Keeping Basic Law V

Nathan Bice
nmb2138@columbia.edu

March 23, 2021

Gottlob Frege's *Grundgesetze der Arithmetik* (Basic Laws of Arithmetic, 1893/1903) is a foundational work in Mathematical Logic. Unfortunately, Bertrand Russell proved that it is inconsistent. Most readers blame the inconsistency on a particular axiom: Basic Law V. However, I will show that we can make a syntactic restriction on the system and keep Basic Law V while developing arithmetic in a consistent way. This will require adding additional axioms asserting the existence of various functions. This approach has historical and philosophical advantages and furthermore has mathematical interest. Historically, one can work through the *Grundgesetze* in a consistent way, using Frege's own definitions and proofs. Philosophically, one can define numbers in such a way that one captures the plausible thesis that a statement of number contains an assertion about a property. Finally, any natural set-theoretic model of this system will be radically *non-well-founded*, including sets that contain themselves. This results in a system quite different (mathematically speaking) than standard contemporary systems of mathematical logic and arithmetic.

Contents

1 Introduction

The goal of this project is to develop a consistent version of Gottlob Frege's *Grundgesetze der Arithmetik* (1893/1903) while maintaining his Basic Law V. This axiom is generally blamed for the contradiction in the *Grundgesetze* resulting from a version of Russell's Paradox, but the axiom only leads to a contradiction *in conjunction with* other aspects of the formal system.

I will show that we can give up other aspects of the formal system while maintaining Basic Law V in a way that still allows us to construct the natural numbers and prove the Dedekind-Peano axioms of arithmetic. The guiding idea is to put a syntactic restriction on which sequences of symbols count as well-formed formulas within the formal system. The cost will be the need to add axioms governing the existence of functions, i.e. axioms stating what is effectively a form of restricted comprehension.

However, if we change the syntax, then in what sense are we really keeping Basic Law V?[1] Basic Law V is the following: two functions have the same value-range if and only if they have the same outputs on the same inputs. It follows that every function has a value-range. This leads to a contradiction in Frege's system via Russell's Paradox.

So, I take the core idea to be that every function has a value-range. As a corollary in Frege's system, every property has an extension. These are the principles that we can maintain. The trouble with the *Grundgesetze*, then, is that not every predicate denotes a function, on pain of contradiction. The language is in this sense is "too rich". By restricting the syntax, we can avoid this unfortunate situation. Our additional axioms will then allow definitions of functions via predicates, albeit with certain constraints.

As further motivation, consider the following approach to developing a simple second-order logic from a classical first-order logic: one might think one could simply allow well-formed formulas to contain predicate symbols and predicate variables in term position, and introduce a new quantifier governing the predicate variables. This would allow one to form sentences such as $P(P)$, which will be interpreted by a model as: "property P has property P". Indeed, properties such as "being a property" are intuitively properties of themselves.

One might then think one could introduce a predicate symbol in the following way: $P(X) := \neg X(X)$. Intuitively, $P(X)$ refers to the property: "is not a property of itself". Now assume $P(P)$. It follows that $\neg P(P)$ by definition, a contradiction. Therefore $\neg P(P)$. But this implies $P(P)$, again a contradiction. But then our second-order logic is inconsistent. This is effectively Russell's paradox applied to properties.

The reason this isn't a problem for our second-order logic is that in any standard syntax one can't simply introduce a predicate symbol using the metalanguage in this way (which formula in the object language does $P(P)$ correspond to?). Instead, one would define the predicate symbol P via a sentence of the *object* language, namely: $\forall X(P(X) \leftrightarrow \neg X(X))$.

[1] I am grateful to Richard Heck and Haim Gaifman for raising this question.

From this sentence one can derive a contradiction. But this doesn't show that our second-order logic is inconsistent, for the contradiction only results from this sentence treated as an assumption. Instead, we can conclude that for any predicate symbol P, the corresponding sentence is a logical falsehood.

The point is: the syntax of our second-order logic is *restricted* in a way that prevents the construction of such paradoxical predicates.

Frege's formal system does not allow sentences of the form $P(P(x))$. First-order functions ("first-level functions") take objects as input, second-order functions take first-order functions as input, etc.[2] However, given Frege's use of value-ranges of functions, although one cannot ascribe a property to itself, one *can* ascribe a property to its own extension, and as a result the self-referential behavior that leads to versions of Russell's paradox resurfaces in Frege's system.[3]

Note also that Frege's formal system *is* effectively a system in which, for example, arbitrary well-formed formulas with one free object variable can be inserted into first-order function variable position. So, where $\mathcal{D}$ is a second-order predicate symbol, the following will be a well-formed formula (translated into a modern notation): $\mathcal{D}(P(x) \to Q(x))$.[4] In this case, $P(x) \to Q(x)$ refers to a function from objects to truth-values, and it is that function that will be mapped by the referent of $\mathcal{D}(\varphi(x))$ to the truth-value of the complete well-formed formula.

I claim that if we give this syntactic assumption up, we can salvage much of Frege's formal system. In particular, we can insist that only predicate symbols, function symbols, terms, and variables can be inserted into variable position in a predicate or functional expression. Once we make this syntactic restriction, in order to capture the content of our example sentence one would first need to assume a definition statement $\forall x(S(x) \leftrightarrow (P(x) \to Q(x)))$, where S is a (presumably unused) predicate symbol in the language. One can then capture the desired content using the sentence $\mathcal{D}(S(x))$. But now our definition statement has been treated as an assumption, and hence any contradiction we derive is a contradiction *relative to that assumption*. We are effectively assuming that there in fact exists a function that we will refer to with the expression $S(x)$ that behaves in the way we have attempted to define, mapping an object to the True if either the function referred to by $P(x)$ maps it to the False or the function referred to by $Q(x)$ maps it to the True.

[2] I should note that in Frege's system, a (first-order) predicate symbol would refer to a function from objects to truth-values.

[3] Note that an *extension* is Frege's analogue of a set of objects and a *value-range* is Frege's analogue of a set of ordered pairs. This analogy isn't perfect because for Frege e.g. binary relations have extensions as well. An extension is the value-range of a function whose range includes at most the two truth-values (compare a characteristic function), and it includes every entity that function maps to the truth-value True and no others.

[4] Strictly speaking, Frege has no need for distinguishing predicate symbols from function symbols in general, and he does not do so in his object language. However, when describing the formal system in his metalanguage (i.e. German with a bit of mathematical symbolism), Frege uses symbols that are obviously meant to correspond to functions from objects to truth-values.

I will show that this syntactic restriction is enough to ensure the relative consistency of our modification of the *Grundgesetze* while maintaining Basic Law V. In particular, I will show that this restriction, even with the addition of axioms governing arithmetic, has a (non-well-founded) model.

Our additional axioms effectively govern the uses of comprehension Frege actually needs in order to construct arithmetic using his definitions and proofs.

What is the value of this approach?

One primary value is *historical*. Using my method, one can effectively read through the *Grundgesetze* in a consistent way, using Frege's own definitions and carrying out Frege's own proofs.

Another value is *philosophical*. One can adopt this restriction while still using Frege's definition of natural number. The guiding idea behind Frege's definition is that each natural number $n > 0$ is the number such that for there to be n Ps is for there to exist distinct $x_1, ...x_n$ that all have property P such that for any y, if y has property P then y is equal to x_i for some $1 \leq i \leq n$.[5]

This is philosophically valuable because it captures the philosophical thesis that a statement of number contains an assertion about a property.[6] Hence, we can explicitly axiomatize the fundamental relation between numbers and properties.

Are there *mathematical* advantages to this approach? It is the opinion of this author that this remains to be seen. This approach is certainly an alternative to standard contemporary conceptions of higher-order logic and formal theories of arithmetic, but any alternative to standard approaches must prove its mathematical value via fruitful results. Furthermore, it may turn out that natural extensions of this theory are equivalent to theories already known.

It *is*, however, mathematically interesting that any natural interpretation of Frege's system will be *non-well-founded*, allowing e.g. extensions that contain themselves.[7] This is thus a non-standard approach to logic and mathematics more generally.

Do these results vindicate Logicism with respect to arithmetic? This would depend upon whether value-ranges are *logical* objects and whether our axioms can be justified as a matter of pure logic. I remain unconvinced.

2 A Note on Syntax and Semantics

I will be modifying the *Grundgesetze* via more contemporary approaches to syntax and semantics, leading to a formal system more appropriate for contemporary logic and mathematics. These modifications will not effect one's ability

[5]For there to be 0 Ps is for there to not exist an x such that x has property P.

[6]Frege defended the similar thesis that a statement of number contains an assertion about a *concept* in (Frege 1884), but this was prior to his sense/reference distinction and hence it is unclear whether he would have endorsed the thesis as I have just stated it. After his sense/reference distinction, Frege did use the expression 'concept' ('Begriff') to refer to the referent of a predicate, so it is natural to assume that he would have endorsed a similar thesis.

[7]A simple example is the property "being a value-range." The extension of this property, itself a value-range by definition, contains itself.

to actually read through Frege's development of arithmetic in the *Grundgesetze* in a consistent way, since the reader could reverse them if desired.

For example, I will *not* be using Frege's own two-dimensional syntax, but will rather translate Frege's syntax into a modern notation. For readers familiar with Frege's syntax, reversing the translation should be a simple exercise. As a result, I must make a few points about the differences between our syntax and Frege's syntax.

Much of Frege's notation can be easily translated into standard contemporary notation, with a few exceptions. Frege introduces a function symbol he calls the horizontal, which refers to a function from the True to the True and anything else to the False. Hence it effectively refers to the property "being the True". When combined with an arbitrary function-symbol, this has the effect of creating a predicate expression or a relation expression, referring to a function mapping its arguments to truth-values.

Frege also has a judgment stroke, meant to inform the reader that the expression following the judgment stroke has been judged to be true. Note also that in Frege's system, only expressions referring to truth-values can be judged. Frege enforces this by insisting that every well-formed formula begins with the horizontal.

Irrespective of Frege's philosophical motivation, for our mathematical purposes both the horizontal and the judgment stroke are unnecessary. We will distinguish predicate symbols and relation symbols from other function symbols, and well-formed formulas will always have a truth-value relative to an interpretation and an assignment. When carrying out a derivation within the formal system, it should be clear from the context which formulas are being asserted as true and which formulas are being treated as assumptions. Within the formal system itself, we have no need to allow for the case of a formula merely being considered rather than assumed or asserted.

Note that strictly speaking, Frege's formal language does *not* satisfy unique readability: there are well-formed formulas within the language that could be constructed in distinct ways. For example, the formula '$—f(a)$' can be read as the application of the horizontal to the result of applying the function-name '$f(\xi)$' to the object-name 'a', but it can also be read as the application of the horizontal to the result of applying the two-place function-name '$\phi(\xi)$' to the function-name '$f(\xi)$' and the object-name 'a', where this two-place function-name refers to the result of applying the referent of its first argument to the referent of its second argument. (Frege 1893, pg. 39)

I will follow the contemporary practice of constructing a formal language that satisfies unique readability.

As stated in the Introduction, our fundamental syntactic departure from Frege will be that we will insist that higher-order function and relation symbols can only be applied to function and relation symbols rather than arbitrary expressions containing the appropriate number (and type) of free variables. Hence, $\mathcal{D}(S(x))$ will be a well-formed formula while $\mathcal{D}(P(x) \to Q(x))$ will not.[8] This

[8]For simplicity, I will consider predicate symbols and properties to be relation symbols and

implies that when using the formal system, one must be careful when introducing abbreviations.

Given this syntactic restriction, Frege's axioms for identity, which quantify over arbitrary functions, ought to be replaced with axioms involving substitution, as is the contemporary custom. However, I will not be using all of Frege's own axioms in the derivation system below, preferring to extend a standard contemporary derivation system with Frege's axioms that go beyond it.[9]

Frege's syntax also goes beyond standard contemporary syntax by including names of value-ranges and a definite article (the analogue of 'the' in English). I will use corner-quotes to construct names of value-ranges: e.g. $\ulcorner v_1, v_2 \urcorner ({f_i}^{1,2}(v_1, v_2))$ refers to the value-range of a first-order function of two variables. it will turn out that only value-ranges of first-order functions and predicates are needed for our purposes. Following Frege, I will also let the symbol $\imath$ correspond to a definite article.

