Frege's Proofs of the Axioms of Arithmetic

Richard G. Heck, Jr.

1 The Dedekind-Peano Axioms for Arithmetic

1. $\mathbb{N}0$
2. $\forall x(\mathbb{N}x \to \exists y.Pxy)$
3(a). $\forall x\forall y\forall z(\mathbb{N}x \land Pxy \land Pxy \to y = z)$
3(b). $\forall x\forall y\forall z(\mathbb{N}x \land \mathbb{N}y \land Pxz \land Pyz \to x = y)$
4. $\neg\exists z(\mathbb{N}z \land Pz0)$
5. $\forall F[F0 \land \forall x\forall y(\mathbb{N}x \land Fx \land Pxy \to Fy) \to \forall n(\mathbb{N}n \to Fn)]$

2 Hume's Principle and Fregean Arithmetic

Hume's Principle:

$$Nx : Fx = Nx : Gx \equiv$$
$$\exists R[\forall x\forall y\forall z(Rxy \land Rxz \to y = z) \land \forall x\forall y\forall z(Rxz \land Ryz \to x = y)\land$$
$$\forall x(Fx \to \exists y(Gy \land Rxy)) \land \forall y(Gy \to \exists x(Fx \land Rxy))]$$

Converse of a relation: $Conv_{\alpha\epsilon}(R\alpha\epsilon, x, y) \equiv Ryx$

Functionality: $Func_{\alpha\epsilon}(R\alpha\epsilon) \equiv \forall x\forall y(Rxy \to \forall z(Rxz \to y = z))$

Frege Arithmetic is the axiomatic theory consisting of standard second-order logic, with Hume's Principle the sole 'non-logical' axiom

3 Frege's Definitions of Arithmetic Notions

Zero: $0 = Nx : x \neq x$

Predecession: $Pmn \equiv \exists F\exists x[Fx \land n = Nz : Fz \land m = Nz : (Fz \land z \neq x)]$

The strong ancestral: $Q^*ab \equiv \forall F[\forall x\forall y(Fx \land Qxy \to Fy) \land \forall x(Qax \to Fx) \to Fb]$

The weak ancestral: $Q^{*=}ab \equiv Q^*ab \lor a = b$

Natural number: $\mathbb{N}x \equiv P^{*=}0x$

4 Frege's Proofs

4.1 Axiom 1

We must show that $\mathbb{N}0$, i.e., that 0 is a natural number. According to Frege's definition, this follows from the trivial fact that $P^{*=}00$.

4.2 Axiom 4

We can in fact prove the stronger claim that:

$$\neg\exists z(Pz0)$$

Suppose, for *reductio*, that $Pz0$. By Frege's definition, we therefore have that, for some *F* and *y*,

$$Nx : Fx = 0 \wedge Fy \wedge Nx : (Fx \wedge x \neq y) = z$$

A fortiori, for some *F* and *y*, $Nx : Fx = 0 \wedge Fy$. But by a result of *Die Grundlagen* §75, if the number of *F*s is zero, then nothing is *F*; so we have a contradiction.

4.3 Frege's Proof of Axiom 5

Axiom 5 follows easily from the following more general result, Theorem 152 of Frege's *Grundgesetze der Arithmetik* (*Basic Laws of Arithmetic*):

$$Q^{*=}an \wedge Fa \wedge \forall x\forall y(Q^{*=}ax \wedge Fx \wedge Qxy \to Fy] \to Fn$$

To prove Axiom 5 from this, generalize '*F*'; substitute '0' for '*a*' and '*P*' for '*Q*', and then apply the definition of '$\mathbb{N}$'. Axiom 5 will then follow by truth-functional logic.

Theorem 152 can easily be derived from the following, which is Theorem 144 of *Grundgesetze*:

$$Q^{*=}an \wedge Fa \wedge \forall x\forall y(Fx \wedge Qxy \to Fy] \to Fn$$

Simply substitute '$Q^{*=}a\xi \wedge F\xi$' for '$F\xi$' in (144) and chug away. The proof mostly consists in applying what I below call the 'basic facts about the ancestral'.

Theorem 144, in turn, may be derived from the following, which is Theorem 128 of *Grundgesetze*, which is also Theorem 81 of *Begriffsschrift*:

$$Q^{*}an \wedge Fa \wedge \forall x\forall y(Fx \wedge Qxy \to Fy) \to Fn$$

Use the definition of the weak ancestral and then argue by dilemma.

