Thoughts and Sharp Boundaries

3.1 The Issue

Frege's view as presented so far includes the theses (1) that ordinary arithmetical discourse involves the expression of true and false thoughts and (2) that some among these thoughts are to be demonstrated by the logicist reduction to be grounded in pure logic. Also included is the claim (3) that the final sentences of Frege's Grundgesetze derivations are themselves intended to express thoughts, ones whose logical grounding is immediately established by those derivations. The question raised in Chapter 2, and to which we'll return in Chapter 4, is that of how Frege conceives the relationship between the thoughts just mentioned in (3) and those mentioned in (1).

Before returning to that question, we digress here to treat a potential difficulty that arises for thesis (1) and thus for the outline just sketched of Frege's logicist strategy.1 The difficulty is this: Frege claims in many places that functions, and hence concepts, must have "sharp boundaries," by which he seems to mean that they are total, in the sense of delivering a value for every argument (or n-tuple of arguments) of appropriate logical type. There seems good evidence, additionally, that he holds the following views: that a function-term that fails to refer to a function thereby fails to refer at all, and that, as a consequence, no sentence in which that function-term appears can have any reference, which is to say that it can have no truth-value.2 If the sentences of ordinary arithmetic are to express truths and falsehoods, then, it is essential, on this view, that their function-terms refer to total functions. The difficulty is that they don't seem to do so. The sense and reference of such function-terms as "+," in ordinary use, are presumed determined largely by our use of them together with relevant mathematical facts. And no such considerations determine what value the reference of such a term should have when applied to the pair of arguments <7, the Eiffel Tower>

It is worth emphasizing here that the difficulty is not just that outlying sentences like "7 + the Eiffel Tower < 14" have no truth-value; the problem is that, given the constellation of theses just outlined, it follows that "7 + 3 = 10," similarly,
has no truth-value, since its function-term fails to refer to a total function and hence fails to refer. All such ordinary sentences fail, then, on this view, to express truth-evaluable thoughts.

We'll begin by looking at the "sharp boundaries" passages in question and at the implications for Frege's project of reading them in the way just suggested, i.e. as demanding that all functions be total. In the end, I'll argue that, despite the textual evidence for this reading, Frege cannot be read as holding that all functions are total. What he does mean in the "sharp-boundaries" passages is significant for our understanding of how he understands the scope of theories and of logic. Most importantly for present purposes, it's not the case that these passages undermine the view of ordinary arithmetical discourse as consisting of the expression of true and false thoughts.

3.2 The Texts

The sharp-boundaries texts include these:

An 1896 claim that occurs in the course of a criticism of Peano:

[\[E\]very concept must have sharp boundaries, so that it is determined for every object whether it falls under the concept or not.\]}

A discussion in *Grundgesetze* that applies the requirement to functions more generally:

[\[A\] first-level function of one argument must always be such as to yield an object as its value, whatever object we may take as its argument. . . .\]

[\[E\]very first-level function of two arguments must have an object as its value for any one object as its first argument and any other object as its second. . . .\]

Similarly for functions of higher level.\]

An 1891 insistence on the

... requirement, as regards concepts, that, for any argument, they shall have a truth-value as their value; that it shall be determinate, for any object, whether it falls under that concept or not.\]

And there are more.\]

If Frege means that, strictly speaking, all functions and concepts must be defined everywhere, even on arguments foreign to the subject matter under discussion, then this requirement undermines a good deal of the picture one might have had of Frege's understanding of language, given his frequent examples of (apparently) truth-evaluable thoughts expressed by ordinary sentences.\]

It's not just e.g. "+" as ordinarily understood that fails to refer to a total function. Predicates and relation-terms whose natural homes are in the physical sciences, for example, are not generally defined over arbitrary objects outside of this range. And even without looking at "outlying" arguments, troubles are easily found close to home: the phrase "the eldest child of . . ." , as applied to a childless adult, gives no determinate answer and so must too, on the requirement of totality, count as reference-less. No sentence employing this function-term can express a true or false thought on the view in question.

In short, if Frege holds that a condition of reference for function-terms is that they refer to total functions, then he cannot coherently hold that ordinary discourse or ordinary scientific inquiry involves, in his words, a "common stock of thoughts," at least not of true or false ones.\]

On one important interpretation of Frege's work, this is just the conclusion that Frege endorses. As Joan Weiner understands the project, Frege was not, after all, concerned with the functioning of language in ordinary, including ordinary scientific, discourse. His theory of sense and reference, on this view, is meant to apply only to languages expressly designed for rigorous scientific inquiry and proof, and meeting some very high standards of completeness. Thought-expression is a feature, in short, just of languages as careful and systematic as that of *Grundgesetze*. As Weiner puts it:

\[G\]iven [Frege's] requirement that each predicate pick out a concept with a sharp boundary, few, if any, of our everyday sentences have comprehensible sense or truth-values.\]

This means that whatever mathematicians were doing prior to the systemization of arithmetic given by Frege's work, they were not expressing, debating, or proving arithmetical truths. The sentences they used expressed neither truths nor falsehoods. And the same must be said for us: since the final systematization of arithmetic has not yet been accomplished, and since our arithmetical education does not inculcate a grasp of senses that determine total functions, we cannot be said to be engaged in expressing, debating, or proving arithmetical truths. Indeed, as Weiner acknowledges, it is a consequence of her interpretation of Frege that, on the Fregean view, "it is not even obvious that we know any truths of arithmetic."\]

For what we know, on the Fregean picture, is what's expressed by some of the sentences we affirm, and none of these, on that picture, express truths.

