I

Frege, Russell, and After

The great logician Gottlob Frege wrote three books, each representing a stage in a grand program for providing a logical foundation for arithmetic and higher mathematics. His *Begriffsschrift* (1879) introduced a comprehensive system of symbolic logic. The first half of his *Grundlagen* (1884) offered a devastating critique of previous accounts of the foundations of arithmetic, while the second half offered an outline of his own ingenious proposed foundation. The two volumes of the *Grundgesetze* (1893, 1903) filled in the technical details of his outline using his logical symbolism, and extended the project from arithmetic, the theory of the natural numbers, to mathematical analysis, the theory of the real numbers. Unfortunately, just as the second volume of the *Grundgesetze* was going to press Bertrand Russell discovered a contradiction in Frege's system.

Subsequent work in logic and foundations of mathematics largely bypassed the *Grundgesetze* until a couple of decades ago, when philosophers and logicians took a new look at Frege's inconsistent system, and recognized that more can be salvaged from it than had previously been thought. In these last years amended and paradox-free versions of Frege's system have been produced; substantial portions of classical mathematics have been developed within such systems; and a number of workers have claimed philosophical benefits for such an approach to the foundations of mathematics.

The thought underlying the present monograph is that however wonderful the philosophical benefits of Frege-inspired reconstructions of mathematics, the assessment of the ultimate significance of any such approach must await a determination of just how *much* of mathematics can be reconstructed, without resort to *ad hoc* hypotheses, on that approach. What is undertaken in the pages to follow is accordingly a survey of various modified Fregean systems, attempting to determine the scope and limits of each. The present work, though entirely independent of Burgess and Rosen (1997), is thus in a sense a companion to the survey of various nominalist strategies in the middle portions of that work. As in that earlier survey, so in the present one, familiarity with intermediate-level logic is assumed. Boolos, Burgess, and Jeffrey (2002) contains more than enough background material, but neither that nor any other specific textbook is presupposed.

Every strategy, if it is to be consistent, must involve some degree of departure from Frege; but some of the approaches to be surveyed here stay much closer to Frege's own strategy than do others. It is sometimes suggested that the closer one stays to Frege, the greater the philosophical benefits. It is not my aim in the present work to argue for or against such claims. What I do insist is that any philosophical gains must be weighed against mathematical losses. For the survey to follow shows that some approaches yield much more of mathematics than others, and it often seems that the less one keeps of Frege, the more one gets of mathematics. Nonetheless, even in the last system to be considered here, which yields all of orthodox mathematics and more also, there remains one small but significant ingredient of Fregean ancestry.

As a necessary preliminary to the survey of attempts to repair Frege's system, that system must itself be reviewed. The underlying logic of the *Begriffsschrift*, the assumption added thereto in the *Grundgesetze* for purposes of developing mathematics, that development itself, the paradox Russell found in the system, and Russell's own attempts to repair it, must each be briefly examined, and the mathematical and philosophical goals a modified Fregean project, or for that matter any present-day foundational program, might set itself must be briefly surveyed.

1.1 Frege's Logic

While Frege is honored as a founder of modern logic, his system will not look at all familiar to present-day students of the subject. To begin with, Frege uses a non-linear notation that no subsequent writer has found it convenient to adopt, and that will not be encountered in any modern textbook. Since the present work is anything but an historical treatise, the notation will be ruthlessly modernized when Frege's system is expounded here.

But even when the notation is modernized, Frege's *higher-order* logic has a grammar that, though still simpler by far than the grammar of German or English or any natural language, is appreciably more complex than the grammar of the *first-order* logic of present-day textbooks. Nonetheless, after Frege's unfamiliar underlying grammatical and ontological assumptions have been expounded, only a few further explanations should be required to enable the reader familiar with first-order logic to understand higher-order logic.

Let us begin, then, with Frege's grammar. For Frege there are two grammatical categories or grammatical types of what he calls *saturated* expressions. The first, here to be called N, or the category of *names*, includes proper names such as "Plato," but also singular definite descriptions such as "the most famous student of Socrates" that are free from indexicals and designate an *object* (which may be a person or place rather than a "thing" in a colloquial sense), independently of context. The second, here to be called S, or the category of *sentences*, includes declarative sentences such as "Plato was the most famous student of Socrates" that are free from indexicals and have a *truth-value*, true or false, independently of context.