Frege's formal language is fully *interpreted*: every expression within the language has a fixed interpretation. Effectively, there is a single domain of objects and functions which includes every referent of a function-name or object-name within the fully interpreted language. Names of value-ranges are object-names formed from function-names.

I have chosen to follow the contemporary practice of allowing the interpretation of non-logical symbols to vary via a model-theoretic semantics. Every model will include a domain of objects and a domain of functions (and relations), where every function (and relation) is assigned a value-range within the domain of objects by the model.

Frege's domain contains two special objects: the True and the False. As we will see, introducing names for these objects (T and F) will give us a useful way of treating properties and relations as functions.

Because of our fundamental syntactic restriction, I will not follow Frege in treating the quantifiers as referring to functions from functions to truth-values, but will instead give them a standard contemporary interpretation. Furthermore, I will not require the interpretation function of a model to assign every well-formed formula a referent (i.e. its truth-value), instead sticking with a standard contemporary definition of truth in a model.

I

I will close the domain of functions and relations under arity reduction (e.g. the existence of a function $f(c, y)$ of arity 1 whenever $f(x, y)$ exists). Moving forward, I will always use e.g. $f(c, y)$ to denote the reduction of the referent of $f(x, y)$ via the referent of constant c at the first argument place. This will greatly simplify our semantics, allowing for a simple treatment of quantification into a value-range expression.

Note that given the motivation behind our fundamental syntactic restriction, we must insist that only function symbols are interpreted by the model as

relations of arity 1, respectively.

[9]If one feels that my axioms for identity require a non-trivial change to the *Grundgesetze*, one is welcome to treat my new identity axioms as part of our *extension* of the *Grundgesetze* rather than the *Grundgesetze* itself.

referring to functions within the domain. For example, $g(x)$ will be interpreted as a name of a function while $P(x) \to Q(x)$ will not. The latter is instead simply a well-formed formula with one free variable.

I will assume that the range of every function is a subset of the domain of objects. This will greatly simplify the syntax and has no effect on our development of arithmetic.[10]

Frege introduces many symbols via definitions in the course of his development of his formal theory of arithmetic. Given our fundamental syntactic restriction, we must ground these definitions in additional axioms governing the existence of functions.

3 Introduction to the Formal System

3.1 Syntax and Semantics

I will now introduce the formal system in more detail. See the Appendix for a rigorous description.

A *language* $\mathscr{L}$ will include the following symbols: (,), ,, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\forall$, $\exists$, $=$, $\ulcorner$, $\urcorner$, $\imath$, T, F. Strictly speaking, T and F will not be considered logical symbols, since their interpretations will be allowed to vary.

Once again, corner-quotes $\ulcorner$ and $\urcorner$ are used to construct names of value-ranges and $\imath$ is a definite article, referring to the unique object contained in an extension if it exists and the extension (or other object) itself otherwise.

Languages always include countably many object variables v_i and *may* include function variables ${f_i}^{n,m}$ or relation variables ${V_i}^{n,m}$, where n denotes the order and m denotes the arity. The reason that not all orders and arities need to be included is that I want to allow for cases in which the domain doesn't contain any functions of a particular order and arity. Note that whenever a function or relation variable is included, there must be countably many variables of that order and arity included in the language as well.

Naturally, a language may also include constants c_i, function symbols ${g_i}^{n,m}$, and relation symbols ${R_i}^{n,m}$. As in the case of variables, if the language contains a function symbol or relation symbol, then it must also contain countably many function or relation variables of that order and arity.

Given our semantics, relation symbols may also be thought of as function symbols, but their range must include no more than the two truth-values True and False.

Terms are defined in the natural way (see the Appendix), with two necessary restrictions:

(R1) Function variables and function symbols of order $n > 1$ can only contain a sequence of terms of order $n-1$ of the appropriate arity within their scope. Function variables and function symbols of order 1 can contain

[10] Frege allowed functions whose range includes functions, but this is unnecessary for our purposes.

a sequence of terms of *any* orders of the appropriate arity within their scope.

(R2) A value-range term can only contain a function variable, function symbol, relation variable, or relation symbol within its scope. $\ulcorner v_1, ..., v_k \urcorner ({V_i}^{1,m}(t_1, ..., t_m))$ is an example.

The reason that in restriction (R1) we allow function variables and symbols of order 1 to contain terms of *any* orders within their scope is that we want to allow for the case in which the function maps the objects denoted by the terms to a particular output *once an appropriate assignment has been given to the variables*. Recall that all functions have only objects within their range. Restriction (R2) ensures that our fundamental syntactic restriction applies to value-range terms.

Well-formed formulas are also defined in the natural way (see the Appendix), with the following two restrictions:

(R3) Relation variables and relation symbols of order $n > 1$ can only contain a sequence of terms or formulas of order $n - 1$ of the appropriate arity within their scope. Relation variables and relation symbols of order 1 may contain a sequence of terms or formulas of *any* orders of the appropriate arity within their scope.

(R4) Relation variables and symbols may only include terms and *atomic* formulas of the appropriate order within their scope.

As with function variables and symbols above, relation variables and symbols of order 1 can be applied to terms and formulas of *any* orders. This ensures that one can speak of relations holding between objects determined by a term *once an appropriate assignment has been given to the variables.*

The point of restriction (R4) is to ensure that our fundamental syntactic restriction holds: higher-order relation variables and symbols can only be applied to terms and relation variables and symbols. They cannot be applied to arbitrary well-formed formulas. For atomic formulas φ_i, ${R_i}^{n,m}(\varphi_1, ..., \varphi_m)$ is a well-formed formula while ${R_i}^{n,m}((\varphi_1 \vee \varphi_k), ..., \varphi_m)$ is not.

Turning now to semantics, a *model* $\mathcal{M}$ is an ordered tuple consisting of a domain of objects O, a domain of functions G, an interpretation function $\mathcal{A}$, and a value-range function $\circ$. Each element of G will have a *type* fixing order and arity.[11] O must contain two distinct objects $T^{\mathcal{M}}$ and $F^{\mathcal{M}}$.[12]

The value-range function $\circ$ will have the domain of functions G as its domain, mapping every element g of G to an element o of O, where o is the *value-range* of g.[13] Of course, o will be constrained by the order and arity of the function g. As stated above, the range of any function is a subset of the domain of objects.

[11] Note that an m-ary function of e.g. order 2 will have a well-defined output for every sequence of m inputs of order 1 of *any* arity. Restricting the arity of the inputs (as Frege does) would require a significantly more complex syntax and semantics.

[12] Must we make this restriction? No, but failing to do so leads to various complications.

[13] Here I use 'g' as a name of a function rather than a name of a symbol. I trust the reader can determine the meaning from the context moving forward.

Relation symbols R_i are interpreted as functions by the model. They will always be interpreted such that their range includes at most the two truth-values $T^{\mathcal{M}}$ (True) and $F^{\mathcal{M}}$ (False), as fixed by the interpretation function $\mathcal{A}$.

The symbols $\ulcorner$ and $\urcorner$ will be used to construct names of value-ranges. For example, $\ulcorner x, y \urcorner (g(x, y))$ refers to the value-range of the function referred to by g. Note that only the variables within corner quotes are bound.

The symbol $\imath$ is interpreted as referring to a function from objects to objects, mapping an object e to an object o if e is an extension and o is the unique element of e, and mapping e to e otherwise. Intuitively, this is Frege's version of the definite article ('the' in English). This symbol does not bind any variables within its scope.

I will close the domain of functions under the following operation: given any function of multiple arguments, one may hold some but not all of those arguments fixed at particular inputs, thereby defining a new function with the corresponding value-range. For example, if $g(x, y)$ refers to a function in the domain and c refers to an object in the domain, then $g(c, y)$ also refers to a function in the domain (in this case, a function of one argument). This will simplify our formal theory of arithmetic below.

Importantly, terms such as ${g_0}^{1,2}(c, y)$ will be interpreted as referring to the *restriction* of function g at first argument place by object c. This restriction is a function ${g_k}^{1,1}(y)$ such that $\forall y\ {g_k}^{1,1}(y) = {g_0}^{1,2}(c, y)$.

In particular, this allows us to quantify into a value-range expression. For example, $\forall x \ulcorner y \urcorner ({g_0}^{1,2}(x, y)) = c$ is a sentence which states that every restriction of function g at first argument place has a value-range equal to object c.

An *interpretation* $\mathcal{I}$ of a language $\mathscr{L}$ will be an ordered pair consisting of a model $\mathcal{M}$ of $\mathscr{L}$ and an *assignment* β mapping each variable in $\mathscr{L}$ to an object in O or function in G of the appropriate type with the appropriate range (as fixed by $\circ$).[14]

Note that this implies that for every non-logical symbol in the language (including variables), there must be at least one element in the domain of the appropriate type.

In the definition of a language above I have chosen to allow the set of variables to vary across languages. This allows more freedom with respect to the domain of functions and relations G, since otherwise every G would need to contain functions and relations of *every* possible order and arity. Recall that for every selected variable type, there must be countably many variables of that type.

Interpretations of arbitrary well-formed formulas are defined in the natural way, with one exception. The semantics for quantifiers is defined in such a way that one extends the language and interpretation with appropriate constants in order to determine truth-value. This allows my use of restrictions of functions mentioned above without additional complications. A rigorous development of

[14]Given the semantics, e.g. $g(x, y)$ can never be interpreted by an interpretation as the restriction of the function $g^{\mathcal{M}}$ by $\beta(x)$ at first argument applied to $\beta(y)$. Instead, $g(x, y)$ will always be interpreted as the result of applying $g^{\mathcal{M}}$, a function of arity 2, to arguments $\beta(x)$ and $\beta(y)$. This is to avoid complicating the syntax and semantics. Of course, this makes little difference in practice.

the semantics is contained in the Appendix.

3.2 Derivation System

Our derivation system is an extension of a standard higher-order derivation system *without* a comprehension scheme, where this extension includes Frege's Basic Laws V and VI. Note, however, that higher-order quantifiers range over the domain of functions and relations G, typed by each distinct variable type, rather than e.g. subsets of the domain of objects O. See the Appendix for a rigorous development of the semantics.

Moving forward, I will use the $\bar{}$ symbol in order to define *meta*-variables which are meant to range over variables of *all* types (e.g. $\bar{x}$). Within a context, all such variables must be of the appropriate types, however.

In order to simplify the semantics (i.e. in order to avoid the annoyance of general function symbol terms appearing in formula position, which Frege addressed with his horizontal), I have distinguished function symbols and variables from relation symbols and variables. Hence, many of our pairs of axiom schemes could in principle be reduced to a single axiom scheme.

The first two additional axiom schemes govern function restriction.

(L1) $\forall f^{n,m+1}\exists h^{n,m}\forall\bar{x}_1,...,\bar{x}_m(f(\bar{x}_1,...,t_i,...\bar{x}_m) = h(\bar{x}_1,...,\bar{x}_m))$, where the scope of h includes all variables $\bar{x}_i$ but does not include the term t_i, which must not contain any variables outside of the scope of a value-range expression.

(L2) $\forall V^{n,m+1}\exists W^{n,m}\forall\bar{x}_1,...,\bar{x}_m(V(\bar{x}_1,...,t_i,...\bar{x}_m) \leftrightarrow W(\bar{x}_1,...,\bar{x}_m))$, where the scope of W includes all variables $\bar{x}_i$ but does not include the term t_i, which must not contain any variables outside the scope of a value-range expression.

Let's also explicitly state that $T \neq F$:

(L3) $\neg T = F$

We will extend the system with Frege's Basic Laws V and VI. Note that given our syntactic distinction between function symbols and relation symbols, Basic Law V has been split into a pair of axioms. Our more general version of Basic Law V, which allows for cases of terms that do not contain variables, is needed to accommodate our allowance for restrictions of functions.

(Va) $\forall g^{n,k}\forall h^{n,k}(\ulcorner\bar{x}_1,...,\bar{x}_m\urcorner(g(t_1,...,t_k)) = \ulcorner\bar{x}_1,...,\bar{x}_m\urcorner(h(t_l,...,t_o)) \leftrightarrow \forall\bar{x}_1,...,\bar{x}_m(g(t_1,...,t_k) = h(t_l,...,t_o)))$, where the sequence $t_1,...,t_k$ and the sequence $t_l,...,t_o$ includes no variables other than $\bar{x}_1,...,\bar{x}_m$ outside the scope of a value-range expression.

