X \in \boldsymbol{N}$,

$$\nu(X) \notin X.$$

Proof. It suffices to show that the property $\nu(X) \notin X$ is inductive. Obviously $\varnothing$ has this property. Supposing that $X \in \boldsymbol{N}$, $\nu(X) \notin X$ but $\nu(X^+) \in X^+$, we have either $\nu(X^+) = \nu(X)$ or $\nu(X^+) \in X$. In the former case $X = X^+$ by Lemma 4, so that $\nu(X) \in X$, a contradiction. In the latter case, by Lemma 2, there is $Y \in \boldsymbol{N}$ such that $Y \subseteq X$ and $\nu(X^+) = \nu(Y)$. Lemma 4 now applies to yield $X^+ = Y \subseteq X$, so again $\nu(X) \in X$, a contradiction. Therefore $\nu(X) \notin X \Rightarrow \nu(X^+) \notin X^+$, and the Lemma follows. ■

Notice that it follows immediately from Lemma 5 that $\boldsymbol{N}$ *is closed under* +, that is, $X \in \boldsymbol{N} \Rightarrow X^+ \in \boldsymbol{N}$.

Now define $0 = \nu(\varnothing)$, $N = \{\nu(X): X \in \boldsymbol{N}\}$, and $s: N \to N$ by $s(\nu(X)) = \nu(X^+)$ for $X \in \boldsymbol{N}$. Then s is well defined and injective on N. (For if $\nu(X) = \nu(Y)$, then, by Lemma 4, $X = Y$, and so $s(\nu(X)) = \nu(X^+) = \nu(Y^+) = s(\nu(Y))$. Conversely, if $s(\nu(X)) = s(\nu(Y))$, then $\nu(X^+) = \nu(Y^+)$, so that, by Lemma 4, $X^+ \approx Y^+$. Lemmas 3 and 5 now imply $X \approx Y$, whence $\nu(X) = \nu(Y)$.) Clearly, also, $0 \neq sn$ for any $n \in N$. The fact that the structure $(N, s, 0)$ satisfies the principle of induction follows immediately from the principle of induction for $\boldsymbol{N}$. Accordingly $(N, s, 0)$ is a model of Peano's axioms, as required.

Remarks. 1. Since the arguments given here are constructive, they may be translated into the internal language of an arbitrary topos, so that the Theorem holds in arbitrary toposes also.

2. The *Zermelo-Bourbaki Lemma* (Lemma 2.1 of [1]) may also be used to give a *nonconstructive* proof of the Theorem. In its set-theoretic form, the Zermelo-Bourbaki

lemma states that, given a map p from a family of subsets of a set E to E such that $p(X) \notin X$ for any $X \in \mathrm{dom}(p)$, there is a subset M of E and a well-ordering $\leq$ of M, such that, writing S_x for $\{y: y < x\}$, (i) $\forall x \in M.\ S_x \in \mathrm{dom}(p)$ and $p(S_x) = x$; (ii) $M \notin \mathrm{dom}(p)$. If we assume the premises of the Theorem and apply the Zermelo-Bourbaki lemma to the set $\{X \in \mathrm{dom}(\nu): \nu(X) \notin X\}$, taking p to be the restriction of ν to this set, we get a well-ordered subset M of E for which $M \notin \mathrm{dom}(p)$, which means that either $M \notin \mathrm{dom}(\nu)$ or $\nu(M) \in M$. In the latter case we may quickly argue as in the proof of 3.1 of [1] to conclude that M is Dedekind infinite, and so yields a model of the Peano axioms. In the former case, we deduce from the properties of $\mathrm{dom}(\nu)$ that the well-ordered set M has no last element and is therefore infinite, again yielding a model of the Peano axioms. It should be noted, however, that the Zermelo-Bourbaki lemma, asserting as it does the existence of well-orderings, is irremediably nonconstructive, since, as is well-known, the existence of a well-ordering on even a two-element set implies the law of excluded middle.

Reference

[1] J. L. Bell. *Type reducing correspondences and well-orderings: Frege's and Zermelo's constructions re-examined.* Journal of Symbolic Logic, vol.60. no.1, March 1995, 209-220.