In addition there are many types of *unsaturated* expressions, with one or more gaps that, if filled in with an expression or expressions of appropriate type(s), would produce an expression of type N or S. Those that, thus filled in, would produce expressions of type S rather than type N will be of most interest here, and these may be called *predicates* in a broad sense. In a notation derived from the much later writers Kasimierz Ajdukiewicz and Yehoshua Bar-Hillel, an expression with k blanks in it, that if filled in with expressions of types $T_1 \ldots, T_k$ will produce an expression of type S, may be said to be a predicate of type $S/T_1 \cdots T_k$. The simplest case is that of *predicates* in the narrow sense of expressions of type S/N. Some other cases are shown in table A at the back of the book.

The label *relational predicate* will be used for the two-place, three-place, and many-place types S/NN, S/NNN, and so on, with the number of places being mentioned explicitly when it is greater than two. The label *higher predicate* will be used for the second-level, third-level, and higher-level types S/(S/N), S/(S/(S/N)), and so on, with the number of the level being mentioned explicitly if it is higher than the second. Similarly with the label *higher relational predicate*, which even with just two places and even at just the second level covers a variety of types, including not only S/(S/N)(S/N) as shown in the table, but also, for instance, S/(S/NN)(S/NN), and such mixed types as S/(S/N)(S/NN) and even S/N(S/N). It is an instructive exercise to look for natural language examples illustrating such possibilities.

In Frege (1892), which after his three books is its author's most famous work, Frege introduced a distinction between the *sense* expressed and the *referent* denoted by an expression. The reader will not go far wrong who thinks of what Frege calls the "sense" of an expression of whatever type as roughly equivalent to what other philosophers would call its "meaning." What the "referent" of an expression is to be understood to be varies from grammatical type to grammatical type.

In the case of a proper name or singular definite description of type N, the referent is the object designated, the thing bearing the name or answering to the description. Clearly two expressions of type N, for instance, "the most famous student of Socrates" and "the most famous teacher of Aristotle," can have different senses even though they have the same referent—in this instance, Plato. Expressions with different senses but the same referent provide different "modes of presentation" of the same object.

The sense of a sentence of type S Frege calls a *thought*, and the reader will not go far wrong who thinks of what Frege calls a "thought" as roughly equivalent to what other philosophers call a "proposition." The referent of a sentence of type S Frege takes to be simply its truth-value. Obviously two sentences of type S, for instance, "Plato is the most famous student of Socrates" and "Plato is the most famous teacher of Aristotle," or "Plato is a featherless biped" and "Plato is a rational animal," can have different senses, though they have the same referent—in these instances, the truth-value *true*.

So much for the referents of complete or saturated expressions. As for the referents of incomplete or unsaturated expressions of types N/… or S/…, they are supposed to be incomplete, like the expressions themselves, containing gaps that when appropriately filled in will produce an object or a truth-value. The referents of expressions of type N/… Frege calls *functions*, and the referents of expressions of type S/…, that is, the referents of predicates, he calls *concepts*. Corresponding to the different grammatical types of predicates are different ontological types of concepts, including concepts of the narrowest, first-level, one-place kind, but also *relational concepts* and *higher concepts*.

For a concept of type S/N, if filling it in with a certain object produces the truth-value *true*, then the object is said to *fall under* the concept. For instance, assuming for the sake of example that Plato may truly be called wise, since the concept denoted by "… is wise," which is to say the concept of being wise, when filled in by the object denoted by "Plato," which is to say the object Plato, produces the truth-value denoted by "Plato is wise", which is to say the truth-value *true*, it follows that the object Plato falls under the concept of being wise. The same terminology is used in connection with other types of predicates: since Socrates taught Plato, the pair Socrates and Plato, in that order, fall under the relational concept of having taught; since Plato is an example of someone who is wise, the concept of being wise falls under the higher concept of being exemplified by Plato.