Theorem 128, in turn, is an easy consequence of Theorem 123 of *Grundgesetze*, which is Theorem 77 of *Begriffsschrift*:

$$Q^{*}an \wedge \forall x(Qax \to Fx) \wedge \forall x\forall y(Fx \wedge Qxy \to Fy) \to Fn$$

For, if Fa and $\forall x\forall y(Fx \wedge Qxy \to Fy)$, then certainly $\forall x(Qax \to Fx)$.

Finally, Theorem 123 is immediate from the definition of the strong ancestral.

4.4 Frege's Proof of Axiom 3(a)

We can in fact prove the stronger:

$$Pwy \wedge Pwz \rightarrow y = z$$

Suppose the antecedent. By the definition of '*P*', we have that, for some *F* and for some *G*, for some *a* and for some *b*:

$$y = Nx : Fx \wedge Fa \wedge w = Nx : (Fx \wedge x \neq a)$$
$$z = Nx : Gx \wedge Gb \wedge w = Nx : (Gx \wedge x \neq b)$$

We thus have that $Fa \wedge Gb \wedge Nx : (Fx \wedge x \neq a) = Nx : (Gx \wedge x \neq b)$ and we need to show that $Nx : Fx = Nx : Gx$.

Since $Nx : (Fx \wedge x \neq a) = Nx : (Gx \wedge x \neq b)$, there is a relation that correlates the *F*s other than *a* one-one with the *G*s other than *b*. That is, for some *R*,

$$\begin{gathered}
\forall x \forall y \forall z (Rxy \wedge Rxz \rightarrow y = z) \wedge \\
\forall x \forall y \forall z (Rxz \wedge Ryz \rightarrow x = y) \wedge \\
\forall x [(Fx \wedge x \neq a) \rightarrow \exists y ((Gy \wedge y \neq b) \wedge Rxy)] \wedge \\
\forall y [(Gy \wedge y \neq b) \rightarrow \exists x ((Fx \wedge x \neq a) \wedge Rxy)]
\end{gathered}$$

We must show, on the assumption that *a* is *F* and *b* is *G*, that some relation also correlates the *F*s one-one with the *G*s. So consider this relation:

$$[R\xi\eta \wedge (\xi \neq a \wedge \eta \neq b)] \vee [\xi = a \wedge \eta = b]$$

Call this relation $Q\xi\eta$. (That there is such a relation follows from the comprehension axiom of second-order logic.)

Claim: $Q\xi\eta$ correlates the *F*s one-one with the *G*s.

Proof of claim: We first show that *Q* is one-one.

Suppose that *Qxy* and *Qxz*; we must show that $y = z$. Either $x = a$ or not. If not, then, by the definition of '*Q*', we must have that *Rxy* and *Rxz*, since the other disjunct cannot hold. But *R* is one-one, so $y = z$. And if $x = a$, then it must be the second disjunct that holds, so $y = b$ and $z = b$, whence $y = z$, again.

Suppose that *Qxz* and *Qyz*. Again, either $z = b$ or not. If not, then *Rxz* and *Ryz*, whence $x = y$, since *R* is one-one. And if $z = b$, then, as before, $x = a = y$.

We now show that *Q* correlates the *F*s with the *G*s. Suppose that *Fx*. We must show that, for some *y*, *Gy* and *Qxy*. Now, either $x = a$ or not. If so, then *Qab*, by the second disjunct; we know that *Gb*, so we are done. If $x \neq a$, then $Fx \wedge x \neq a$, whence for some *y*, $Gy \wedge y \neq b \wedge Rxy$, since *R* correlates the *F*s other than *a* with the *G*s other than *b*. But then *Qxy*, by the definition of '*Q*', since the first disjunct holds. Similarly, if *Gy*, then either $z = b$ or not. If so, then *Qab*, by the second disjunct, and since *Fa*, we are done. If $z \neq b$, then $Gy \wedge y \neq b$, whence, for some *x*, $Fx \wedge x \neq a \wedge Rxy$ and so Qxy.