(Vb) $\forall P^{n,k}\forall Q^{n,k}(\ulcorner\bar{x}_1,...,\bar{x}_m\urcorner(P(t_1,...,t_k)) = \ulcorner\bar{x}_1,...,\bar{x}_m\urcorner(Q(t_l,...,t_o)) \leftrightarrow \forall\bar{x}_1,...,\bar{x}_m(P(t_1,...,t_k) \leftrightarrow Q(t_l,...,t_o)))$, where the sequence $t_1,...,t_k$ and the sequence $t_l,...,t_o$ includes no variables other than $\bar{x}_1,...,\bar{x}_m$ outside the scope of a value-range expression.

(VI) $\forall\bar{x}(\bar{x} = \iota(\ulcorner\bar{y}\urcorner(\bar{x} = \bar{y})))$

We will add additional axioms governing arithmetic as we proceed.

3.3 Arithmetic

In a section titled "Special definitions" (Frege 1893, pg. 52), Frege immediately introduces a new symbol $\frown$ meant to be interpreted as a function of two object variables, intended to be an object and a value-range, where this function outputs the result of applying a function having that value-range to the object.

Since this definition is fundamental to Frege's development of arithmetic, I will describe it in detail. This will also show my general approach to formulating Frege's definitions within our framework.

Given our fundamental syntactic restriction, we will replace Frege's definition by two axioms introducing the relation symbol R_0 and the function symbol $\frown$:

(A1) $\forall x\forall y\forall z(R_0(x,y,z) \leftrightarrow (\exists f(y = \ulcorner w\urcorner f(w) \wedge f(x) = z) \vee \exists V(y = \ulcorner w\urcorner V(w) \wedge V(x) \wedge z = T)))$

So, the relation referred to by R_0 holds of x, y, and z if y is the value-range of a function that maps x to z or y is the extension of a property that x has and z is the True (i.e. the output of that property on input x).

(A2) $\forall x\forall y(x \frown y = \iota(\ulcorner z\urcorner R_0(x,y,z)))$

So, $a \frown b$ refers to the result of applying the function or relation of which b is the value-range to a. If b is the value-range of a property that a does not have, then this expression will refer to the empty extension.

Note that $a \frown b$ will also refer to the empty extension if b does not refer to a value-range. If b refers to the extension of a property that the referent of a has, then $a \frown b$ will refer to the True.

Frege makes extensive use of the analogue of this function for functions of two arguments. I will axiomatize it separately, introducing the relation symbol R_1 and the function symbol $\frown^{1,3}$:[15]

(A3) $\forall x_1\forall x_2\forall y\forall z(R_1(x_1,x_2,y,z) \leftrightarrow (\exists f(y = \ulcorner w_1, w_2\urcorner f(w_1,w_2) \wedge f(x_1,x_2) = z) \vee \exists V(y = \ulcorner w_1, w_2\urcorner V(w_1,w_2) \wedge V(x_1,x_2) \wedge z = T)))$

R_1 is simply the four-place analogue of R_0.

(A4) $\forall x_1\forall x_2\forall y(\frown^{1,3}(x_1,x_2,y) = \iota(\ulcorner z\urcorner R_1(x_1,x_2,y,z)))$

[15] My axiomatization behaves slightly differently from Frege's definition when the third term of $\frown^{1,3}$ does not refer to a value-range (Frege's will refer to the False; mine will refer to the empty extension). Some definitions below also behave slightly differently. This won't make any difference moving forward. In several cases I also use second-order quantification where Frege did not in order to greatly simplify proofs below.

What follows are a series of definitions needed to define equinumerosity and natural number. Let me make a few remarks. See the Appendix for an explicit development.

Frege's classical definition of "the number of" is the following: the number of y's, where y is the extension of a property, is the extension of a property $R(y, z)$ (i.e. the restriction of a relation R by y) that holds of z exactly if there exists a bijective relation holding between the objects in extensions y and z. Translated into set-theoretic terms, the number of y's would be the set of all sets equinumerous with y.

Frege defines 0 and 1 explicitly: 0 is the number of the value-range of the property I call C_0, defined as $x \neq x$ (i.e. a property with empty extension). 1 is the number of the value-range of the property C_1 holding of extensions equal to the extension of C_0. The remaining natural numbers can be explicitly defined at this point, but Frege himself defines them via the successor function.

Frege then defines the property "being a cardinal number" as follows: x is a cardinal number if and only if there exists a y such that the number of y's is x.

The next task is to define the successor relation. To do so, Frege defines a relation that holds between two numbers x and y if x is the number of objects with some property P distinct from some object with property P and y is the number of all objects with property P. The successor value-range is then defined as the value-range of this relation.

Frege then states six theorems without proof that characterize the definitions introduced so far. See the Appendix for my own proofs of these theorems.

Theorem 1: If x is the successor of 0, then $x = 1$.

Theorem 2: For any x, if the number of a property P with extension x is 1, then there exists a y with property P.

Theorem 3: For any x, y, z, if the number of a property P with extension x is 1 and y has property P and z has property P, then $y = z$.

Theorem 4: For any x, if x is the extension of a property P such that for any y, if y has P and any z has P then $y = z$, then if there is any object at all with property P, then the number of P's is 1.

Theorem 5: Any object has at most one predecessor and at most one successor.

Theorem 6: Any cardinal number other than 0 is such that there exists a predecessor of it.

Frege nowhere argues that the successor of any number is distinct from that number. It certainly isn't. Consider a bijection between the extension containing all natural numbers except for 0 and the extension containing all natural numbers. These two extensions have the same number and hence this number is its own successor. A contemporary mathematician may prefer a different definition of the successor relation.

To fully characterize the natural numbers, Frege introduces a definition of what it is for an object to follow another object in a series. This definition is a bit complex. Here it is in full:

y follows x in the z-series if and only if z is the extension of a binary relation W and for any property V, if whenever V holds of v_1 it follows that any v_2 that v_1 bears W to also has property V, then if for any v_3 such that x bears W with v_3, v_3 has property V, then y has property V.

In effect, this relation holds whenever from any property being transmitted down the z-series and x being such that bearing the relation to an object implies that object has that property, it follows that y has that property. That is, in this case all properties transmitted down the series are such that if x transmits them then y also has them.

Frege can then define the property of being a natural number. A*natural number* is any object either identical to 0 or such that it follows 0 in the successor-series.

Frege states a final theorem without proof. See the Appendix for my own proof.

Theorem 7: If n belongs to the series beginning with 0 defined by the successor relation, then the number of P's, where P is the property of belonging to to the successor-series ending with n, is the successor of n.

Hence, Frege can now characterize all natural numbers explicitly as particular extensions using his successor relation.

I will end this section by stating the Dedekind-Peano Axioms, which I prove in the Appendix.

(DP1) For every natural number n, the successor of n is a natural number.

(DP2) For all natural numbers n and m, $n = m$ if and only if the successor of $n =$ the successor of m.

(DP3) 0 does not have a predecessor.

(DP4) If P is a property such that $P(0)$ and for every natural number n, $P(n)$ implies P holds of the successor of n, then $P(n)$ is true for every natural number n.

This concludes our construction of the natural numbers within our modification of Frege's formal system.

3.4 Consistency

I follow the standard procedure of exhibiting a model of our axioms. What might a natural model look like?

In any natural model of Frege's system, every function must have a value-range. This value-range must be an object. But since functions are defined for all objects in the domain, this means that every function must have a well-defined output *given its own value-range as input*.

So, suppose we translate value-ranges into sets in the natural way (value-ranges of functions are sets of ordered tuples, etc.). This means that each value-range will be contained within the first component of an ordered tuple contained within *itself*. *Any* natural set-theoretic model of Frege's formal system will be radically *non-well-founded*. Every value-range provides a counterexample to the Axiom of Foundation.

Hence, I will not use ZFC to construct the needed model. Instead, I have chosen to use the set theory generated from removing the Axiom of Foundation from ZFC and replacing it by the Anti-Foundation Axiom of Peter Aczel (1988):

(AFA) Every accessible pointed directed graph corresponds to a unique set.

Hence, we will construct our model by describing an appropriate accessible pointed directed graph. An accessible pointed directed graph is a directed graph with a distinguished node (the *root*) such that for any node in the graph, there is at least one path from the root to that node.

The correspondence between such a graph and a set is the following: each edge leads from a set to one of its elements. So, the graph with a single node and no edges corresponds to the empty set. Such graphs can lead to non-well-founded sets. For example, the graph containing a single node with an edge from that node to itself corresponds to the Quine atom $x = \{x\}$.

Before proceeding to the proof, let's consider a more complex example that better corresponds to the sets we will need. Suppose that we have two sets T and F such that $T = \{(T, T), (F, F)\}$ and $F = \{(T, F), (F, T)\}$. Intuitively, T is the extension of the property "being identical to T" (this is Frege's actual definition) and F is the extension of the property "being identical to F", assuming a two-element domain.

As is customary, let any ordered pair $(a, b) = \{\{a\}, \{a, b\}\}$. We will now use (AFA) to prove that our two sets exist. For F, our graph will consist of a root node F with two edges leading to two distinct nodes (T, F) and (F, T). Node (T, F) will have edges leading to two further nodes $\{T\}$ and $\{T, F\}$. Node $\{T\}$ will have a single edge leading to a new node T while node $\{T, F\}$ will have two edges leading to node T and node F, respectively.

Node (F, T) is similar to node (T, F), with edges leading to node $\{F\}$ and node $\{T, F\}$, the former which has a single edge leading back to node F. Finally, node T has two edges leading to nodes (T, T) and (F, F), each of which has a single edge leading to node $\{T\}$ or node $\{F\}$, respectively.

The graph of T is identical except the root node is now T rather than F.

Let's call our theory FA.

Theorem 8: FA is consistent.

A sketch of a proof is contained in the Appendix.

4 Conclusion

I have demonstrated that by modifying the syntax of the formal system contained within Frege's *Grundgesetze* we can keep Basic Law V while developing arithmetic in a consistent way. This requires adding additional axioms asserting the existence of various functions.

The motivating idea is that the problem with Frege's formal system is *not* that every function has a value-range. Instead, the problem is that the language is "too rich", allowing one to form predicates that don't refer to functions on pain of contradiction. The solution was to only allow function symbols in variable position within higher-order functional expressions.

This has the historical advantage of allowing one to use Frege's own definitions and carry out Frege's own proofs. It has the philosophical advantage of allowing one to use Frege's own definition of number, which captures the philosophical thesis that a statement of number contains an assertion about a property. Finally, we saw that any natural model of Frege's formal system will be radically non-well-founded.

It remains to be seen whether this approach can be extended to Frege's treatment of the real numbers. I hope to have convinced the reader that this is a promising path to take.

5 Appendix

5.1 Syntax

A *language* $\mathscr{L}$ will include the following symbols: (,), ,, $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\forall$, $\exists$, $=$, $\ulcorner$, $\urcorner$, $\imath$, T, F. Strictly speaking, T and F will not be considered logical symbols, since their interpretations will be allowed to vary.

A language will also include the following object variables: v_0, v_1, v_2,... These variables may also be considered variables of type $0, 0$.

A language may also include some combination of the following types of variables (each type corresponds to a particular choice of n and m, with $n, m \geq 1$), but countably many for each selected type: ${f_0}^{n,m}$, ${f_1}^{n,m}$, ${f_2}^{n,m}$,..., ${V_0}^{n,m}$, ${V_1}^{n,m}$, ${V_2}^{n,m}$,...

Lower-case variables v_i will range over objects (*object variables*) while lower-case variables ${f_i}^{n,m}$ will range over functions (*function variables*). Upper-case variables ${V_i}^{n,m}$ will range over functions whose range consists of at most the two truth-values (*relation variables*). The first superscript indicates the order (first-order, second-order, etc.) and the second superscript indicates the arity.[16]

A language may also include constants c_0, c_1, c_2,..., m-ary function symbols of nth order (with $n, m \geq 1$) ${g_0}^{n,m}$, ${g_1}^{n,m}$,..., and m-ary relation symbols of nth order (with $n, m \geq 1$) ${R_0}^{n,m}$, ${R_1}^{n,m}$, ${R_2}^{n,m}$,... $=$ is also a (first-order, binary)

[16] One could extend the syntax to allow for what Frege called "unequal-leveled functions", e.g. a function that takes an object as its first argument and a function as its second argument. This will turn out to be unnecessary for our purposes.

relation symbol and $\imath$ is a (first-order, unary) function symbol. If a language includes function or relation symbols of a particular type, then it must also include countably many of the corresponding variables of that type.