Two concepts are called *coextensive* if they apply to the same items, or in other words, if whatever falls under either falls under the other. Now the crucial difference between the *referents* of predicates, which is to say concepts, and the *senses* of predicates is that, according to Frege, coextensive concepts are the same. Thus if every featherless biped is a rational animal and vice versa, then though the senses of "… is a featherless biped" and "… is a rational animal" are different, the concept of being a featherless biped and the concept of being a rational animal are the same.

It sounds odd to say so, and the degree of oddity is a measure of the degree of departure of Frege's technical usage of "concept" from the ordinary usage of "concept," which tends to suggest the sense rather than the referent of a predicate. The label "concept" was in fact chosen by Frege well before he recognized the importance of systematically distinguishing sense and reference. By hindsight it seems he might have done well, after recognizing the importance of that distinction, to revise his terminology. I was tempted to substitute "classification" for "concept" in the foregoing short exposition, but have stuck with "concept" because it is still used by most of the writers with whose views I will be concerned.[1]

In Fregean terminology, then, to the grammatical categories of names, sentences, and predicates there correspond the ontological categories of objects, truth-values, and "concepts." The formal language of Frege's higher-order logic is more complex than the formal languages of the first-order logic expounded in present-day textbooks in that it makes provision for predicates of all types, denoting concepts of all types.

So much for the grammar behind the logic. Turning to logic itself, modern textbooks introduce the student to the notions of a first-order *language*, with the symbol = for identity, and usually other *non-logical* symbols (*n*-place relation symbols, including 0-place ones or sentence symbols, and *n*-place function symbols, including 0-place ones or constants). Also introduced are *rules of formation*, and the notion of a *term* (built up from variables and constants using function symbols), *atomic formula* (obtained by putting terms in the places of relation symbols), and *formula* (built up from atomic formulas using the logical operators ~, &, ∨, →, ↔, ∀, ∃), along with the ancillary notions of *free* and *bound* variables, and *open* and *closed* formulas (with some variables free and with all variables bound, respectively).[2]

Also introduced are certain *rules of deduction*, which may take widely different forms in different books—the different formats for deductions used in different books going by such names as "Hilbert-style" and "Gentzen-style" and "Fitch-style"—but which lead in all books to equivalent notions of what it is for one formula to be *deducible* from others, and hence to equivalent notions of a *theory*, consisting of all the formulas, called *theorems* of the theory, that are deducible from certain specified formulas, called the non-logical *axioms* of the theory.[3]

Familiarity with all these notions must be presupposed here, but certain conventions that are not covered in all textbooks may be briefly reviewed. To begin with, it proves convenient in practice, when writing out formulas, to make certain departures from what in principle ought to be written. In particular, for the sake of conciseness and clarity certain *defined* symbols are added to the *primitive* or official symbols. The simplest case is the introduction of $\neq$ for *distinctness* by the usual definition, as follows:

(1) $$x \neq y \leftrightarrow \sim x = y$$

What it means to say that $\neq$ is "defined" by (1) is that officially $\neq$ isn't part of the notation at all, and that the left side of (1) is to be taken simply as an unofficial abbreviation for the right side of (1). The "definition" (1) is thus not a substantive axiom, but merely an abbreviation for a tautology of the form $p \leftrightarrow p$. The slash notation for negation may also be used with certain non-logical relation symbols when their shape permits.

Another slightly less simple case, often not covered in textbooks, is that of $\exists!$ for *unique existence*, with two of the several equivalent usual definitions being as follows:

(2a) $\exists!x\phi(x) \leftrightarrow \exists x\phi(x)\ \&\ \forall x_1 \forall x_2(\phi(x_1)\ \&\ \phi(x_2) \rightarrow x_1 = x_2)$

(2b) $\exists!x\phi(x) \leftrightarrow \exists x \forall y(\phi(y) \leftrightarrow y = x)$

Here the first and second conjuncts of the conjunction on the right in (2a) are called the *existence* and *uniqueness* clauses, respectively.