One difficulty for this view is that most of Frege's discussions of thoughts and their expression in language have to do not with the formulas of a precisely
specified formal system, but with ordinary sentences of natural language. In his most pointed discussions of sense and reference, e.g. in the trio of papers from 1891–92 and in the later "Der Gedanke" and related essays, he never claims that the sentences of ordinary discourse fail to express thoughts. And in his discussions of the ways in which rigorous formal languages offer improvements on ordinary language, he never claims that a fault of ordinary sentences is that they fail to express thoughts. Ordinary language is sloppy, to be sure, on Frege's view, in ways already touched upon above, but that the sloppiness is sufficient to undermine the attempt to use the sentences of such a language to express truths and falsehoods is never cited by Frege as one of its difficulties.

A more serious problem for the view that only formal languages like that of Grundgesetze are capable of expressing true and false thoughts, in virtue of the total definition of the functions referred to in such a language, is that the functions referred to in Grundgesetze are very clearly not total. Consider Frege's treatment of course-of-values names. At Grundgesetze 1 §10, in the course of laying down the references of his terms, Frege discusses the fact that course-of-values names have not yet been assigned determinate reference via the stipulation.

The discussion is as follows:

Although we have laid it down that the combination of signs

\[ \varepsilon \Phi(x) = \ulcorner \Psi(x) \urcorner \]

has the same denotation as

\[ \forall x(\Phi(x) = \Psi(x)) \]

this by no means fixes completely the denotation of a name like "\( \varepsilon \Phi(x) \)". We have only a means of always recognizing a course-of-values if it is designated by a name like "\( \varepsilon \Phi(x) \)" by which it is already recognizable as a course-of-values. But we can neither decide, so far, whether an object is a course-of-values that is not given us as such, and to what function it may correspond, nor decide in general whether a given course-of-values has a given property unless we know that this property is connected with a property of the corresponding function. ... How may this indefiniteness be overcome? By its being determined for every function when it is introduced, what values it takes on for courses-of-values as arguments, just as for all other arguments. Let us do this for the following functions considered up to this point."

Here Frege lists the three functions considered up to now, and notes that two of them are reducible to the third, i.e. to the identity-function:

Since in this way everything reduces to consideration of the function \( \varepsilon = \xi \), we ask what value this has if a course-of-values occurs as argument. Since up to now we have introduced only the truth-values and courses-of-values as objects, it can only be a question of whether one of the truth-values can perhaps be a course-of-values."

Consider a particular course-of-values, say, \( \varepsilon(e = c) \). The indeterminacy in question is, with respect to this case, the fact that sentences of the form "\( \ulcorner = \varepsilon(e = c) \)" and "\( \varepsilon(e = c) = \ulcorner \)" have no truth-value in cases in which the blank is filled in by anything other than a term of the form "\( \varepsilon \Phi(x) \)". One might think that the problematic indeterminacy of reference on the part of "\( \varepsilon(e = c) \)" is due to the fact that a vast number of identity-questions regarding its purported reference have been left unanswered. But that this would be a misleading way of construing the difficulty is made clear by Frege's means of addressing it. Frege does not solve the problem by providing a means of answering those unanswered questions in general. He does not bring it about that the concept-phrase "\( \ulcorner = \varepsilon(e = c) \)" determinately holds or fails to hold of each object. He instead brings it about that every way of completing this concept-phrase with a singular term of the language, i.e. every sentence formed in this way, has a determinate truth-value. As he says, since the only objects that have been "introduced" up to this point (i.e., the only ones to which one can refer in the language) are truth-values and courses-of-value, and since the identity-questions about courses-of-value have been settled already, the only question left to be settled is that of whether the concept-phrase "\( \ulcorner = \varepsilon(e = c) \)" is satisfied by terms for either of the truth-values. Similarly for each of the course-of-value terms: for each such term \( \varepsilon(\Phi(e)) \), it must be determined whether "\( \ulcorner = \varepsilon(\Phi(e)) \)" is satisfied by terms for either of the truth-values.

Frege settles the issue by a simple arbitrary stipulation: "[I]et us lay it down that \( \varepsilon(e) \) is to be the True and \( \varepsilon(e = e) \) is to be the False." This neatly solves the problem: each of the problematic sentences now has a determinate truth-value. The stipulation does not, of course, have the result that the concept-phrases of the form "\( \ulcorner = \varepsilon(\Phi(e)) \)" refer to total concepts: nothing has been said to determine whether any such phrase is satisfied by objects not named in the language. Similarly, nothing has been said to determine whether the courses-of-values in question fall under or fail to fall under concepts not yet referred to in the language. Frege's remarks on this circumstance are simply that:

With this we have determined the courses-of-values so far as is here possible. As soon as there is a further question of introducing a function that is not completely reducible to functions known already, we can stipulate what value it is to have for courses-of-values as arguments; and this can then be regarded as much as a further determination of the courses-of-values as of that function."
The fact that the concepts referred to by Grundgesetze concept-expressions are very clearly not total does not, as Frege sees it, stand in the way of the entirely unproblematic functioning of the language of Grundgesetze. Indeed, the failure of totality does not even stand in the way of the expression of determinate, truth-evaluable thoughts by the formulas of that language. As Frege puts it in Grundgesetze I §32, having argued that each of the ways of combining function- and argument-expressions of the language has a well-defined reference (in the sense that no formula constructed out of them lacks a truth-value):

...we have in every correctly-formed proposition of the concept-script [here Frege refers to a sentence prefixed with the judgment-stroke] a judgment that a thought is true; and here a thought certainly cannot be lacking.19

In short, the sentences of Grundgesetze express unproblematic thoughts, ones that have determinate truth-values. This despite the fact that the functions referred to via their predicative parts are all very partial, defined only over an extremely limited collection of objects.