Note that ${R_0}^{1,1}$ is a first-order predicate symbol. Note also that given our semantics, relation symbols can also be thought of as function symbols, although they will always be interpreted as referring to functions whose range consists of at most the two truth-values. One may extend the definition a language to allow for uncountably many non-logical symbols in the obvious way.

For the purposes of developing arithmetic, we will only need to consider second-order languages. I will state rules for constructing well-formed formulas containing function symbols of arbitrary order, however.

Moving forward, I will use the symbol $^-$ to indicate that I am using *meta*-variable meant to range over variables of *all* types, e.g. $\bar{x}$.

A *term* is any expression generated by finitely many applications of the following rules:

(T1) All constants c_i are terms of order 0.

(T2) All object variables v_i are terms of order 0.

(T3) If $(t_1, ..., t_m)$ is a sequence of terms of order $n-1$, for $n, m \geq 1$, then any ${f_i}^{n,m}(t_1, ..., t_m)$ or ${g_i}^{n,m}(t_1, ..., t_m)$ is a term of order n.

(T4) If $(t_1, ..., t_m)$ is a sequence of terms of *any* orders, then any ${f_i}^{1,m}(t_1, ..., t_m)$ or ${g_i}^{1,m}(t_1, ..., t_m)$ is a term of order 1.

(T5) If $\bar{x}_1, ..., \bar{x}_k$ are variables and ${g_i}^{n,m}(t_1, ..., t_m)$ is a function symbol followed by a sequence of terms of appropriate order, then $\ulcorner \bar{x}_1, ..., \bar{x}_k \urcorner({g_i}^{n,m}(t_1, ..., t_m))$ is a term of order 0.

(T6) If $\bar{x}_1, ..., \bar{x}_k$ are variables and ${f_i}^{n,m}(t_1, ..., t_m)$ is a function variable followed by a sequence of terms of appropriate order, then $\ulcorner \bar{v}_1, ..., \bar{v}_k \urcorner({f_i}^{n,m}(t_1, ..., t_m))$ is a term of order 0.

(T7) If $\bar{v}_1, ..., \bar{v}_k$ are variables and ${R_i}^{n,m}(t_1, ..., t_m)$ is a relation symbol followed by a sequence of terms of appropriate order, then $\ulcorner \bar{x}_1, ..., \bar{x}_k \urcorner({R_i}^{n,m}(t_1, ..., t_m))$ is a term of order 0.

(T8) If $\bar{x}_1, ..., \bar{x}_k$ are variables and ${V_i}^{n,m}(t_1, ..., t_m)$ is a relation variable followed by a sequence of terms of appropriate order, then $\ulcorner \bar{x}_1, ..., \bar{x}_k \urcorner({V_i}^{n,m}(t_1, ..., t_m))$ is a term of order 0.

(T9) If t_1 is a term, then $\imath(t_1)$ is a term of order 0.

Note that the expression ${f_0}^{1,2}(c_0, c_1)$ is a term of order 1 rather than a term of order 0. When placed within the scope of a second-order function symbol, the two constants are ignored and it is instead the function symbol that is treated as argument. This avoids additional syntactic complications.

Frege himself would include arbitrary well-formed formulas as terms. Recall our fundamental syntactic restriction. The expression $\ulcorner v_0 \urcorner({R_0}^{1,1}(v_0) \wedge {R_1}^{1,1}(v_0))$ is *not* a term.

A *formula* is any expression generated by finitely many applications of the following rules:

(F1) If t_1 and t_2 are terms, then $=(t_1, t_2)$ is an atomic formula of order 1.

(F2) If $t_1, ..., t_m$ are terms of order $n-1$, then any ${R_i}^{n,m}(t_1, ..., t_m)$ or ${V_i}^{n,m}(t_1, ..., t_m)$ is an atomic formula of order n.

(F3) If $\varphi_1, ..., \varphi_m$ are *atomic* formulas of order $n-1$, then any ${R_i}^{n,m}(\varphi_1, ..., \varphi_m)$ or ${V_i}^{n,m}(\varphi_1, ..., \varphi_m)$, where no φ_j is an identity formula, is an atomic formula of order n.

(F4) If $t_1, ..., t_m$ are terms of *any* orders, then any ${R_i}^{1,m}(t_1, ..., t_m)$ or ${V_i}^{1,m}(t_1, ..., t_m)$ is an atomic formula of order 1.

(F5) If φ is a formula, then $\neg\varphi$ is also a formula.

(F6) If φ and ψ are formulas, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are also formulas.

(F7) If φ is a formula and $\bar{x}$ is a variable, then $\forall\bar{x}\varphi$ and $\exists\bar{x}\varphi$ are also formulas.

Variables of the appropriate type within the scope of a quantifier or $\ulcorner\urcorner$ expression are *bound*. Also, within the scope of a higher-order function symbol or relation symbol that is not itself within the scope of a first-order function symbol or relation symbol, every variable of order lower than the order of the function's arguments is *bound*.

5.2 Semantics

A *model* $\mathcal{M}$ is an ordered tuple consisting of a domain of objects O, a domain of functions G, an interpretation function $\mathcal{A}$, and a value-range function $\circ$. Each element of G will have a *type* fixing order and arity.[17]

The value-range function $\circ$ will have the domain of functions G as its domain, mapping every element g of G to an element o of O, where o is the *value-range* of g.[18] Of course, o will be constrained by the order and arity of the function g. As stated above, the range of any function is a subset of the domain of objects.

The symbols $\ulcorner$ and $\urcorner$ will be used to construct a name of a value-range. For example, $\ulcorner x, y \urcorner(g(x, y))$ refers to the value-range of the function referred to by g. Note that only the variables within corner quotes are bound.

[17]Note that an m-ary function of e.g. order 2 will have a well-defined output for every sequence of m inputs of order 1 of *any* arity. Restricting the arity of the inputs (as Frege does) would require a significantly more complex syntax and semantics.

[18]Here I use 'g' as a name of a function rather than a name of a symbol. I trust the reader can determine the meaning from the context moving forward.

The symbol $\imath$ is interpreted as referring to a function from objects to objects, mapping an object e to an object o if e is an extension and o is the unique element of e, and mapping e to e otherwise. Intuitively, this is Frege's version of the definite article ('the' in English). This symbol does not bind any variables within its scope.

I will close the domain of first-order functions under the following operation: given any first-order function of multiple arguments, one may hold some but not all of those arguments fixed at particular inputs, thereby defining a new function with the corresponding value-range. For example, if $g(x, y)$ refers to a function in the domain and c refers to an object in the domain, then $g(c, y)$ also refers to a function in the domain (in this case, a function of one argument). This simplifies our formal theory of arithmetic and quantification into value-range expressions.

More carefully, the *value-range function* $\circ$ maps each element of G to a set of ordered pairs in O whose first component is either an object in O or function in G of the appropriate type if the element of G has arity 1 or an ordered n-tuple consisting of objects in O or functions in G of the appropriate type if the element of G has arity $n > 1$. All elements in the n-tuple must be of the same order. The second component is always an object in O, and is of course subject to the constraint that for each first component there is a unique second component in the set of ordered pairs. The output of $\circ$ will always be referred to as the *value-range* of its input, and if the second component of the ordered pairs in the output are always truth-values (as fixed by the interpretation function below) then the value-range will also be referred to as the *extension* of the input.

$\circ$ maps each relation in G of arity 1 to a set containing objects in O or functions or relations in G of the appropriate type. $\circ$ maps each relation in G of arity $n > 1$ to a set of ordered tuples in O of objects in O or functions in G of the appropriate type. Any such value-range is also known as an *extension*.

Furthermore, a model $\mathcal{M}$ of a language $\mathscr{L}$ includes an interpretation function $\mathcal{A}$ mapping T and F to objects $T^{\mathcal{M}}$ and $F^{\mathcal{M}}$ in O, where $T^{\mathcal{M}} \neq F^{\mathcal{M}}$, each constant symbol c_i to an object ${c_i}^{\mathcal{M}}$ in O, each relation symbol ${R_i}^{n,m}$ to a function ${R_i}^{n,m,\mathcal{M}}$ in G of order n and arity m whose range consists of at most the two truth-values (as fixed by $\circ$), and each function symbol ${g_i}^{n,m}$ to a function ${g_i}^{n,m,\mathcal{M}}$ in G of order n with arity m.

$=^{\mathcal{M}}$ is the first-order identity relation, a function mapping a pair of objects to $T^{\mathcal{M}}$ if the first object is identical to the second object and $F^{\mathcal{M}}$ otherwise.

For any variable of *any* type, the application of $\mathcal{A}$ to that variable results in that variable as output. For example, ${v_i}^{\mathcal{M}} = v_i$.

For any term t_i containing constants, function symbols, relation symbols, or variables as components where either all terms contain variables or all terms do not contain variables, ${t_i}^{\mathcal{M}}$ consists of the application of $\mathcal{A}$ to each such component.[19]

For any term t_i containing a function symbol or relation symbol or variable

[19] Note that e.g. ${g_1}^{1,1}(v_1)^{\mathcal{M}}$ consists of the sequence ${g_1}^{1,1,\mathcal{M}}, v_1$. The first element is a function in G and the second element is a variable in $\mathscr{L}$ (an element of the *syntax*).

of order 1 followed by a sequence of terms $t_1, ..., t_n$ within its scope including some but not all terms t_j containing no variables, ${t_i}^{\mathcal{M}}$ consists of the result of holding the arguments t_j fixed by virtue of their interpretations ${t_j}^{\mathcal{M}}$. For example, ${g_1}^{1,2}(c_1, v_1)^{\mathcal{M}}$ is the sequence consisting of the function of arity 1 in G resulting from holding the first argument place of ${g_1}^{1,2,\mathcal{M}}$ fixed at ${c_1}^{\mathcal{M}}$ followed by the variable v_1.

We must also fix the interpretations of our non-standard symbols. For any variables $x_1, ..., x_k$ and $y_1, ..., y_m$ of order 0 and any function symbol ${g_i}^{o,m}$ or relation symbol ${R_i}^{o,m}$, $\mathcal{A}$ maps $\ulcorner x_1, ..., x_k \urcorner({g_i}^{o,m}(y_1, ..., y_m))$ and $\ulcorner x_1, ..., x_k \urcorner({R_i}^{o,m}(y_1, ..., y_m))$ to $\circ({g_i}^{o,m,\mathcal{M}})$ and $\circ({R_i}^{o,m,\mathcal{M}})$, respectively.

If in ${g_i}^{o,m}(t_1, ..., t_m)$ or ${R_i}^{o,m}(t_1, ..., t_m)$ the sequence $t_1, ..., t_m$ includes terms that do not contain any variables in addition to free object variables, then $\mathcal{A}$ maps $\ulcorner x_1, ..., x_k \urcorner({g_i}^{o,m}(t_1, ..., t_m))$ and $\ulcorner x_1, ..., x_k \urcorner({R_i}^{o,m}(t_1, ..., t_m))$ to the result of applying $\circ$ to the function or relation resulting from ${g_i}^{o,m}$ or ${R_i}^{o,m}$ with the appropriate arguments held fixed by the interpretations of the terms t_i which do not contain any variables.

If in ${g_i}^{n,m}(t_1, ..., t_m)$ or ${R_i}^{n,m}(t_1, ..., t_m)$ the sequence $t_1, ..., t_m$ does not contain any free variables, then $\mathcal{A}$ maps $\ulcorner \bar{x}_1, ..., \bar{x}_k \urcorner({g_i}^{n,m}(t_1, ..., t_m))$ and $\ulcorner \bar{x}_1, ..., \bar{x}_k \urcorner({R_i}^{n,m}(t_1, ..., t_m))$ to $\circ({g_i}^{n,m,\mathcal{M}})$ and $\circ({R_i}^{n,m,\mathcal{M}})$, respectively.