Yet another and less simple case may arise when one has assumed as an axiom or deduced as a theorem something of the form $\exists!\phi(x)$. It may then be convenient to "give a name to" this x by introducing a constant c and assuming $\phi(c)$. The assumption $\phi(c)$ is then called an *implicit* definition, and $\exists!\phi(x)$

the *presupposition* of that definition. Any formula $\psi(c)$ can then be regarded indifferently as abbreviating either of the following:

(3a) $\forall x(\forall y(\phi(y) \leftrightarrow y = x) \rightarrow \psi(x))$

(3b) $\exists x(\forall y(\phi(y) \leftrightarrow y = x) \;\&\; \psi(x))$

I say "indifferently" because (3a) can be deduced from (3b) and vice versa using the presupposition $\exists!\phi(x)$. Note that $\phi(c)$ expands into a logical consequence of that presupposition. Note also that if $\psi(c)$ is of the form $\sim\theta(c)$, it is really a matter of indifference whether one first unpacks the abbreviation in $\theta(c)$ and then applies negation, or first applies negation and then unpacks the abbreviation in $\sim\theta(c)$. For the two results, which read as follows:

(4a) $\sim\forall x(\forall y(\phi(y) \leftrightarrow y = x) \rightarrow \theta(x))$

(4b) $\forall x(\forall y(\phi(y) \leftrightarrow y = x) \rightarrow \sim\theta(x))$

are deducible from each other as were (3a) and (3b); and the same holds for other compounds than negation.

Still yet another and even less simple case is the many-place analogue of the abbreviatory convention just discussed. When one has assumed or deduced $\forall y\exists!x\psi(y, x)$, one may "give a name to" this x by introducing a *function symbol* f and assuming $\forall y\psi(y, f(y))$. The unpacking to eliminate f proceeds analogously to the unpacking to eliminate c in the one-place case. And what has just been said about a one-place function symbol f applies also to many-place function symbols.

Bertrand Russell's famous *theory of descriptions* provides a general notation, which I will write as iota ι, that attaches to a variable x and a formula having x as a free variable to form a term that behaves like an n-place function symbol, where n is the number of free variables other than x. In this notation our constant or 0-place function symbol c above would be $\iota x\phi(x)$, while $f(y)$ would be $\iota x\psi(y, x)$. In both, x counts as a bound variable. Contexts containing ι-terms are expanded as indicated above.[4]

Certain abbreviatory conventions apply specifically to the displaying of laws (axioms and theorems) of a theory. These again may be illustrated by the case of identity, where we have the logical law of *indiscernibility of identicals*, which in many textbook presentations is taken as a logical axiom and in others appears as a logical theorem. Indiscernibility may be stated as follows:

(5) $$x = y \rightarrow (\phi(x) \leftrightarrow \phi(y))$$

One of the conventions illustrated by this way of stating the law is that though we allow ourselves to speak in the singular of "the" axiom or theorem of indiscernibility, what one really has here is an axiom *scheme*, meaning a rule to the effect that all formulas of a certain form are to be counted as axioms, or a theorem *scheme*, meaning a result to the effect that all formulas of a certain form are theorems. What is displayed in (5) is not "the" law, but the general form of an *instance* of the scheme, wherein ϕ may be any formula.[5]

Actually, (5) does not yet fully display the general form of an instance of the law, since it is conventional in displaying laws to omit initial universal quantifiers. Thus what an instance really looks like is this:

(5a) $$\forall x \forall y(x = y \rightarrow (\phi(x) \leftrightarrow \phi(y)))$$

And actually, even (5a) still does not yet fully exhibit the general form of an instance of the law, since it is conventional in displaying laws to omit *parameters*, or additional free variables, that may be present. Thus what an instance *really* looks like is this:

(5b) $$\forall u_1 \ldots \forall u_k \forall x \forall y(x = y \rightarrow (\phi(x, u_1, \ldots, u_k) \leftrightarrow \phi(y, u_1, \ldots, u_k)))$$

Like all the other conventions that have been described so far, these conventions about displaying axioms and theorems of first-order logic will apply also to higher-order logic.

THESE CONVENTIONS HAVE BEEN briefly described here, I say, because not all textbooks cover them. There is one further important topic that very few textbooks cover, that of *many-sorted* first-order languages and theories. A many-sorted language is just like an ordinary one, except that there is more than one style of variables. For instance, in a geometrical theory about points and lines, it may be convenient to have one style of variable x, y, z, … for points, and another style of variable ξ, υ, ζ, … for lines. Certain obvious changes in the rules of formation and deduction then have to be made.