What, then, does Frege mean by the sharp-boundaries requirement, given that neither ordinary (arithmetical) discourse nor the extremely careful language of Grundgesetze come close to meeting the requirement of totality?

Here it will be important to distinguish the requirement of totality from what we'll call the requirement of linguistic completeness. The latter is the requirement for a given language that each of its well-formed concatenations of symbols has a determinate reference. This is a requirement that Frege takes to be crucial for languages that are intended for the expression of rigorous arguments, and especially for those like his own concept-script, with respect to which the conditions of rigorous proof are to be laid down syntactically. For the presence in a language of well-formed phrases lacking reference is an invitation to fallacious reasoning. As Frege puts it:

It seems to be demanded by scientific rigor that we should have provisions against an expression's possibly coming to have no reference; we

must see to it that we never perform calculations with empty signs in the belief that we are dealing with objects. People have in the past carried out invalid procedures with divergent infinite series.20

Not just mathematical calculations, but ordinary reasoning as well can be derailed by the presence of non-referring parts of language. Even the law of excluded middle, as Frege points out, will lead us astray if we rely on it in contexts in which the application of a concept-expression to an object-name gives no definite result. Particularly important to Frege, given his interest in induction, are the Sorites-style fallacies that are invited by the use of concept-expressions that fail to deliver a truth-value on some completions by object-names.21 In short, a language that fails the requirement of linguistic completeness is one that's not suited for maximum argumentative clarity and rigor. Frege is explicit in his application of the requirement of linguistic completeness to the language of Grundgesetze, going so far as to present a detailed proof that the requirement is satisfied.22

As to the relationship between the requirement of linguistic completeness and the requirement of totality: if a given theory or discourse is concerned just with a specific range of objects—say, the integers—then the language used there will satisfy the requirement of linguistic completeness as long as its function-expressions refer to functions defined on all of the arguments from this range. The requirement of linguistic completeness for such a language is considerably weaker than the requirement that its function-terms refer to total functions, i.e. to ones whose domain includes outlying objects.

Frege clearly has just such restricted languages in mind much of the time. The central example of this is just the language of Grundgesetze itself. With respect to the context of ordinary arithmetic, additionally, Frege notes in "Function and Concept" that:

So long as the only objects dealt with in arithmetic are the integers, the letters a and b in "a + b" indicate only integers; the plus-sign need be defined only between integers.23

These two examples are of course counterexamples to the requirement of totality, and further evidence that Frege's sharp-boundaries requirement is not in fact the strong totality requirement.

What, then, does Frege mean when he claims that functions must be defined for "all arguments"? Here, the evidence points to the view that by "all arguments," Frege means to include all of those arguments in the domain of the science or theory at hand. Frege's real interest, I'll suggest, is the requirement of linguistic completeness, a requirement in which he has a clear vested interest. Consider the first sharp-boundaries passage quoted above, together with a little more of its context:
We can also argue for this requirement on the ground that the law of the excluded middle must hold. For that implies that every concept must have sharp boundaries, so that it is determined for every object whether it falls under the concept or not.24

The "requirement" to which Frege refers is spelled out in the previous paragraph, which consists largely in a criticism of Peano for having failed to define his function-terms "num" and "e" in such a way that every result of completing their argument-places with names from Peano's language will have a determinate reference:

[If, in our example, we take as before "then u is not a class" as a consequent, we must also understand by u something that is not a class, and hence in this case too "num u" must refer to something, if it is to be possible to judge whether the condition "if num u does not coincide with num v" is satisfied, and equally it must also be possible to judge whether a given relation maps u into v or v into u for the case where u is not a class. So the definitions of "num u" and of mapping ("f e vfu") must be formulated accordingly. Since one cannot know from the outset in which sentences these signs will occur and what restrictions will be thereby placed on the meanings of the letters, the definitions are to be constructed in such a way that a reference is guaranteed these combinations of signs for every reference of the letters.25]

That is: the requirement in question is the all-important requirement of linguistic completeness, not the requirement that the function-signs be defined over outlying objects.

Similarly for the Grundgesetze's sharp-boundaries passages quoted above. Despite his talk of functions defined over "any" arguments, Frege's practice makes it clear that the only objects meant here are those that have names in the language of Grundgesetze, i.e. courses-of-value and truth-values. Once again, the only coherent way to read these passages is as insisting not on the strong requirement of totality for functions, but on the requirement of linguistic completeness.