For any term t_i that does not contain free variables, if ${t_i}^{\mathcal{M}}$ is in the range of $\circ$ and is an extension with a unique element o that is not a tuple (that is, o is the first component of the unique ordered pair in ${t_i}^{\mathcal{M}}$ with $T^{\mathcal{M}}$ as its second component, if it exists, and o is not a tuple), then $\mathcal{A}$ maps $\imath(t_i)$ to o. In any other case, $\mathcal{A}$ maps $\imath(t_i)$ to ${t_i}^{\mathcal{M}}$.

An *interpretation* $\mathcal{I}$ of a language $\mathscr{L}$ will be an ordered pair consisting of a model $\mathcal{M}$ of $\mathscr{L}$ and an *assignment* β mapping each variable in $\mathscr{L}$ to an object in O or function in G of the appropriate type with the appropriate range (as fixed by $\circ$).[20]

Note that this implies that for every non-logical symbol in the language (including variables), there must be at least one element in the domain of the appropriate type.

In the definition of a language above I have chosen to allow the set of variables to vary across languages. This allows more freedom with respect to the domain of functions G, since otherwise every G would need to contain functions of *every* possible order and arity. Recall that for every selected variable type, there must be countably many variables of that type.

Let next define the interpretation $\mathcal{I}$ of arbitrary terms using $\circ\mathcal{M}$.

For any term t_i that does not contain a function symbol outside of the scope of $\ulcorner\urcorner$, ${t_i}^{\circ\mathcal{M}} = {t_i}^{\mathcal{M}}$.

For any function symbol ${g_i}^{n,m}$ outside of the scope of $\ulcorner\urcorner$ and sequence of terms $t_1, ..., t_m$ which do not contain free variables outside the scope of $\ulcorner\urcorner$,

[20] Given the semantics, e.g. $g(x, y)$ can never be interpreted by an interpretation as the restriction of the function $g^{\mathcal{M}}$ by $\beta(x)$ at first argument applied to $\beta(y)$. Instead, $g(x, y)$ will always be interpreted as the result of applying $g^{\mathcal{M}}$, a function of arity 2, to arguments $\beta(x)$ and $\beta(y)$. This is to avoid complicating the syntax and semantics. Of course, this makes little difference in practice.

${g_i}^{n,m}(t_1, ..., t_m)^{\circ\mathcal{M}}$ is the second component of the ordered pair with first component $({t_1}^{\circ\mathcal{M}}, ..., {t_m}^{\circ\mathcal{M}})$ contained in $\circ({g_i}^{n,m,\mathcal{M}})$.

If any t_i do contain free variables outside the scope of $\ulcorner\urcorner$, then ${g_i}^{n,m}(t_1, ..., t_m)^{\circ\mathcal{M}} = {g_i}^{n,m}(t_1, ..., t_m)$ (i.e. it remains an expression of the language).

(I1) For terms t_i that do not contain free variables outside the scope of $\ulcorner\urcorner$, $\mathcal{I}(t_i) = {t_i}^{\circ\mathcal{M}}$.

(I2) For terms t_i that *do* contain free variables t_j outside the scope of $\ulcorner\urcorner$, extend the language by introducing new constants, function symbols, or relation symbols of the appropriate type: r_{t_j}. Now extend the model $\mathcal{M}$ to $\mathcal{M}'$, which is just like $\mathcal{M}$ except for each t_j, ${r_{t_j}}^{\mathcal{M}'} = \beta(t_j)$. Convert t_i to a term t'_i with all free variables outside the scope of $\ulcorner\urcorner$ replaced by the terms just introduced. Finally, $\mathcal{I}(t_i) = {t'_i}^{\circ\mathcal{M}'}$.

It remains to define truth in a model relative to an assignment. If a formula is not true in a model relative to an assignment, then it is false in that model relative to that assignment. If a formula contains no free variables or only contains free variables within the scope of $\ulcorner\urcorner$, then we can also say that it is simply true in a model.

(T1) $=(t_i, t_j)$ is true in $\mathcal{M}$ relative to assignment β if $\mathcal{I}(t_i) = \mathcal{I}(t_j)$.

(T2) ${R_i}^{n,m}(t_1, ..., t_m)$ is true in $\mathcal{M}$ relative to assignment β if the second component of the ordered pair contained in $\circ({R_i}^{n,m,\mathcal{M}})$ with first component $(\mathcal{I}(t_1), ..., \mathcal{I}(t_m))$ is $T^{\mathcal{M}}$.

(T3) ${V_i}^{n,m}(t_1, ..., t_m)$ is true in $\mathcal{M}$ relative to assignment β if the second component of the ordered pair contained in $\circ(\beta({V_i}^{n,m}))$ with first component $(\mathcal{I}(t_1), ..., \mathcal{I}(t_m))$ is $T^{\mathcal{M}}$.

(T4) Where $\varphi_1, ..., \varphi_m$ are atomic formulas either beginning with a relation symbol R_j (including $=$) or a relation variable V_j, ${R_i}^{n,m}(\varphi_1, ..., \varphi_m)$ is true in $\mathcal{M}$ relative to assignment β if the second component of the ordered pair contained in $\circ({R_i}^{n,m,\mathcal{M}})$ with first component $(u_1, ..., u_m)$ is $T^{\mathcal{M}}$. If there are no terms of order 0 within the scope of $\ulcorner\urcorner$ without free variables and all the terms in φ_i all contain free variables outside of the scope of $\ulcorner\urcorner$ or all do not contain free variables except those within the scope of $\ulcorner\urcorner$, then u_i is the interpretation of the relation symbol or relation variable beginning φ_i by model $\mathcal{M}$ relative to assignment β. If either there are terms of order 0 within the scope of $\ulcorner\urcorner$ or some but not all terms in φ_i do not contain free variables except those within the scope of $\ulcorner\urcorner$, then u_i is the appropriate restriction (at the relevant argument places by the interpretations of the relevant terms) of the interpretation of the expression beginning φ_i by model $\mathcal{M}$ relative to assignment β.[21]

[21] In the former case, one first restricts functions within the scope of $\ulcorner\urcorner$ and then restricts the outer function. Recall that functions can only be restricted by objects.

(T5) Where $\varphi_1, ..., \varphi_m$ are atomic formulas either beginning with a relation symbol R_j (including $=$) or a relation variable V_j, ${V_i}^{n,m}(\varphi_1, ..., \varphi_m)$ is true in $\mathcal{M}$ relative to assignment β if the second component of the ordered pair contained in $\circ(\beta({V_i}^{n,m}))$ with first component $(u_1, ..., u_m)$ is $T^{\mathcal{M}}$. If there are no terms of order 0 within the scope of $\ulcorner\urcorner$ without free variables and all the terms in φ_i all contain free variables outside of the scope of $\ulcorner\urcorner$ or all do not contain free variables except those within the scope of $\ulcorner\urcorner$, then u_i is the interpretation of the relation symbol or relation variable beginning φ_i by model $\mathcal{M}$ relative to assignment β. If either there are terms of order 0 within the scope of $\ulcorner\urcorner$ or some but not all terms in φ_i do not contain free variables except those within the scope of $\ulcorner\urcorner$, then u_i is the appropriate restriction (at the relevant argument places by the interpretations of the relevant terms) of the interpretation of the expression beginning φ_i by model $\mathcal{M}$ relative to assignment β.

(T6) $\neg\varphi$ is true in $\mathcal{M}$ relative to assignment β exactly if φ is false in $\mathcal{M}$ relative to assignment β.

(T7) $(\varphi \wedge \psi)$ is true in $\mathcal{M}$ relative to assignment β exactly if both φ and ψ are true in $\mathcal{M}$ relative to assignment β.

(T8) $(\varphi \vee \psi)$ is true in $\mathcal{M}$ relative to assignment β exactly if either φ or ψ is true in $\mathcal{M}$ relative to assignment β or both are.

(T9) $(\varphi \rightarrow \psi)$ is true in $\mathcal{M}$ relative to assignment β exactly if either $\neg\varphi$ is true in $\mathcal{M}$ relative to assignment β or ψ is true in $\mathcal{M}$ relative to assignment β.

(T10) $(\varphi \leftrightarrow \psi)$ is true in $\mathcal{M}$ relative to assignment β exactly if either both φ and ψ are true in $\mathcal{M}$ relative to assignment β or both φ and ψ are false in $\mathcal{M}$ relative to assignment β.

(T11) For any variable $\bar{x}$, $\forall\bar{x}\varphi$ is true in $\mathcal{M}$ relative to assignment β exactly if we were to extend the language with an appropriate constant, function symbol, or relation symbol of the relevant type, *any* extension of the interpretation function $\mathcal{A}$ to include that symbol as input would be such that if we were to replace each instance of $\bar{x}$ free in φ by that new symbol, the resulting formula would be true in $\mathcal{M}$ relative to assignment β.[22]

(T12) For any variable $\bar{x}$, $\exists\bar{x}\varphi$ is true in $\mathcal{M}$ relative to assignment β exactly if we were to extend the language with an appropriate constant, function symbol, or relation symbol of the relevant type, there is *some* extension of the interpretation function $\mathcal{A}$ to include that symbol as input that would be such that if we were to replace each instance of $\bar{x}$ free in φ by that new symbol, the resulting formula would be true in $\mathcal{M}$ relative to assignment β.

[22]Note that this allows one to restrict a function relative to a quantifier. $\forall x(\ulcorner y \urcorner g(x, y) = c)$ will be true if any *restriction* of g at first argument place has value-range c.

5.3 Derivation System

The derivation system is described in section 3.2 above.

5.4 Towards a Formal Theory of Arithmetic

Moving forward, I will introduce axioms for convenience and axioms of necessity. When an axiom is introduced merely for convenience, I will flag it with an asterisk. Axioms introduced for convenience define function symbols that are never actually used as arguments for higher-order function symbols, and hence one could technically do without them in what follows. I have chosen to keep my axioms nearly as weak as possible (minimalism over elegance). I will follow Frege's own definitions quite closely. Note that many of Frege's definitions are designed to use value-ranges to avoid the use of second-order functions.

5.4.1 Defining the Natural Numbers

In a section titled "Special definitions" (Frege 1893, pg. 52), Frege immediately introduces a new symbol $\frown$ meant to be interpreted as a function of two object variables, intended to be an object and a value-range, where this function outputs the result of applying a function having that value-range to the object. Given our fundamental syntactic restriction, we will replace Frege's definition by two axioms introducing the relation symbol R_0 and the function symbol $\frown$:

(A1) $\forall x \forall y \forall z (R_0(x,y,z) \leftrightarrow (\exists f (y = \ulcorner w \urcorner f(w) \wedge f(x) = z) \vee \exists V (y = \ulcorner w \urcorner V(w) \wedge V(x) \wedge z = T)))$

So, the relation referred to by R_0 holds of x, y, and z if y is the value-range of a function that maps x to z or y is the extension of a property that x has and z is the True (i.e. the output of that property on input x).

(A2) $\forall x \forall y (x \frown y = \imath(\ulcorner z \urcorner R_0(x,y,z)))$

Note that $a \frown b$ will refer to the empty extension if b does not refer to a value-range. If b refers to the extension of a property that the referent of a has, then $a \frown b$ will refer to the True.

Frege makes extensive use of the analogue of this function for functions of two arguments. I have chosen to axiomatize it separately, introducing the relation symbol R_1 and the function symbol $\frown^{1,3}$:[23]

(A3) $\forall x_1 \forall x_2 \forall y \forall z (R_1(x_1,x_2,y,z) \leftrightarrow (\exists f (y = \ulcorner w_1, w_2 \urcorner f(w_1,w_2) \wedge f(x_1,x_2) = z) \vee \exists V (y = \ulcorner w_1, w_2 \urcorner V(w_1,w_2) \wedge V(x_1,x_2) \wedge z = T)))$

R_1 is simply the four-place analogue of R_0.

[23] My axiomatization behaves slightly differently from Frege's definition when the third term of $\frown^{1,3}$ does not refer to a value-range (Frege's will refer to the False; mine will refer to the empty extension). Several of my definitions below are also slightly different. This won't make any difference moving forward. In several cases I will also use second-order quantification where Frege did not in order to greatly simplify proofs below.