As to formation rules and formulas, the usual rules say that $u = v$ is an atomic formula for any variables u and v. In many-sorted logic one has a different identity symbol for each sort of variable. So in our geometrical example we would have atomic formulas of the kinds $x =_{point} y$ and $\xi =_{line} \upsilon$, and not atomic formulas identifying a point and a line. Similarly, for non-logical symbols there may be restrictions as to which sorts of variables can go into which places. These changes affect only the rules for forming terms and atomic formulas. The rules for forming more complex formulas from simpler ones by logical operations remain unchanged. In particular, ∀ and ∃ may be applied to any sort of variable.

As to deduction rules and deducibility, the usual rules for quantifiers allow—in one format or another—the inference from $\forall u\phi(u)$ to $\phi(v)$ and from $\phi(v)$ to $\exists u\phi(u)$ for any variables u and v. But in our geometrical example we would want to allow only inference from $\forall x\phi(x)$ to $\phi(y)$ and from $\forall\xi\psi(\xi)$ to $\psi(\upsilon)$, and not from $\forall x\phi(x)$ to $\phi(\upsilon)$ or from $\forall\xi\psi(\xi)$ to $\psi(y)$. Similarly, for ∃ there are restrictions as to which sorts of variables can be substituted where. These changes affect only the rules of deduction involving quantifiers. The rules for ~, &, ∨, →, ↔ remain unchanged.

In principle, a two-sorted theory can always be replaced by a one-sorted theory with a single style of variables *x*, *y*, *z*, … by introducing a one-place relation symbol P, called a *sortal predicate*, and replacing quantifications ∀x(…) and ∀ξ(…) by ∀*x*(P*x* → …) and ∀*x*(~P*x*→ …), and ∃x(…) and ∃ξ(…) by ∃*x*(P*x* & …) and ∃*x*(~P*x* & …). For instance, a geometric theory about points X and lines ξ can be reduced to a one-sorted theory about points-or-lines *x* by introducing a predicate "is a point," and replacing quantifications "for every point x" and "for every line ξ" by "for every *x*, if *x* is a point, then" and "for every *x*, if *x* is not a point, then." If there is, say, a constant c of sort x in the original language, we need to add explicitly as an axiom, since it is no longer implicit in the notation, that c is a point: Pc. If there is, say, a function symbol † in the original two-sorted language that takes arguments of sort ξ and gives values of sort x, we need to add explicitly as an axiom, since it is no longer implicit in the grammar, that † applied to a line gives a point: ~P*x* → P†*x*. And similarly for function symbols of more places. A similar reduction can be carried out for three-sorted theories, using two sortal predicates P and Q (and speaking of the items of the third sort as the *x* such that ~P*x* & ~Q*x*). But retaining the many-sorted formulation is often in practice more convenient and more illuminating.

For the reader comfortable with first-order logic, including its many-sorted variant version just described, higher-order logic may be introduced as simply one special many-sorted first-order theory, with certain distinctive formulas as axioms, in one special many-sorted first-order language, with certain distinctive relation symbols as primitives. From the perspective of first-order logic, these distinctive primitives and axioms are considered non-logical; from the perspective of higher-order logic, they are considered logical. Whether they are or are not "logical" in a philosophically interesting sense is obviously relevant to assessing the philosophical significance of the development of mathematics within a system based on higher-order logic. But this philosophical issue is irrelevant to the technical definitions of higher-order formula and higher-order deducibility, which as already indicated are the same as for many-sorted first-order logic, which in turn are, apart from some obvious restrictions on which sorts of variables can turn up in which positions, the same as for ordinary, textbook first-order logic, with which it is assumed the reader is familiar.