3.3 Piecemeal Definition and New Objects

Perhaps the most important context in which Frege discusses what looks like the requirement of totality, certainly the context in which the sharp-boundaries requirement is raised in the most sustained and most forceful way, concerns Frege's prohibition against what he calls "piecemeal definition." What he means by this, and the reasons for prohibiting it in a good scientific system, are laid out in Grundgesetze as follows:

A definition of a concept (of a possible predicate) must be complete; it must unambiguously determine, as regards any object, whether or not it falls under the concept (whether or not the predicate is truly ascribable to it). . . . We may express this metaphorically as follows: the concept must have a sharp boundary. . . . [A] concept that is not sharply defined is wrongly termed a concept. Such quasi-conceptual constructions cannot be recognized as concepts by logic; it is impossible to lay down precise laws for them. The laws of excluded middle is really just another form of the requirement that the concept should have a sharp boundary.

Now from this it follows that the mathematicians' favorite procedure, piecemeal definition, is inadmissible. The procedure is this: First they give the definition for a particular case—e.g. for positive integers—and make use of it; then, many theorems later, there follows a second definition for another case—e.g. for negative integers and zero. Here they often commit the further mistake of making specifications all over again for the case they have already dealt with. Even if in fact they avoid contradictions, in principle their method does not rule them out. What is more, as a rule they do not attain to completeness, but leave over some cases, as to which they make no specification; and many are naive enough to employ the word or symbol for these cases too, as if they had assigned a meaning to it. . . . But the chief mistake is that they are already using the symbol or word in theorems before it has been completely defined—often, indeed, with a view to further development of the definition itself. So long as it is not completely defined, or known in some other way, what a word or symbol means, it may not be used in an exact science—least of all with a view to further development of its own definition.26

The dangers associated with piecemeal definition, then, include (1) the possibility of introducing a contradiction by providing later stipulations incompatible with earlier ones, and (2) the possibility of failing to cover some cases, with the result that some combinations of symbols are left without reference.

But if piecemeal definition is ruled out, then what are we to say of those episodes in the history of mathematics in which familiar function-signs have had to have their reference extended so as to be applicable to newly-introduced objects? Frege's answer is as follows:

Now, of course it must be admitted that scientific progress, which has been effected by conquering wider and wider domains of numbers,
rehearsal of the lesson to be learned from the "Caesar" passage. The relevant part for our purposes is as follows:

Let us cast a final brief glance back over the course of our enquiry. After establishing that number is neither a collection of things nor a property of such, yet at the same time is not a subjective product of mental processes either, we concluded that a statement of number asserts something objective of a concept. We attempted next to define the individual numbers 0, 1, etc., and the step from one number to the next in the number series. Our first attempt broke down, because we had defined only the predicate which we said was asserted of the concept, but had not given separate definitions of 0 or 1, which are only elements in such predicates. This resulted in our being unable to prove the identity of numbers. It became clear that the number studied by arithmetic must be conceived not as a dependent attribute, but substantively.

In short, the difficulty with the original proposal has nothing to do with the totality of the concept of number. The difficulty illustrated by the Caesar example is that this proposal fails to treat numbers as objects. In so doing, it also fails to treat the relevant predicates, e.g. "... = 0" and "... is a number," as function-denoting phrases of any kind.

3.6 Quantification

One final issue, and one about which Frege is simply not very clear, concerns the status of quantified sentences involving function-expressions that do not refer to total functions. Consider the case of first-level concepts. If a concept-expression \(g(x)\) refers to a concept that is undefined on some arguments (specifically, on arguments with which the science in question is not concerned), then what are the truth-conditions of the universally-quantified sentence \(\forall x g(x)\)?

Frege notes the connection between the issue of totality and quantification as follows, in a passage that, taken in isolation, would seem to conflict with much of what he has said in the passages quoted above:

As soon as people aim at generality in propositions they will need in arithmetical formulae not only symbols for definite objects—e.g. the proper name "2"—but also letters that only indicate and do not designate [i.e., bound variables—PB]; and this already leads them, quite unawares, beyond the domain within which they have defined their symbols. One may try to avoid the dangers thus arising by not making the letters indicate objects in general (as I did), but only those of a domain with fixed boundaries. Let us suppose for once that the concept number has been sharply defined; let it be laid down that italic letters are to indicate only numbers; and let the sign of addition be defined only for numbers. Then in the proposition \(a + b = b + a\) we must mentally add the conditions that \(a\) and \(b\) are numbers; and these conditions, not being expressed, are easily forgotten. But let us deliberately not forget them for once! By a well-known law of logic, the proposition

"if \(a\) is a number and \(b\) is a number then \(a + b = b + a\)

can be transformed into the proposition

"if \(a + b\) is not equal to \(b + a\), and \(a\) is a number, then \(b\) is not a number"

and here it is impossible to maintain the restriction to the domain of numbers.\(^{36}\)

Frege might be taken here to reject the view that the range of values of a given variable can be restricted to a specific domain, e.g. to numbers. But it is important to see that this is not his point. Frege's argument here is intended to expose the difficulty involved in restricting the definition of function-signs, e.g. of the plus-sign, to a proper part of the collection of objects already under discussion. As he puts it in the previous paragraph:

Such a restriction would have to be incorporated into the definition, which would thus take some such form as: "If \(a\) and \(b\) are numbers, then \(a + b\) means ..." We should have a conditional definition. But the sign of addition has not been defined unless every possible complex symbol of the form \("a + b\)" has a definite meaning, whatever meaningful proper names may take the places of "\(a\)" and "\(b\)."\(^{37}\)

The difficulty is the one already noted above with conditional definitions. We can't coherently restrict the functions referred to by our function-signs to those arguments that satisfy some predicate of the language, since this strategy leads immediately to well-formed combinations of symbols that lack reference. It leaves out not only those singular terms formed by concatenating the function-sign with argument-signs from outside the chosen domain, but also those quantified sentences that involve the function-sign and the negation of the predicate. The point, again, is not that our functions must be total, but rather that they must be defined on all of the arguments over which we quantify in the science in question.