(A4) $\forall x_1 \forall x_2 \forall y(\frown^{1,3} (x_1, x_2, y) = \iota(\ulcorner z \urcorner R_1(x_1, x_2, y, z)))$

Frege next introduces a new function symbol in order to define a function that maps an object to the True if that object is the extension of a binary relation $R(x, y)$ such that each object only bears the relation to a unique object, if any. That is, in such a case $R(x, y)$ and $R(x, z)$ implies $y = z$. Frege calls such a relation *single-valued*. I will introduce a new predicate symbol I via the following axiom:[24]

(A5*) $\forall x(I(x) \leftrightarrow (\exists R(x = \ulcorner y, z \urcorner R(y, z)) \wedge \forall y \forall z(\frown^{1,3} (y, z, x) = T \rightarrow \forall w(\frown^{1,3} (y, w, x) = T \rightarrow z = w))))$

With the preliminaries out of the way, we can move toward the definitions of the natural numbers. The first step is to define equinumerosity between properties. To begin, we will introduce a three-place relation symbol R_2, where the relation holds between the extension of a binary relation and two property extensions if the binary relation is single-valued and correlates every object with the first property with a unique object with the second property.

(A6) $\forall x \forall y \forall z(R_2(x, y, z) \leftrightarrow (I(x) \wedge \exists V(y = \ulcorner v \urcorner V(v)) \wedge \exists W(z = \ulcorner v \urcorner W(v)) \wedge \forall w((w \frown y) = T \rightarrow \exists v(\frown^{1,3} (w, v, x) = T \wedge (v \frown z) = T))))$

Next we'll introduce a new function symbol $\rangle$ which maps objects to value-ranges:[25]

(A7*) $\forall x(\rangle(x) = \ulcorner y, z \urcorner R_2(x, y, z))$

Consider the expression $\frown^{1,3} (y, z, \rangle(x))$. This expression will refer to the True if and only if y and z are extensions of properties and x is the extension of a single-valued relation R that correlates every object in y with an object in z. Hence, for any x that is the extension of a binary relation, $\rangle(x)$ will refer to an extension of all pairs of property extensions that are correlated by that relation.

Next we will introduce an expression $\triangleleft$ to handle the converse of a binary relation, moving closer to our definition of equinumerosity. First, I will close the domain of first-order binary relations under converse. Then I will introduce a new relation $R_3(x, y)$. Then I will define the new expression:

(A8) $\forall V \exists W \forall x \forall y(V(x, y) \leftrightarrow W(y, x))$

(A9) $\forall x \forall y(R_3(x, y) \leftrightarrow \exists V(x = \ulcorner v, w \urcorner V(v, w) \wedge \exists W(y = \ulcorner v, w \urcorner W(v, w) \wedge \forall v \forall w(V(v, w) \leftrightarrow W(w, v)))))$

This relation holds if x and y are the extensions of first-order binary relations that are converses of each other.

[24] This axiom is optional because $I(x)$ will not be used as the input to a second-order function such as the value-range function and thereby could always be replaced by a formula.

[25] The function $\rangle$ is optional because Frege only uses it in cases in which we could replace it with R_2.

(A10) $\forall x(\triangleleft(x) = \iota\ulcorner y \urcorner R_3(x, y))$

Consider the following sentence: $\frown^{1,3} (a, b, \rangle(c)) = T \wedge \frown^{1,3} (b, a, \rangle(\triangleleft c)) = T)$. This sentence will be true exactly if c refers to the extension of a binary relation R and a and b refer to the extensions of properties such that R correlates the first property with the second and the converse of R correlates the second property with the first.

Hence, the sentence $\exists x(\frown^{1,3} (a, b, \rangle(x)) = T \wedge \frown^{1,3} (b, a, \rangle(\triangleleft x)) = T))$ states that properties with extensions a and b are equinumerous. Note that given our allowance for a restricted domain of functions within a model, this sentence will only be true if there does in fact exist such a relation. That is to say, the existence of a relation with extension x is not guaranteed.

Now we can state Frege's classical definition of "the number of Ps": the extension of the property "being equinumerous with P".[26] I will use the symbol # for this purpose. First we'll introduce the relation R_4, then the function referred to by #:[27]

(A11) $\forall y \forall z(R_4(y, z) \leftrightarrow \exists x(\frown^{1,3} (y, z, \rangle x) = T \wedge \frown^{1,3} (z, y, \rangle \triangleleft x) = T))$

(A12*) $\forall y(\#(y) = \ulcorner z \urcorner R_4(y, z))$

Next, we will introduce predicates C_0 and C_1 that can be used to define 0 and 1 explicitly:

(A13.0) $\forall x(C_0(x) \leftrightarrow \neg(x = x))$

(A13.1) $\forall x(C_1(x) \leftrightarrow x = \ulcorner y \urcorner C_0(y))$

Since every object is identical with itself, the extension of C_0 is empty. The extension of C_1 only contains the extension of C_0. We can now explicitly define the numbers 0 and 1 (note that the following are definitions rather than axioms):

(D0) $0 = \#\ulcorner x \urcorner C_0(x)$

(D1) $1 = \#\ulcorner x \urcorner C_1(x)$

Frege then defines the property "being a cardinal number", which we will express using C, as follows:

(A14) $\forall x(C(x) \leftrightarrow \exists y \# y = x)$

At this point, we could define the remaining natural numbers explicitly by introducing predicates C_i. Frege instead defines a successor relation and uses it to define the remaining natural numbers. Theorem 7 will characterize the remaining natural numbers as the numbers of particular properties.

To do so, Frege specifies a sentence that he interprets as stating that b follows a immediately in the number series. To follow him, we will define the relations R_5 and R_6:

[26] Actually we will follow Frege in defining this function and property in terms of *extensions*.
[27] The axiom defining # is for convenience since we will not be using this function as an argument for a second-order function, and hence it could be replaced by a term throughout.

(A15) $\forall x \forall y \forall z(R_5(x, y, z) \leftrightarrow x \frown y = T \wedge \neg x = z)$

This relation holds exactly if y is the extension of a property that x has where x is distinct from z.

(A16) $\forall x \forall y(R_6(x, y) \leftrightarrow \exists z \exists w(\#\ulcorner v \urcorner R_5(v, z, w) = x \wedge w \frown z = T \wedge \#z = y))$

This relation holds exactly if there is some property P such that x is the number of objects with that property that are distinct from some object that also has that property and y is the number of all the objects with that property. So, if the number of P's is 3, then x will be 2 and y will be 3.

Frege then defines the successor extension as the extension of R_6:

(D2) $s = \ulcorner x, y \urcorner R_6(x, y)$

Hence, $\frown^{1,3}(0, 1, s) = T$ states that 1 is the successor of 0.

Frege then states the following theorems without proof. I will sketch proofs below. As usual, we will need to assert the existence of various functions due to our fundamental syntactic constraint.

Theorem 1: $\forall x(\frown^{1,3}(0, x, s) = T \rightarrow x = 1)$

If x is the successor of 0, then $x = 1$.

Theorem 2: $\forall x(\#x = 1 \rightarrow \exists y(y \frown x = T))$

For any x, if the number of a property P with extension x is 1, then there exists a y with property P.

Theorem 3: $\forall x \forall y \forall z((\#x = 1 \wedge y \frown x = T \wedge z \frown x = T) \rightarrow y = z)$

For any x, y, z, if the number of a property P with extension x is 1 and y has property P and z has property P, then $y = z$.

Theorem 4: $\forall x((\forall y(y \frown x = T \rightarrow \forall z(z \frown x = T \rightarrow y = z))) \rightarrow ((\exists w(w \frown x = T) \rightarrow \#x = 1)))$

For any x, if x is the extension of a property P such that for any y, if y has P and any z has P then $y = z$, then if there is any object at all with property P, then the number of P's is 1.

Theorem 5: $I(\triangleleft s) \wedge I(s)$

Any object has at most one predecessor and at most one successor.

Theorem 6: $\forall x(\exists y \#y = x \rightarrow (\neg x = 0 \rightarrow \exists z \frown^{1,3}(z, x, s) = T))$

Any cardinal number other than 0 is such that there exists a predecessor of it.

Here I will sketch proofs of these theorems.

For theorems 1, 4, and 5, I will assert the closure of single-valued binary relations with single-valued converses under exchange of one element with another (A17). Consider a bijective function mapping a to b and c to d. Simply map a to d and c to b instead.[28]

For theorem 5, we will also need two additional axioms (A18) and (A19) allowing the definition of a new relation via the addition and removal of elements of a bijective relation. Intuitively if R holds between a and b, c and d, and e and f, all of which are distinct, then one should be able to define a new bijective relation also holding between g and h, and one should also be able to define a new bijective relation not holding between e and f.

Finally, for theorem 6, I will close the domain of properties under the following restriction: if the extension of a property is nonempty, then there exists a property that holds of all and only instances of that property except some particular object (A20).

(A17) $\forall V((I(\ulcorner y,z\urcorner V(y,z)))\wedge I(\lhd\ulcorner y,z\urcorner V(y,z))\rightarrow((\forall x\forall y\forall z\forall w((V(x,y)\wedge V(z,w)\wedge y\neq w)\rightarrow\exists W(\forall v_1\forall v_2((x\neq v_1\wedge z\neq v_1\wedge y\neq v_2\wedge w\neq v_2)\rightarrow((V(v_1,v_2)\leftrightarrow W(v_1,v_2))\wedge W(x,w)\wedge W(z,y)\wedge\neg W(x,y)\wedge\neg W(z,w))))))))$

(A18) $\forall V((I(\ulcorner y,z\urcorner V(y,z)))\wedge I(\lhd\ulcorner y,z\urcorner V(y,z))\rightarrow((\forall x\forall y((\neg\exists z(V(x,z)\vee V(z,y))\rightarrow\exists W(\forall v_1\forall v_2((x\neq v_1\wedge y\neq v_2)\rightarrow(V(v_1,v_2)\leftrightarrow W(v_1,v_2)))\wedge W(x,y)))))))$

(A19) $\forall V((I(\ulcorner y,z\urcorner V(y,z)))\wedge I(\lhd\ulcorner y,z\urcorner V(y,z))\rightarrow((\forall x\forall y(V(x,y)\rightarrow\exists W(\forall v_1\forall v_2((x\neq v_1\wedge y\neq v_2)\rightarrow(V(v_1,v_2)\leftrightarrow W(v_1,v_2)))\wedge\neg W(x,y))))))$

(A20) $\forall V(\exists xV(x)\rightarrow\exists W\exists y(\forall z(y\neq z\rightarrow(V(z)\leftrightarrow W(z)))\wedge V(y)\wedge\neg W(y)))$

Theorem 1: $\forall x(\frown^{1,3}(0,x,s)=T\rightarrow x=1)$

Proof. Assume $\frown^{1,3}(0,c,s)=T$. Hence by (A3) and (A4) 0 and c bear a relation with extension s. Hence by (D2) and (A16) there is some property P such that c is the number of objects with that property and 0 is the number of objects that have that property other than some object that has that property.

It remains to show that P must be equinumerous with C_1. To do so, we must show that there exists a relation that correlates P with C_1 such that its converse correlates C_1 with P.

There must exist a property Q holding of all and only objects with property P other than some particular object such that the number of Q's is 0. Hence there exists a relation R such that R correlates C_0 with Q and its converse correlates Q with C_0. Since every object is self-identical, any single-valued binary relation trivially correlates C_0 with Q by (A6). But if any object had property Q, then the converse of R would have to map that object to an object with property C_0, and hence no object has property Q. Hence there is a unique object a with property P.

[28] Once we define the natural numbers, repeated applications of (A17) can be used to prove that e.g. $\forall P(\#\ulcorner x\urcorner P(x)=3\leftrightarrow\exists y\exists z\exists w(P(y)\wedge P(z)\wedge P(w)\wedge y\neq z\wedge y\neq w\wedge z\neq w\wedge\forall v(P(v)\rightarrow(v=y\vee v=z\vee v=w)))))$.

If x is an element of the extension of C_1, then x is the extension of C_0, i.e. x is the empty extension. But by (A17) there exists a single-valued relation S whose converse is a single-valued relation such that S holds between a and the empty extension (simply modify the identity relation, mapping a to the empty extension and the empty extension to a). S thereby establishes equinumerosity between P and C_1, thus completing the proof.

□

Theorem 2: $\forall x(\#x = 1 \rightarrow \exists y(y \frown x = T))$

Proof. Assume $\#c = 1$. Hence there exists a property P with extension c such that the number of P's is 1. Hence C_1 is equinumerous with P. Hence there exists a single-valued relation R with a single-valued converse that correlates C_1 with P such that its converse correlates P with C_1. But if P had no instances, then by (A6) C_1 couldn't have any instances either. But the extension of C_0 is an instance of C_1. Hence P has an instance.