4.5 Frege's Proof of Axiom 3(b)

We can again prove the stronger:

$$Pwz \wedge Pyz \rightarrow w = y$$

Again, by the definition of '*P*', we have that, for some *F* and *G*, and for some *a* and *b*:

$$z = Nx : Fx \wedge Fa \wedge w = Nx : (Fx \wedge x \neq a)$$
$$z = Nx : Gx \wedge Gb \wedge y = Nx : (Gx \wedge x \neq b)$$

So $Fa \wedge Gb \wedge Nx : Fx = Nx : Gx$ and and must show that $Nx : (Fx \wedge x \neq a) = Nx : (Gx \wedge x \neq b)$. Let *R* be a one-one correspondence between the *F*s and the *G*s. Since *Fa*, there is some object *m* such that *Ram* and *Gm*; since *Gb*, there is some object *n* such that *Rnb* and *Fn*. Consider the relation:

$$[R\xi\eta \wedge \xi \neq a \wedge \eta \neq b] \vee [x = n \wedge y = m]$$

Call this relation $Q\xi\eta$. (That there is such a relation follows from the comprehension axiom of second-order logic.)

Claim: $Q\xi\eta$ correlates the *F*s other than *a* one-one with the *G*s other than *b*.

Proof of claim: Exercise: *Very* similar to the preceding proof of 3(a).

5 Frege's Proof of Axiom 2

Using our notation, and adding a few indices, we may abbreviate the relevant portions of §§82–3 of *Grundlagen* as follows:

> §82. It is now to be shown that—subject to a condition still to be specified—
>
> $$(0)\ P(n, Nx : P^{*=}xn)$$
>
> And in thus proving that there exists a Number which follows in the series of natural numbers directly after *n*, we shall have proved at the same time that there is no last member of this seris. Obviously, this proposition cannot be established on empirical lines or by induction.
>
> To give the proof in full here would take us too far afield. I can only indicate briefly the way it goes. It is to be proved that
>
> $$(1)\ Pda \wedge P(d, Nx : P^{*=}xd) \rightarrow P(a, Nx : P^{*=}a)$$
>
> It is then to be proved, secondly, that
>
> $$(2)\ P(0, Nx : xP^{*=}0)$$
>
> And finally, (0) is to be deduced. The argument here is an application of the definition of [the strong ancestral], taking the relevant concept to be $P(\xi, Nx : xP^{*=}\xi)$.

> §83. In order to prove (1), we must show that
>
> $$(3)\ a = Nx : P^{*=}xa$$
>
> And for this, again, it is necessary to prove that
>
> $$(4)\ \forall x[(P^{*=}xa \wedge x \neq a) \equiv P^{*=}xd]$$
>
> For this we need
>
> $$(5)\ \forall x(P^{*=}0x \rightarrow \neg P^{*}xx)$$
>
> And this must once again be proved by mean of our definition of [the strong ancestral]. . . .
>
> It is this that obliges us to attach a condition to the proposition that $P(n, Nx : xP^{*=}n)$—the condition, namely, that $P^{*=}0n$. For this there is a convenient abbreviation, which I define as follows: the proposition $P^{*=}0n$ is to mean that same as '*n* is a finite Number'. We can thus formulate the last proposition above as follows: no finite Number follows in the natural series of numbers after itself.

In fact, though Frege does not make this as explict as one might like, not only does the 'condition' that *n* must be a natural number need to be added to (0); corresponding conditions must also be added to other propositions. This fact makes following his proof somewhat difficult to follow.[1]

Following George Boolos, let us call the following the *basic facts* about the ancestral:

$$Q^{*=}ab \wedge Qbc \rightarrow Q^{*}ac$$
$$Qab \wedge Q^{*=}bc \rightarrow Q^{*}ac$$

Proofs: Exercise.

We shall also make frequent use of the following, which are Theorems 102 and 103 of *Grundgesetze*:

$$(102)\ m = Nx : (Fx \wedge x \neq y) \wedge Fy \rightarrow P(m, Nx : Fx)$$
$$(103)\ Fy \rightarrow P(Nx : (Fx \wedge x \neq y), Nx : Fx)$$

Proofs: Suppose the antecedent of (102). Then we have

$$m = Nx : (Fx \wedge x \neq y) \wedge Fy \wedge Nx : Fx = Nx : Fx$$

Apply the definition of '*P*'. That proves (102), and (103) then follows by the laws of identity.