But what, exactly, is the domain of objects over which we quantify in a given science? Frege does not, in explaining the truth-conditions of quantified sentences, refer to an explicitly demarcated range. He explains those truth-conditions variously in terms of the truth of each instance of the matrix and
in more familiar "objectual" ways. In each case, the general idea is that a universally-quantified sentence talks about all of the arguments of a certain kind. The unrestricted universality in terms of which Frege explains quantification has often been interpreted as implying, or being, the view that e.g. first-level quantification always talks about all objects, second-level about all first-level functions, etc. But this can't be right. Recall Frege's claim that as long as we're talking just about e.g. the integers, it's entirely coherent to define the addition sign just over the integers. The sentence "∀x∀y(x + y = y + x)" is, when the sign is used in this way, true. Hence the range of the first-level quantifier is straightforwardly the range of objects under discussion: in this case, the integers. That Frege is less explicit about the precise limits of the scope of a given domain than his successors have come to be involves a notable lack of comprehensive precision on his part. But given the context of his work, in which, as now, it is entirely commonplace to take one's audience to understand the range of mathematical (or other) objects under discussion, this comparative lack of precision should not strike us as particularly surprising.

It is also worth noting in this context that our own habit now of always explicitly specifying a domain of quantification, rather than relying on a general understanding of the boundaries of the theory in question, is a result of the modern conception of theories as multiply-interpretable sets of sentences whose domains require explicit demarcation on each application. From the Fregean point of view, in which sentences are fully interpreted and theories are sets of thoughts, there is always a determinate subject matter under discussion that already sets the boundaries of quantification. When we're dealing with the arithmetic of the integers, for Frege, "∀x" means "every integer"; when with the reals, "∀x" means "every real," and so on.

That Frege admits the coherence and the scientific acceptability of theories with restricted ranges (e.g., mathematical theories with particular subject-matters) does not by itself tell us much about how he views the ranges of his own quantifiers in Begriffsschrift and Grundgesetze. Given Frege's view of logic as a purely general science, one might have expected him to view the purely-logical setting of Begriffsschrift and Grundgesetze as imposing a requirement of complete generality on the ranges of the quantifiers of the languages \(L_o\) and \(L_e\). What we have seen is that, at least with respect to \(L_o\), this doesn't seem to be the case. Frege's strategy in Grundgesetze is to define \(L_o\) 's first-level function-terms over just courses-of-value and truth-values. And presumably, then, Frege takes it that the language as a whole, including its quantifiers, is restricted in semantic scope to those objects. (For otherwise, his quantified sentences will all lack truth-values.)

Frege holds that the fundamental principles of logic are purely general: that they hold for all objects and for all functions of appropriate type. The limited range of Grundgesetze's quantifiers brings to light the connection between the fundamental laws of logic and the thoughts expressed by Grundgesetze's axioms.

The latter are restrictions of the former, restrictions to the domain of the system. This it not to say that there is any barrier from the Fregean point of view to the expression of the purely general rules of logic, or to the use of absolutely-unrestricted quantification. But the requirement of rigor for Grundgesetze, i.e. the requirement of linguistic completeness, means that the quantifiers of \(L_o\) can only range over that domain over which its function-terms are defined.

3.7 Conclusion

Despite some often-quoted passages that seem to indicate that Frege takes all functions to be total, there is in the end no coherent way to read him as endorsing this requirement. In the notorious passages in question, Frege insists that argumentative rigor demands (1) the rejection of conditional and piecemeal definitions, and (2) the satisfaction of the requirement of linguistic completeness. Hence in any context in which argumentative rigor is at a premium, it is important that every function be defined over every argument of the domain under discussion. That even the strictest demands of rigor fail to imply a requirement of totality on the part of functions is made most clear by the fact that the function-expressions of Frege's own extraordinarily careful and rigorous language of Grundgesetze fail to refer to total functions.

Furthermore, the fact that the function-expressions of ordinary language and of ordinary scientific discourse are not defined over outlying objects and concepts—e.g. the fact that "... plus..." as ordinarily used is not defined over mountains or proper sets—tells us nothing, from the Fregean point of view, about the capacity of those expressions to appear in sentences expressing truth-evaluable thoughts. And finally, the fact that the function-expressions of ordinary discourse fail to have sharp boundaries even within the intended domain of discourse—e.g. that "heap" is vague, that "the eldest son of ..." often delivers no value when attached to the name of a person—tells us just that these expressions are not ideally suited to the rigors of argument, not that they cannot be used, according to Frege's semantic theory, in the expression of truths and falsehoods.