□

Theorem 3: $\forall x \forall y \forall z((\#x = 1 \wedge y \frown x = T \wedge z \frown x = T) \rightarrow y = z)$

Proof. Assume $\#c = 1 \wedge a \frown c = T \wedge b \frown c = T$. Hence there is a property P such that c is the extension of P, a has property P, and b has property P. Furthermore, C_1 is equinumerous with P. But then there exists a single-valued relation R that correlates C_1 with P with a single-valued converse that correlates P with C_1. The converse of R must hold between a and the empty extension, since that is the only element of the extension of C_1. But then it also must hold between b and the empty extension. But then R holds between the empty extension and a and between the empty extension and b. But R is single-valued, and hence $a = b$.

□

Theorem 4: $\forall x((\forall y(y \frown x = T \rightarrow \forall z(z \frown x = T \rightarrow y = z))) \rightarrow ((\exists w(w \frown x = T) \rightarrow \#x = 1)))$

Proof. Assume $\forall y(y \frown a = T \rightarrow \forall z(z \frown a = T \rightarrow b = z))$ and $\exists w(w \frown a = T))$. Hence a is the extension of a property P such that there exists a unique instance of P. Call this instance a. It remains to show that C_1 is equinumerous with P. But this follows from axiom (A17) in the same manner as in the proof of Theorem 1 (note that equinumerosity can easily be shown to be symmetric via the converse of the correlating relation).

□

Theorem 5: $I(\triangleleft s) \wedge I(s)$

Proof. Let R be a relation with extension s (e.g. R_6). Suppose $R(a, b)$ and $R(a, c)$. We want to show that $b = c$. a is the number of P's for some property P and is also the number of W's for some property W. b is the number of Q's for some property Q, and c is the number of S's for some property S. Furthermore, P holds of all and only objects with property Q other than some particular object o_1 and W holds of all and only objects with property S other than some particular object o_2 (not necessarily the same as o_1).

Since a is both the number of P's and the number of W's, there exists a relation R correlating P and W with the appropriate converse. It remains to show that there is some relation correlating Q and S with the appropriate converse. Suppose there is some x such that $R(o_1, x)$ and some y such that $R(y, o_2)$. Apply (A17) to define a relation just like R except it holds between o_1 and o_2 and between y and x. This new relation is the needed correlating relation between Q and S.

Now suppose no such x or y exists. Apply (A18) to construct a new relation just like R except it holds between o_1 and o_2. If x exists but not y, apply (A19) to drop the relation between o_1 and x and then (A18) to construct another new relation that holds between o_1 and o_2. If y exists and x does not, apply (A19) and (A18) in the same way. In each case, the final relation is the needed correlating relation between Q and S.

Now suppose $R(b, a)$ and $R(c, a)$, for different choices of a, b, c. The proof is analogous.

□

Theorem 6: $\forall x(\exists y \# y = x \to (\neg x = 0 \to \exists z \frown^{1,3} (z, x, s) = T))$

Proof. Assume there exists a b such that $a = \#b$ and $a \neq 0$. Since $a \neq 0$, b is nonempty. So, apply (A20) to a property with extension b, resulting in a property with some extension c holding of all and only elements of b except some particular element. It remains to show $\exists x(x = \#c)$. Simply use the identity relation as the needed single-valued relation to complete the proof.

□

Note that we have *not* established that the successor of any number is distinct from that number. It certainly isn't. Consider a bijection between the set of all natural numbers except for 0 and the set of all natural numbers. We *will* be able to establish this about the successor relation applied to the natural numbers themselves, however. A contemporary mathematician may prefer a different definition of the successor relation.

One advantage of Frege's approach is that we are explicitly establishing the relation between the natural numbers and properties, in accordance with the philosophical thesis that a statement of number contains an assertion about a property.

Frege next turns to the task of defining the relation "one object follows after an object in a series". We'll first define a relation R_7 and then the function $\star$:

(A21) $\forall x, y, z(R_7(x, y, z) \leftrightarrow (\exists W(z = \ulcorner w, v \urcorner W(w, v)) \wedge (\forall V((\forall v_1(V(v_1) \to \forall v_2(\frown^{1,3} (v_1, v_2, z) = T \to V(v_2)))) \to ((\forall v_3(\frown^{1,3} (x, v_3, z) = T \to V(v_3))) \to V(y)))))$

This relation holds of x, y, z exactly if z is the extension of a binary relation W and for any property V, if whenever V holds of v_1 it follows that any v_2 that v_1 bears W to also has property V, then if for any v_3 such that x bears W with v_3, v_3 has property V, then y has property V. That is, y follows x in the z-series. In effect, this relation holds whenever from any property being transmitted down the z-series and x being such that bearing the relation to an object implies that object has that property, it follows that y has that property.

(A22) $\forall x(\star(x) = \ulcorner y, z \urcorner R_7(y, z, x))$

So, $\frown^{1,3} (x, y, \star z) = T$ implies y follows x in the z-series.

Frege points out that $\frown^{1,3} (x, y, \star z) = T \vee x = y$ implies that either y follows x in the z-series or $x = y$. Frege calls this y belonging to the z-series starting with x or x belonging to the z-series ending with y (depending on whether x or y is held fixed). Frege uses this to define another function similar to $\star$, which we'll define with R_8 and †:[29]

(A23) $\forall x, y, z(R_8(x, y, z) \leftrightarrow (\frown^{1,3} (x, y, \star z) = T \vee x = y))$

(A24*) $\forall x(\dagger(x) = \ulcorner y, z \urcorner R_8(y, z, x))$

Frege says n is a *finite* cardinal number (natural number) if n belongs to the cardinal number series starting with 0. As is well-known, this property is not definable in first-order logic. Let's define this property explicitly:

(D3) $\forall x(N(x) \leftrightarrow R_8(0, x, s))$

Frege ends the section with another theorem stated without proof. The theorem states that if n belongs to the series beginning with 0 defined by the successor relation, then the number of P's, where P is the property of belonging to the s-series ending with n, is the successor of n.[30] I will sketch a proof below.

Let's first establish that 0 doesn't follow anything in the successor series. To do so, I'll define the property P_0 of being a cardinal number containing a non-empty extension. I'll then establish that if y is the successor of x and $y \neq x$, then the number of the property of belonging to the s-series ending with y is the successor of the number of the property of belonging to the s-series ending with x. To do this, I'll define the relation R_9 holding between z and w if z is the number of a property holding of less objects than some property that w is the number of ("less" here in the sense characterized by equinumerosity).

I'll also define the relation R_{10} holding between x and y if y has a predecessor belonging to the s-series beginning with x. This will be used in the proof of Lemma 2.

[29] For our purposes the † function is unnecessary since Frege always uses it in a way that could be replaced by our R_8.

[30] Compare a popular set-theoretic definition of the natural numbers: 0 is the empty set and any $n > 0$ is the set $\{0, .., n-1\}$.

To prove the theorem, I'll define the property P_1 of being an x such that x's successor is identical to the number of the property of belonging to the s-series ending with x. Other properties and relations besides these four could have been chosen as well.

(A25) $\forall x(P_0(x) \leftrightarrow (\exists z \exists y((\#z = x) \wedge (y \frown z = T))))$

(A26) $\forall x \forall y(R_9(x,y) \leftrightarrow \exists P \exists Q(x = \#\ulcorner z \urcorner P(z) \wedge y = \#\ulcorner z \urcorner Q(z) \wedge \forall V((w = \ulcorner v, z \urcorner V(v,z) \wedge I(w)) \rightarrow \exists v(Q(v) \wedge \forall z(V(v,z) \rightarrow \neg P(z))))))$

This relation holds if x is the number of some property and y is the number of some property such that for any single-valued binary relation there is some instance of y's property that doesn't bear the relation to an instance of x's property.

(A27) $\forall x \forall y(R_{10}(x,y) \leftrightarrow \exists z(\frown^{1,3}(z,y,s) \wedge R_7(x,z,s)))$

This relation holds between x and y if y has a predecessor that belongs to the s-series beginning with x.

(A28) $\forall x(P_1(x) \leftrightarrow \exists y(\frown^{1,3}(x,y,s) = T \wedge y = \#(\ulcorner z \urcorner R_7(z,x,s))))$

Lemma 1: $\forall x \frown^{1,3}(x, 0, \star s) \neq T$

Proof. P_0 is transmitted across the s-series, since the successor of any number is the number of some property instantiated by one additional object. In particular P_0 holds of the successor of x, if it exists. But $\neg P_0(0)$. This completes the sketch. □

Lemma 2: $\forall x \forall y((\frown^{1,3}(x,y,s) = T \wedge x \neq y) \rightarrow \frown^{1,3}(\#(\ulcorner z \urcorner R_8(z,x,s)), \#(\ulcorner z \urcorner R_8(z,y,s)), s) = T)$

Proof. Suppose y is the successor of x and $y \neq x$. It remains to show that the property of being an element of the s-series ending with x holds of all but one instance of the property of being an element of the s-series ending with y, and in particular that adding x's successor y to the series adds one additional element.

Consider the property "following x in the s-series". This property is transmitted down the s-series and also holds of y's successor, but it does not hold of x. It does not hold of x because one can restrict the relation $R_9(z,w)$ to a property by letting $z = x$, and this property transmits down the s-series and holds of y but it does not hold of x. Hence y is not an element of $\ulcorner z \urcorner R_8(z,x,s)$.

If z is a member of the s-series ending with x, then z is also a member of the s-series ending with y (any transmitting property will transmit from x to y). Suppose $z \neq y$ and z is a member of the s-series ending with y.

If $z = x$ then trivially z is a member of the s-series ending with x, so suppose $z \neq x$. But then if z transmits properties across the s-series to y, then z must also transmit properties to y's predecessor x, since y has a unique predecessor by Theorem 5. To see this, note that the property of having a predecessor that belongs to the s-series beginning with z is transmitted across the s-series and holds of the successor of z (restrict R_{10} to define this property). However, if x does not belong to the s-series beginning with z, then this property doesn't hold of y. Hence z is a member of the s-series ending with x. This completes the proof.

□

Theorem 7: $\forall n(\frown^{1,3} (0, n, \dagger s) = T \rightarrow (\frown^{1,3} (n, \#(\ulcorner x \urcorner R_8(x, n, s)), s) = T))$

Proof. Assume $\frown^{1,3} (0, n, \dagger s) = T$. So n belongs to the series beginning with 0 defined by the successor relation. We need to show that the number of P's, where P is the property "belonging to the s-series ending with n", is the successor of n.

Note that the property of being a cardinal number, which we defined above via the predicate symbol C (A14), is transmitted across the s-series. Therefore n is a cardinal number.

Suppose $n = 0$. Then n belongs to the series beginning and ending with 0 defined by the successor relation. But $\ulcorner x \urcorner R_8(x, n, s)$ only contains 0 by Lemma 1. Use (A17) to modify the identity relation and establish equinumerosity with $C1$, establishing that $\#(\ulcorner x \urcorner R_8(x, n, s)) = 1$. The result then follows from Theorem 1.

Suppose $n \neq 0$. $P_1(1)$ given Lemma 2 and that $\#(\ulcorner x \urcorner R_8(x, 0, s)) = 1$, as we have just shown. It suffices to show that P_1 is transmitted across the s-series. Suppose x has P_1. Then the successor of x is the number of the property Q_1 of belonging to the s-series ending with x. Call the successor y, and call y's successor z. If $x = y$ then the result is trivial since then $y = z$ as well, hence assume $x \neq y$. By Lemma 2, $z = \#(\ulcorner w \urcorner R_8(w, y, s))$. This completes the proof.

□

We have thereby characterized the natural numbers as the numbers of particular properties.

5.4.2 Proving the Dedekind-Peano Axioms

To prove the Dedekind-Peano axioms, we will add an additional closure principle (A29) to our system such that for any property P there exists a property Q that holds of an object exactly if either that object has P or that object is not a natural number. The Dedekind-Peano Axioms are then straightforward consequences of our earlier results.

(A29) $\forall P \exists Q \forall x(Q(x) \leftrightarrow (P(x) \vee \neg R_8(0, x, s)))$

(DP1) For every natural number n, the successor of n is a natural number.