What the distinctive primitives and axioms of higher-order logic are can almost be guessed from the earlier discussion of Fregean ontology. As to primitives, there are variables of various types T, with the corresponding identity symbols $=_{\mathrm{T}}$, giving rise to such atomic formulas as the following:

$$x =_{\mathrm{N}} y \qquad X =_{\mathrm{S/N}} Y \qquad R =_{\mathrm{S/NN}} S \qquad \mathbf{X} =_{\mathrm{S/(S/N)}} \mathbf{Y}$$

Besides these we will want for each type $\mathrm{T} = \mathrm{S/T}_1 \dots \mathrm{T}_k$ a $(k + 1)$-place relation symbol ∇_{T} to express that a given concept of type T has given items of types $\mathrm{T}_1, \dots, \mathrm{T}_k$ falling under it. These symbols will give rise to such atomic formulas as the following:

$$\nabla_{\mathrm{S/N}} Xx \qquad \nabla_{\mathrm{S/NN}} Rxy \qquad \nabla_{\mathrm{S/(S/N)}} \mathbf{X}X$$

Such is the language of full *higher-order* logic.[6] If we drop everything above second-level concepts, the result is the language of (*polyadic*) *third-order* logic. If we drop everything above first-level concepts, the result is the language of (*polyadic*) *second-order* logic. If we drop all relational concepts of more than two places, the result is the language of *dyadic higher-order* logic. If we drop all relational concepts whatsoever, the result is the language of *monadic higher-order* logic.

This is all that needs to be said about the *official* notion of formula for higher-order logic. Unofficially, $=_{\mathrm{N}}$ is conventionally just written as $=$, while all other $=_{\mathrm{T}}$ are written as $\equiv$. As for ∇, not only are subscripts dropped, but the very symbol itself is not written, so in practice "x falls under X" and "x and y in that order fall under R" and "X falls under $\mathbf{X}$" are written just Xx and Rxy and $\mathbf{X}X$, or sometimes $X(x)$ and $R(x, y)$ and $\mathbf{X}(X)$.

Turning from primitives to axioms, there are just two of these, called *comprehension* and *extensionality*.

$$(6) \qquad \exists X \forall x (Xx \leftrightarrow \phi(x))$$
$$(7) \qquad X \equiv Y \leftrightarrow \forall z (Xz \leftrightarrow Yz)$$

The usual conventions for stating axioms apply here: really (6) is an axiom scheme, the formulas displayed are to be prefixed with universal quantifiers, and there may be parameters.[7] A further convention is that when a formula is stated as an axiom of monadic second-order logic, unless explicitly indicated otherwise, the analogous polyadic and higher-order formulas are also to be taken as axioms. Thus in addition to (6) and (7), the following are also comprehension and extensionality axioms:

$$(6a) \quad \exists R \forall x \forall y (Rxy \leftrightarrow \psi(x, y)) \qquad (6b) \quad \exists \mathbf{X} \forall X (\mathbf{X}X \leftrightarrow \theta(X))$$
$$(7a) \quad R \equiv S \leftrightarrow \forall x \forall y (Rxy \leftrightarrow Sxy) \qquad (7b) \quad \mathbf{X} \equiv \mathbf{Y} \leftrightarrow \forall Z (\mathbf{X}Z \leftrightarrow \mathbf{Y}Z)$$

Some consequences immediately deducible from these axioms may be noted. (7) implies that the X in (6) is unique, and we may give a name to it, calling it «x: $\phi(x)$», read "the concept of being an x such that $\phi(x)$." In Russellian notation «x: $\phi(x)$» = $\iota X \forall x (Xx \leftrightarrow \phi(x))$. We similarly use the notation «x, y: $\psi(x, y)$» and «X: $\theta(X)$».[8]

Extensionality admits of several equivalent formulations. As formulated in (7) above, together with the indiscernibility of identicals it yields the following:

$$(8) \qquad \forall z (Xz \leftrightarrow Yz) \rightarrow (\phi(X) \leftrightarrow \phi(Y))$$

On most contemporary approaches, ≡ is not even included among the official primitives, but rather is regarded as an unofficial abbreviation, in which case (7) is not included among the official axioms, but rather is regarded as the definition of ≡. I will henceforth fall in with this practice. When ≡ is thus taken as defined rather than primitive, it is either (8) that is called the axiom of extensionality, or else the following instance of (8), which together with (6b), actually yields the general scheme (8):

(9) $$\forall z(Xz \leftrightarrow Yz) \rightarrow (\mathbf{X}X \leftrightarrow \mathbf{X}Y)$$

For reasons that will be explained later in this chapter, Frege did not need extensionality in *any* formulation as an axiom for the development of mathematics within his system. Extensionality is needed, however, to express fully his notion of what a concept *is*.