Notes

1. For a fuller treatment of this issue, see Blanchette [forthcoming].
2. Grundgesetze (hereafter Gg) 2 §56: "a concept that is not sharply defined is wrongly termed a concept." See also the §62 discussion of "inadmissible sham concepts."
4. Gg 2 §63.
To return to the central issue: Frege claims to be engaged in a demonstration that the truths of arithmetic are purely logical, a demonstration that proceeds via proofs of arithmetical truths from purely-logical ones. But the proofs he gives do not in any obvious sense terminate with truths of arithmetic. They terminate instead with thoughts about such unfamiliar things as courses-of-values and highly convoluted functions defined on them. As we have seen, Frege represents the proven truths as the results of analyses of arithmetical truths: the central idea is that those arithmetical truths are clarified or elucidated via an analysis of them and of their components, and only subsequently are the thus-analyzed truths subjected to proof. As Frege puts it in 1914:

"The effect of the logical analysis ... will then be precisely this—to articulate the sense clearly. Work of this kind is very useful; it does not, however, form part of the construction of the system, but must take place beforehand. Before the work of construction is begun, the building stones have to be carefully prepared. . . ."

The preparation of the "building stones" is the analysis of arithmetical truths, an analysis that yields those fully-articulated truths that will then be subject to proof. But the rather large gulf between the analysandum-thoughts and their analysans-counterparts leaves it far from clear that what Frege gives can in any reasonable sense be called merely an "analysis" of the ordinary contents of arithmetical discourse. More to the point: because of Frege's lack of clarity about the nature of what he calls "analysis," it is not obvious that the proof from purely logical principles of his highly-complex and unfamiliar analysans-thoughts can really be taken as a guarantee of the purely logical nature of the truths of arithmetic.

The purpose of this chapter is to clarify, insofar as this is possible, Frege's understanding of the role of conceptual analysis in his proof-theoretic project.
used wherever a special value must be placed on the validity of proofs. . . .” For general applications, it is essential that a derivation from premise-sentences make unquestionably clear that the thought expressed by its conclusion-sentence follows logically from the thoughts expressed by those premise-sentences.

We'll repeat here the definition of reliability introduced in Chapter 6, in this case as applied to a language that, like Frege's, has a single canonical reading: A deductive system D for such a language is reliable iff, for each of its sentences \( \phi \) and each of its sets \( \Sigma \) of sentences:

If \( \phi \) is derivable in \( D \) from \( \Sigma \), then the thought expressed by \( \phi \) follows logically from the thoughts expressed by the members of \( \Sigma \).

It is crucial for Frege's purposes that his system be reliable, and also essential that its reliability be relatively transparent. The fundamental difficulty revealed by Russell's paradox was, again, that the enriched system of Grundgesetze was not in fact reliable.

The systems of Begriffsschrift and Grundgesetze, which we'll call "B" and "G" respectively, each include a clearly-specified language, a collection of axiom-sentences, and a handful of rules for deriving sentences one from another. In order that the systems be obviously reliable, it is essential that each axiom-sentence express a thought that isn't just a logical truth, but one that is so simple and self-evident that its purely-logical status is beyond question. Similarly, it is essential that the derivation-rules sanction only the most self-evidently logical inferences, that they be applied correctly, and that the introduction of defined terms causes no problems.

Because there are no gaps in the chains of inference, every "axiom," every "assumption," "hypothesis," or whatever you wish to call it, upon which a proof is based is brought to light; and in this way we gain a basis upon which to judge the epistemological nature of the law that is proved. Of course the pronouncement is often made that arithmetic is merely a more highly developed logic; yet that remains disputable so long as transitions occur in the proofs that are not made according to acknowledged laws of logic, but seem rather to be based upon something known by intuition. Only if these transitions are split up into logically simple steps can we be persuaded that the root of the matter is logic alone. I have drawn together everything that can facilitate a judgment as to whether the chains of inference are cohesive and the buttresses solid. If anyone should find anything defective, he must be able to state precisely where, according to him, the error lies: in the Basic Laws, in the Definitions, in the Rules, or in the application of the Rules at a definite point. If we find everything in order, then we have accurate knowledge of the grounds upon which each individual theorem is based.  

The possibility in principle of finding something "defective" in the deductive system is the possibility of finding that an axiom ("basic law") or rule doesn't state or instantiate an "acknowledged law of logic" or that the introduction of a defined term presupposes either some falsehood or some truth of a non-logical kind.

Frege is of course convinced that neither B nor G has such defects. At the time of Begriffsschrift, with its relatively-simple language and semantics, Frege has little to say about definitions other than the standard constraint that they be entirely stipulative and eliminable in principle. He is more explicit about the correctness of his axioms and derivation-rule. In the early sections of Begriffsschrift, for example, Frege introduces the conditional sign and its associated derivation-rule, modus ponens, as follows:

\[ \text{§5 If } A \text{ and } B \text{ stand for contents that can become judgments . . . , there are the following four possibilities:} \]

1. \( A \) is affirmed and \( B \) is affirmed;
2. \( A \) is affirmed and \( B \) is denied;
3. \( A \) is denied and \( B \) is affirmed;
4. \( A \) is denied and \( B \) is denied.