Proof. Since we defined natural numbers n as objects belonging to the s-series beginning with 0, for for any $n > 0$ and any property transmitted across the s-series that is also held by 1, n must have that property. But this property transmits across the s-series, so the successor of n also has this property. Hence the successor of n is also an element of the s-series beginning with 0.

The existence of a successor is a consequence of Theorem 7.

□

(DP2) For all natural numbers n and m, $n = m$ if and only if the successor of $n =$ the successor of m.

Proof. Immediate consequence of Theorem 5.

□

(DP3) 0 does not have a predecessor.

Proof. Suppose 0 has a predecessor p. Then there is some property P such that p is the number of P's except some particular object and 0 is the number of P's. But if 0 is the number of P's, then no object has P, since there could be no relation that correlates a property held by some object with C_0.

□

(DP4) If P is a property such that $P(0)$ and for every natural number n, $P(n)$ implies P holds of the successor of n, then $P(n)$ is true for every natural number n.

Proof. Recall the natural numbers are defined as the objects belonging to the s-series beginning with 0. If P transmitted across *every* s-series the proof would be trivial, but we can only assume that P transmits across this particular s-series.

Instead, we will use our closure axiom (A29) to define a new property Q such that $Q(x) \leftrightarrow (P(x) \vee \neg R_8(0, x, s))$. That is, Q holds of an object iff either that object has P or that object is not a natural number. $Q(1)$ since $P(1)$ since $P(0)$.

Suppose $Q(b)$. Then either $P(b)$ or b is not a natural number. If b is a natural number, then its successor c has P as well by assumption. If b is not a natural number, then if it has a successor c, c is not a natural number either. This is because if c were a natural number, then it would either be 0 or its predecessor would also be a natural number (this follows analogously to our proof of Lemma 2: the property of having a predecessor belonging to the s-series beginning with 0 transmits down the s-series and holds of 1). But c can't be 0 because then b couldn't exist. Hence $Q(c)$.

Hence Q is transmitted across *every* s-series. But then every natural number has Q, given $Q(1)$. Hence every natural number has P.

□

5.5 The Consistency of our Formal Theory of Arithmetic

We will follow the standard procedure of exhibiting a model of our axioms. However, we will not use ZFC to construct this model, since our system violates the axiom of foundation (*any* value-range provides a counterexample). Hence, we will use another set theory to construct our model.

I have chosen to use the set theory generated from removing the Axiom of Foundation from ZFC and replacing it by the Anti-Foundation Axiom of Peter Aczel (1988):

(AFA) Every accessible pointed directed graph corresponds to a unique set.

Call our theory FA.

Theorem 8: FA is consistent.

Proof. As described above, we will define an accessible pointed directed graph piece-by-piece in order to prove the existence of a model $\mathcal{M}$ of the axioms of FA. The needed interpretation function $\mathcal{A}$ will become clear as we proceed. I will focus on the definitions of the domain of objects O and the domain of functions G. I identify any element of G with its value-range. Hence, G is a subset of O and $\circ$ is the identity function from G to O.

O will contain two additional elements T and F, where F is the empty set $\{\}$ and T is the Quine atom $T = \{T\}$. Hence, with respect to our graph, F is the unique node with no edges leading from it and T is the unique node with a single edge leading from it to itself. All elements of G and all other elements of O will be sets of ordered pairs.

I will describe a construction procedure for building each element of G piece-by-piece, which will correspond to adding nodes and directed edges to our graph. This construction procedure is in-principle recursive, consisting of finitely many applications of distinct sets of countably many steps. Since every function is defined over every element of the domain, each time we introduce a new function we will need to add the appropriate ordered pairs to each function that has already been defined. We will also need to respect our various closure principles. When I speak of "adding an ordered pair to a set", what I mean with respect to our graph is that we should add an edge from that set to that ordered pair.

Each needed ordered pair (a, b) corresponds to a node in our graph with two edges leading to nodes $\{a\}$ and $\{a, b\}$, the first of which has a single edge leading to node a and the second of which has two edges leading to node a and node b, respectively. Let's define an ordered triple (a, b, c) as the ordered pair $((a, b), c)$ and an ordered quadruple (a, b, c, d) as the ordered pair $((a, b, c), d)$. Each element of G will correspond to a node with edges leading to each ordered pair contained within that element.

Our first element of G is the identity relation $=$, which will consist of a set of ordered pairs, each of whose first component is an ordered pair and second component is either T or F. There will be one first component corresponding to every combination of elements of O. For each such first component (a, b), the second component will be T if $a = b$ and F otherwise. Since $=$ has an extension (itself), we will begin the construction of $=$ by adding the three ordered pairs $((=,=), T), ((T, T), T)$, and $((F, F), T)$ to $=$, followed by adding every other combination of pairs of elements of the domain as first components with second component F (e.g. $((T, F), F)$).

We will close our domain G under the operations defined in axioms $(L1)$ and $(L2)$ of our derivation system. This will force our model to have an infinite domain and will thereby force all elements of G to themselves be infinite sets. Currently, G contains the single element $=$. Since $=$ is a binary relation, $(L2)$ implies the existence of three properties given the current state of our domain O: "being identical to T", "being identical to F", and "being identical to $=$". Each property, being identical to its own extension, will require the addition of the obvious additional ordered pairs to $=$. Given $(L2)$, this implies the existence of another three properties corresponding to being identical to each of those extensions. This will require adding the needed ordered pairs to $=$, resulting in new properties, etc.

Let's next add the function $\imath$ to G. Recall that $\imath$ maps each element of O that is an extension with a unique element contained in O to that unique element and maps any other element of O to itself. Hence, for each property that we just defined via $(L2)$ and $=$, $\imath$ will contain an ordered pair consisting of that property and the unique element of its extension. For every other element of O, $\imath$ will contain an ordered pair consisting of that element and itself. As usual, we must now add the ordered pair $((\imath, \imath), T)$ to $=$, followed by adding ordered pairs corresponding to every other combination of elements of O with $\imath$. Furthermore, we must add the ordered pair $(\imath, F)$ to each property defined via $(L2)$ and $=$. Finally, $(L2)$ and $=$ imply the existence of an additional property "being identical to $\imath$" with its own extension, which forces the addition of the needed ordered pairs to each element of G, followed by adding a new property of being identical to that extension, etc.

$(L3)$ is immediate from our definitions of T and F.

Basic Law V, which has been split into axioms (Va) and (Vb), is trivially true in our model. This is because we have identified every function with its value-range.

Axiom (VI) is a straightforward consequence of our definition of $\imath$.

We can now move to our arithmetical axioms $(A1)$ - $(A29)$. I will describe the relation R_0 defined by $(A1)$ in detail in order to give the reader a general idea of the approach.

$(A1)$ defines the three-place relation R_0, where R_0 holds of x, y, and z if y is the value-range of a function that maps x to z or y is the extension of a property that x has and $z = T$ (i.e. the output of that property on input x). So, for example, the extension e of the property "being identical to T" is such that $((T, e, T), T)$ is an element of R_0, since T has this property. For any elements x of O other than T and any elements y of O, $((x, e, y), F)$ is an element of R_0.

Hence, we must add the appropriate ordered pairs to R_0 corresponding to the extensions of the properties we defined via $(L2)$ and $=$, in addition to ordered pairs corresponding to $\imath$ applied to each element of O (e.g. $((T, \imath, T), T)$). These will be the current ordered pairs with second component T. All other ordered pairs will consist of an ordered triple as first component and F as second component, where the ordered triples are every combination of elements of O that did not already appear in the previous step. Add ordered pairs whose first component includes the extension of R_0 itself as a component in the obvious way.

Given ($L2$), we will close G under restrictions of R_0 via holding one of its three arguments fixed at any element of O, and furthermore we must close G under the restriction of any binary relation resulting from restricting R_0. As usual, we must then allow $=$ to be restricted by the extensions of any of these properties or relations, which will result in the definitions of new properties, etc. Furthermore, we must add the appropriate ordered pairs to every element of G. One could of course define this procedure as a single countable list of steps.

($A2$) defines a two-place function $\frown$ based upon R_0. $a \frown b = c$ if either b is the value-range of a one-place function mapping a to c or b is not the value-range of a one-place function and c is the empty extension (the set of all ordered pairs consisting of an element of O as first component and F as second component). Since we have identified functions with their value-ranges and hence properties are not distinguished from functions in general, if b is the extension of a property that a does *not* have then c will be F. This is not true in the general case (c could be the empty extension in other models).[31]

So, we must simply add the needed ordered pairs to $\frown$, close G under the closure operations defined above (including ($L1$), since $\frown$ is a two-place function), and add the new ordered pairs to each element of G. This is another example of countably many required steps.

($A3$) and ($A4$) are analogous to ($A1$) and ($A2$), meant to introduce the function $\frown^{1,3}$ (grounded on R_1) mapping x, y, and z to w if w is the output of a two-place function with extension z to inputs x and y.

($A5$) introduces a property I holding of the extensions of binary relations if they are single-valued, i.e. $R(x,y)$ and $R(x,z)$ implies $y = z$. Our current binary relations include the identity relation, which is single-valued, and various restrictions of R_0 and R_1. So, we must add the appropriate ordered pairs to I, close G under the appropriate closure operations, and add the appropriate ordered pairs to all elements of G.

The general strategy should now be clear. The reader is welcome to check the cases that I do not explicitly discuss here. There are six remaining closure principles:

($A8$) introduces a new closure principle, requiring G to contain the converse of any binary relation in G.

($A17$) - ($A19$) introduce additional closure principles that will need to be respected, closing bijective binary relations under exchange of a pair of elements

[31] It is worth noting that, as Quine pointed out in (1954, pg. 155), there is a contradiction lurking nearby. Consider the formula $x \frown x = y$. One might think that this formula defines a relation between x and y. Call this relation $R(x,y)$. Now restrict R at second argument with F, resulting in the property $R(x,F)$. This property seems to hold of objects exactly if those objects are either extensions which are not elements of themselves or are value-ranges of functions mapping their own value-ranges to F. Let the extension of this property be e. Now consider the sentence $e \frown e = T$. If e is an element of itself, then e cannot be an element of itself, and if e is not an element of itself, then e must be an element of itself. Hence, we cannot allow the definition of such a relation without an explicit definition statement. It is also worth noting that given our semantics, $\ulcorner x \urcorner x \frown x = F$ ignores the function symbol $\frown$ and instead refers to the extension of the property "being identical to F".

with each other (i.e. from $R(a,b)$ and $R(c,d)$ to $R^*(a,d)$ and $R^*(b,c)$), removal of a pair of elements from the relation, and addition of a pair of elements to the relation. $(A20)$ closes the domain of nonempty properties under the existence of properties holding of one less element.

$(A29)$ closes the domain of properties P under the existence of properties Q such that Q holds of an object exactly if either that object has P or that object is not a natural number.

This sketch of an in-principle recursive procedure results in the construction of an accessible pointed directed graph with countably many nodes and edges. The corresponding set is our needed model $\mathcal{M} = (O, G, \mathcal{A}, \circ)$ of FA. Hence, FA is consistent.

$\square$

References

[1] Aczel, Peter (1988), *Non-Well-Founded Sets*, CSLI Lecture Notes 14. Stanford, CA: Stanford University, Center for the Study of Language and Information.

[2] Ebbinghaus, H.D., J. Flum, and W. Thomas (1994), *Mathematical Logic*, New York, NY: Springer.

[3] Frege, Gottlob (1884), *Die Grundlagen der Arithmetik*, translated by J. L. Austin as *The Foundations of Arithmetic*, Northwestern: Basil Blackwell, 1980.

[4] Frege, Gottlob (1893), *Grundgesetze der Arithmetik: Band I*, translated by P. Ebert and M. Rossberg as *Basic Laws of Arithmetic: Volume I*, Oxford: Oxford University Press, 2013.

[5] Frege, Gottlob (1903), *Grundgesetze der Arithmetik: Band II*, translated by P. Ebert and M. Rossberg as *Basic Laws of Arithmetic: Volume II*, Oxford: Oxford University Press, 2013.

[6] Quine, W.V.O. (1954), "On Frege's Way Out", reprinted in: Quine (1966), pp. 146-58.

[7] Quine, W.V.O. (1966), *Selected Logic Papers*, Cambridge, MA: Harvard University Press.