Before closing this section, it will be well to illustrate with an example that plays an important role in Frege's attempt to develop mathematics on a purely logical foundation. To begin with, consider three conditions, each of which may or may not hold for a given relational concept R:

Reflexivity	$\forall x Rxx$
Symmetry	$\forall x \forall y(Rxy \rightarrow Ryx)$
Transitivity	$\forall x \forall y \forall z(Rxy \ \& \ Ryz \rightarrow Rxz)$

Comprehension then gives the existence, and extensionality the uniqueness, conditions in each of the following:

$$\exists! S \forall x \forall y(Sxy \leftrightarrow (x = y \vee Rxy))$$
$$\exists! S \forall x \forall y(Sxy \leftrightarrow (Rxy \vee Ryx))$$
$$\exists! S \forall x \forall y(Sxy \leftrightarrow \forall Z((\forall u(Rxu \rightarrow Zu) \ \& \ \forall u \forall v(Zu \ \& \ Ruv \rightarrow Zv)) \rightarrow Zy))$$

We may give S a name in each case, writing ρR or σR or τR for the S in the three cases. It is easily seen that ρR is reflexive and that σR is symmetric. It is less easily seen but also true—this result having been about the first non-trivial theorem of distinctively higher-order logic, proved in the *Begriffsschrift*—that τR is transitive. ρR and σR and τR are called the reflexive and symmetric and transitive *closures* of R, respectively. If Rxy is the intransitive "x is a parent of y," then τR amounts to the transitive "x is an ancestor of y." On account of this example the transitive closure is also called the *ancestral*. The notion of ancestral was especially important to Frege in his attempted development of mathematics within his system.[9]

In his *Grundlagen*, Frege attempted to situate his philosophical position among others by using Kant's threefold division of knowledge into the *analytic*, the *synthetic a priori*, and the *a posteriori*. In these terms, Frege agreed with the conclusion of Leibniz that arithmetic is analytic, while exposing the fallacy in the argument Leibniz offered in an attempt to establish that conclusion. Frege respectfully disagreed with Kant's claim that arithmetic is synthetic *a priori*, while agreeing with the corresponding claim for geometry—a surprisingly old-fashioned position in the era of non-Euclidean geometry. Frege ridiculed Mill's claim that arithmetic is *a posteriori*.

What was at issue in Frege's disagreements with his predecessors and contemporaries was not the classification of *actual* knowledge. Frege was not interested in how young children learn the basic laws of arithmetic, or how our remote ancestors learned them. He did not claim anyone before himself ever *had* proved the basic laws of arithmetic from principles of pure logic plus appropriate definitions of arithmetical terms in logical terms, without appeal to any distinctively arithmetical "intuition." Frege's claim was, rather, that this *can* be done, even if no one before him ever did it, and that because it can be done arithmetic ranks as analytic.

A potential source of doubt is the principle of mathematical induction, which might be, and by some was, cited as an arithmetical law that could only be established by "intuition."[10] According to this principle, if a condition is fulfilled by zero, and is fulfilled by the successor of any natural number fulfilling it, then it is fulfilled by all natural numbers. Frege already at the time of the *Begriffsschrift* had an idea of how the claim that "intuition" is indispensable for mathematical induction could be refuted.

The Fregean strategy would be first to define the notions of zero and successor, and then to define the greater-than relation as the ancestral of the successor relation.[11] We can then *define* the natural numbers as zero together with those objects that are greater than zero. If we call a concept *inductive* if zero falls under it and the successor of every number falling under it falls under it, then the Fregean definition is more or less equivalent to defining natural numbers as those objects that fall under all inductive concepts. Except for the substitution of talk of a number falling under a concept for talk of a condition being fulfilled by a number—and the two ways of speaking are equivalent by the comprehension axiom—this definition makes the principle of mathematical induction true *almost by definition*.

1.2 Frege's Mathematics

When it came to working out the derivation of arithmetic from logic in detail in his *Grundgesetze*, Frege found he needed one further axiom beyond anything in his *Begriffsschrift*. What need to be explained next are: first, what this additional assumption—which Frege numbered as axiom or basic law V—amounted to; second, how Frege proposed to obtain the most basic laws of arithmetic—the so-called *Peano postulates*—

(continued)