Frege's goal in §§82–3 of *Die Grundlagen* is to prove

$$(0')\ P^{*=}0n \rightarrow P(n, Nx : P^{*=}xn)$$

(This is Theorem 155 of *Grundgesetze*.) Frege writes that $(0')$ is to be derived from (1) and (2) by means of a "method of inference" which amounts to an application of the definition of the ancestral, i.e., by mathematical induction. We must thus prove the basis case

[1] There are several issues in this vicinity about how Frege's proof is to be read. See George Boolos and Richard Heck, "*Die Grundlagen der Arithmetik* §§82–83", in George Boolos, *Logic, Logic, and Logic*, ed. by Richard Jeffrey (Cambridge MA: Harvard University Press, 1998), pp. 315–38.

$$(2)\ P(0, Nx : P^{*=}x0)$$

(which is Theorem 154 of *Grundgesetze*) and the induction step

$$(1')\ P^{*=}0d \wedge Pda \wedge P(d, Nx : P^{*=}xd) \rightarrow P(a, Nx : P^{*=}xa)$$

($1'$) is Frege's proposition (1), with the additional condition that *d* is a natural number. (The universal generalization of ($1'$), with respect to '*d*' and '*a*', is Theorem 150 of *Grundgesetze*.)

Frege says that "in order to prove" ($1'$), we must prove

$$(3)\ a = Nx : (P^{*=}xa \wedge x \neq a)$$

In fact, however, (3) is not provable. At most, we can prove ($3'$), which adds the condition that *a* is a natural number:

$$(3')\ P^{*=}0a \rightarrow a = Nx : (P^{*=}xa \wedge x \neq a)$$

It is natural to suppose, and it was for a long time almost universally was supposed, that ($3'$) is what Frege intends us to prove. And if one makes this assumption, then it is possible to prove ($1'$) from ($3'$) and then to conclude ($0'$) as sketched. However, Boolos observed that Frege's proof, so reconstructed, suffers from redundancy. Substituting '$P^{*=}\xi a$' for '$F\xi$' and '*a*' for '*y*' in (103), we have:

$$P^{*=}aa \rightarrow P(Nx : (P^{*=}xa \wedge x \neq a), Nx : P^{*=}xa)$$

But $P^{*=}aa$, trivially; and by ($3'$), if $P^{*=}0a$, then $a = Nx : (P^{*=}xa \wedge x \neq a)$. So

$$P^{*=}0a \rightarrow P(a, Nx : P^{*=}xa)$$

And except for the change of variable, that just is ($0'$): Neither (2) nor ($1'$) is needed.

Frege says that we must prove (3) in order to prove (1). Note, however, that the variable '*a*' that occurs in (3) is the same variable that occurs in the consequent of (1). Surely, what Frege intends is not that we prove ($3'$) outright but that we prove ($3'$) by deriving (3) *from the antecedent* of ($1'$) and then deriving the consequent of ($1'$) from (3). That is: Frege intends us to prove something like

$$(3\dagger)\ P^{*=}0d \wedge Pda \wedge P(d, Nx : P^{*=}xd) \rightarrow a = Nx : (P^{*=}xa \wedge x \neq a)$$

If we so read the proof, then, not only is it elegant, it is, except for a minor permutation of steps, essentially the proof later given formally in *Grundgesetze*.

Suppose we could prove $(3\dagger)$. We can then derive ($1'$) from it as follows. Recall that ($1'$) is:

$$(1')\ P^{*=}0d \wedge Pda \wedge P(d, Nx : P^{*=}xd) \rightarrow P(a, Nx : P^{*=}xa)$$

Suppose the antecedent. The antecedent of $(3\dagger)$ therefore holds, so $a = Nx : (P^{*=}xa \wedge x \neq a)$. But (102) yields:

$$a = Nx : (P^{*=}xa \wedge x \neq a) \wedge P^{*=}aa \rightarrow P(a, Nx : P^{*=}xa)$$

But $P^{*=}aa$, trivially, so $P(a, Nx : P^{*=}xa)$.

Now, Frege says that we are to derive (3)—i.e., $(3\dagger)$—from

$$(4)\ \forall x(P^{*=}xa \wedge x \neq a \equiv P^{*=}xd)$$

But he can not have meant us to prove this outright, either. Unless some relationship between *d* and *a* is assumed, nothing like (4) can possbily be proven. The assumption, made in $(1')$, that *Pda*, must therefore still be in force. Moreover,

$$(4')\ P^{*=}0d \wedge Pda \rightarrow \forall x(P^{*=}xa \wedge x \neq a \equiv P^{*=}xd)$$

cannot be proven without the assumption that *d* is a natural number, so it must be $(4')$ that Frege intends us to prove. (And, as it happens, $(4')$ is Theorem 149α of *Grundgesetze*.)