Now \[ \vdash (B \rightarrow A) \]
stands for the judgment that the third of these possibilities does not take place, but one of the three others does.

\[ \text{§6. The definition given in §5 makes it apparent that from the two judgments:} \]

\[ \vdash (B \rightarrow A) \text{ and } \vdash B \]

the new judgment

\[ \vdash A \]

follows. Of the four cases enumerated above, the third is excluded by

\[ \vdash (B \rightarrow A) \]

and the second and fourth by

\[ \vdash B \].
so that only the first remains.\(^3\)

Similarly for the axioms. Formula (1), Frege's first axiom, is introduced as follows:

\[ \vdash (a \rightarrow (b \rightarrow a)) \]

says "The case in which \(a\) is denied, \(b\) is affirmed, and \(a\) is affirmed is excluded." This is evident, since \(a\) cannot at the same time be denied and affirmed.\(^5\)

The justification of formula (2) is lengthier:

\[ \vdash (c \rightarrow (b \rightarrow a)) \rightarrow ((c \rightarrow b) \rightarrow (c \rightarrow a)) \]

means "The case in which

\[(c \rightarrow b) \rightarrow (c \rightarrow a)\]

is denied and

\[(c \rightarrow (b \rightarrow a))\]

is affirmed does not take place." But

\[(c \rightarrow (b \rightarrow a))\]

means the circumstance that the case in which \(a\) is denied, \(b\) is affirmed, and \(c\) is affirmed is excluded. The denial of

\[(c \rightarrow b) \rightarrow (c \rightarrow a)\]

says that \((c \rightarrow a)\) is denied and \((c \rightarrow b)\) is affirmed. But the denial of \((c \rightarrow a)\) means that \(a\) is denied and \(c\) is affirmed. Thus the denial of

\[(c \rightarrow b) \rightarrow (c \rightarrow a)\]

means that \(a\) is denied, \(c\) is affirmed, and \((c \rightarrow b)\) is affirmed. But the affirmation of \((c \rightarrow b)\) and that of \(c\) entails the affirmation of \(b\). That is why the denial of

\[(c \rightarrow b) \rightarrow (c \rightarrow a)\]

has as a consequence the denial of \(a\) and the affirmation of \(b\) and \(c\). Precisely this case is excluded by the affirmation of

\[(c \rightarrow (b \rightarrow a)).\]

Thus the case in which

\[(c \rightarrow b) \rightarrow (c \rightarrow a)\]

is denied and

\[(c \rightarrow (b \rightarrow a))\]

is affirmed cannot take place, and that is what the judgment

\[ \vdash (c \rightarrow (b \rightarrow a)) \rightarrow ((c \rightarrow b) \rightarrow (c \rightarrow a)) \]

asserts.\(^6\)

In each of these cases, Frege's task is the simple one of pointing out that, given what the signs mean, "it is apparent" that the inference-rule is reliable, or "it is evident" that instances of the axiom-sentence scheme express only truths, given just the fundamental logical principle against self-contradiction. Each of the remaining axiom-sentence schemes is accompanied by discussion of this kind, sometimes simply pointing out in German what a formula of the kind in question says, letting its self-evidence speak for itself, and sometimes adding to this a bit of discussion designed to bring the reader to see that what's expressed is indeed an "acknowledged law" of logic, i.e. that it can be seen to be true just on the basis of rudimentary logical reasoning.

It will be important in what follows to note that in these passages of Begriffsschrift Frege is not attempting the impossible task of "justifying" the fundamental logical principles in virtue of which contents or thoughts follow from or contradict one another. There is of course no way to do this, since any such attempted justification would presuppose the very principles it attempts to justify. As Frege says with respect to the law of non-contradiction, "If other persons presume to acknowledge and doubt a law in the same breath, it seems to me an attempt to jump out of one's own skin against which I can do no more than urgently warn them." Nothing like (noncircular) argument, or justification, can be given in cases like these, in which what's at issue is a fundamental principle of logic. Frege's purpose in these early pages of Begriffsschrift is simply to justify the inclusion of specific axiom-sentences and sentential derivation-rules in his formal system. In each of the brief
discussions to this effect, he relies on his audience to make a number of simple logical inferences and to recognize as correct a handful of basic logical principles. The argument for the reliability of B simply takes these fundamental principles and entailment-patterns for granted and uses them to establish the reliability of the sentential derivation-rules and axiom-sentences under discussion.8

By the time of Grundgesetze and the increased complexity of its language, its semantics, and its deductive system, Frege recognizes that establishing the adequacy of the formal system requires a more extended argument regarding the formulas themselves. In order that axioms always express logical truths, and that derivation-rules applied to sentences expressing truths always yield sentences expressing truths, it is crucial that each sentence of the formal language express a determinate thought with exactly one truth-value. This requires among other things that, as Frege puts it, "every correctly-formed name is to denote something, a principle that is essential for full rigor."9 In keeping with the importance of this requirement, Frege gives a series of careful treatments of terms and their denotations. Some examples of this approach as found in Grundgesetze 1 are as follows:

In §10, as discussed in Chapter 3 above, Frege removes an indeterminacy regarding certain sentences involving course-of-value names, an indeterminacy left by his earlier (§3) introduction of those names. The indeterminacy is removed by a two-step procedure. First, Frege proves that the system as defined to this point can be consistently extended by adding to it any of a number of arbitrary stipulations identifying a particular course of values with the True and another with the False.10 Second, he gives just such a stipulation. This is the stipulation that the True is to be identified with the course of values of the concept (or of any concept) under which it and it alone falls; similarly for the False.