We now derive (3†) from $(4')$. Suppose that $P^{*=}0d$, that Pda, and that $P(d, Nx : P^{*=}xd)$. We must show that $a = Nx : (P^{*=}xa \wedge x \neq a)$. Since *P* is one-one, we have that

$$a = Nx : P^{*=}xd$$

But the antecedent of $(4')$ holds, so

$$\forall x(P^{*=}xa \wedge x \neq a \equiv P^{*=}xd)$$

whence, by Hume's Principle,[2]

$$Nx : (P^{*=}xa \wedge x \neq a) = Nx : P^{*=}xd$$

But then, by identity,

$$a = Nx : (P^{*=}xa \wedge x \neq a)$$

which was our target.

Now, $(4')$ is to be derived from Frege's

$$(5)\ \forall x[P^{*=}0x \rightarrow \neg P^{*}xx]$$

(Note that (5) is Theorem 145 of *Grundgesetze*.) For the proof of $(4')$, we also need the following general fact about the ancestral, which is Theorem 141 of *Grundgesetze*:

$$(141)\ P^{*}xy \rightarrow \exists x(Pzy \rightarrow P^{*=}xz)$$

(The proof is left as an exercise.) Now, recall that (4') is

$$P^{*=}0d \wedge Pda \rightarrow \forall x(P^{*=}xa \wedge x \neq a \equiv P^{*=}xd)$$

So suppose the antecedent. Left-to-right: Suppose that $P^{*=}xa$ and that $x \neq a$. Then $P^{*}xa$. By (141), for some *z*, Pza and $P^{*=}xz$. Since '*P*' is one-one, $z = d$; but then $P^{*=}xd$. Right-to left: Suppose that $P^{*=}xd$. Since Pda, we have that $P^{*}xa$, by the basic facts, whence certainly $P^{*=}xa$. And if $x = a$, then $P^{*}aa$. But since $P^{*=}0d$ and Pda, $P^{*=}0a$, which contradicts (5); so $x \neq a$.

The proof of

$$(5)\ \forall x[P^{*=}0x \rightarrow \neg P^{*}xx]$$

[2] Note that, in fact, only a very weak consequence of Hume's Principle is needed here, namely:

$$\forall x(Fx \equiv Gx) \rightarrow Nx : Fx = Nx : Gx$$

Boolos suggests that this principle should be regarded as logical.

is straightforward, given (141). We proceed by mathematical induction, so we must prove that

$$\neg P^*00$$

and that

$$P^{*=}0x \wedge \neg P^*xx \wedge Pxy \rightarrow \neg P^*yy$$

Suppose for *reductio* that P^*00. By (141), then, for some *z*, $P^{*=}z0$ and $Pz0$. But this contradicts Axiom 4, which states that $\neg\exists z.Pz0$.

Now suppose that $P^{*=}0x$, that $\neg P^*xx$, and that Pxy. For *reductio*, suppose that P^*yy. By (141), for some *z*, Pzy and $P^{*=}yz$. Since Pxy and Pzy, $x = z$, by Axiom 3(b). So $P^{*=}yx$. But then Pxy and $P^{*=}yx$, so P^*xx, by the basic facts, contradicting our hypotheses.

That completes the proof of $(1')$.

Finally, we complete the proof $(0')$ by proving (2). Recall that (2) is

$$P(0, Nx : P^{*=}x0)$$

Proof: Since $P^{*=}00$, trivially, (103) delivers:

$$P(Nx : (P^{*=}x0 \wedge x \neq 0), Nx : P^{*=}x0)$$

Now, suppose that P^*x0. Then, by (141), for some *z*, $P^{*=}0z$ and $Pz0$. But that again contradicts Axiom 2, so $\neg\exists x.P^*x0$. Hence, if $P^{*=}x0$, then $x = 0$, for the other alternative, P^*x0, is impossible. That is, $\neg\exists x[P^{*=}x0 \wedge x \neq 0]$, and so, by a result of *Die Grundlagen* §75, $0 = Nx : (P^{*=}x0 \wedge x \neq 0)$, and (2) follows by identity.

That completes Frege's proof of Axiom 2 and so of the Dedekind-Peano axioms for arithmetic.