In §11, Frege introduces the function \( \forall \), his version of the definite article. He notes that the definite article of ordinary language (schematically, "the F") carries with it the dual "logical danger" of forming ambiguous names (in the case in which there are two Fs) and of forming empty names (in the case in which there are no Fs).11 To ward off this danger, Frege assigns an arbitrary but determinate denotation to his newly-introduced terms in all such cases. He closes the section by pointing out that this strategy works, i.e. that "this danger about the definite article is here completely circumvented. . . ." Neither empty nor ambiguous names can be formed via the new operator, given the rules in terms of which Frege introduces it.

§29 contains an extended discussion of the conditions under which singular terms and function-names denote. §§30 and 31 provide a detailed argument to the effect that correctly-formed proper names and names of functions all denote.12

In §32, Frege notes that each name of the formalism has not just a denotation but also a sense, and that in the case of sentences, this sense is a thought.13 Having set out in §33 the principles of correct definition and applied these in §34 to the definition of the symbol "\( \forall \)" Frege notes in §35 that the definition is indeed a good one in the sense that, as is essential for the "correctness of ... proofs," the defined function-term has a denotation.

With respect to rules of derivation, Frege's discussions are similar to those of Begriffsschrift, The principle that he calls the "interchangeability of subcomponents" is introduced in §12 via the following line of reasoning:

\[ \Delta \rightarrow (\Lambda \rightarrow \Phi) \]

is the False if \( \Delta \) and \( \Lambda \) are the True while \( \Phi \) is not the True; in all other cases it is the True. From this there follows the interchangeability of \( \Lambda \) and \( \Delta \):

\[ \Lambda \rightarrow (\Delta \rightarrow \Phi) \]

is the same truth-value as

\[ \Delta \rightarrow (\Lambda \rightarrow \Phi) \]

Modus ponens is introduced as follows:

From the propositions \( \vdash (\Delta \rightarrow \Gamma) \) and \( \vdash \Delta \) we may infer \( \vdash \Gamma \); for if \( \Gamma \) were not the True, then since \( \Delta \) is the True, \( (\Delta \rightarrow \Gamma) \) would be the False.14

Similarly for the next rule:

The following method of inference is a little more complicated. From the two propositions \( \vdash (\Delta \rightarrow \Gamma) \) and \( \vdash (\Phi \rightarrow \Delta) \) we may infer the proposition \( \vdash (\Phi \rightarrow \Gamma) \). For \( (\Phi \rightarrow \Gamma) \) is the False only if \( \Phi \) is the True and \( \Gamma \) is not the True. But if \( \Phi \) is the True then \( \Delta \) too must be the True, for otherwise \( (\Phi \rightarrow \Delta) \) would be the False. But if \( \Delta \) is the True then if \( \Gamma \) were not the True then \( (\Delta \rightarrow \Gamma) \) would be the False. Hence the case of \( (\Phi \rightarrow \Gamma) \)'s being the False cannot arise; and \( (\Phi \rightarrow \Gamma) \) is the True.15

Axioms are justified similarly, in much the same way as in Begriffsschrift, though with more explicit appeal to truth-values. Basic Law I, which is Formula (1) of Begriffsschrift, is introduced as follows:

\[ (\Gamma \rightarrow (\Lambda \rightarrow \Gamma)) \]
could be the False only if both \( \Gamma \) and \( \Delta \) were the true while \( \Gamma \) was not the True. This is impossible; therefore

\[
(a \rightarrow (b \rightarrow a))
\]

Similarly for Basic Law IV:

\[-\Delta \text{ and } --\Delta \text{ are always different, and always truth-values. Now, since } \Gamma \text{ is in any case always a truth-value, it must coincide either with } -\Delta \text{ or with } --\Delta. \text{ It follows from this that } (-(-\Gamma = -\Delta) \rightarrow (-\Gamma = -\Delta)) \text{ is always the True; for it could be the False only if } (-(-\Gamma = -\Delta) \text{ were the True (i.e., if } -\Gamma = -\Delta \text{ were the False)) and } -\Gamma = -\Delta \text{ were not the True (i.e., if } -\Delta \text{ were the False). In other words, } \Gamma = -\Delta \text{ would be the False only if both } -\Gamma = -\Delta \text{ and } -\Gamma = -\Delta \text{ were the False, which, as we just saw, is not possible. Therefore,}
\]

\[
(-(-a = -b) \rightarrow (-a = -b)).
\]

And so on for the remaining axioms: Frege points out in each case that, given what the signs mean, it is impossible for an instance of an axiom to refer to the False.

These metatheoretic passages consist essentially in (1) arguments to the effect that the syntactic formation-rules and meaning-stipulations together guarantee that every well-formed piece of language has an appropriate reference and that well-formed sentences express determinate thoughts; and (2) arguments, or in simple cases mere pointings-out, that each axiom and rule of derivation is reliable. The arguments and discussions take up a very small portion of the two formal works and are by no means at the center of Frege's attention. These passages do, however, suffice for Frege's metatheoretic purposes: they show (or would have shown, if they hadn't contained errors) that Frege's formal systems were reliable, and hence that Frege's means of demonstrating his logicist thesis, and of setting up a framework for rigorous logical demonstrations in general, was a good one.

From the vantage point of one who takes the formal system as an important tool, and not as an independently-interesting topic for investigation in its own right, this is just what one should expect to see as metatheory.

7.2 Universalism and Metatheory

7.2.1 The Issue

The account just given, on which Frege provides careful arguments regarding the adequacy, in various senses, of his formal systems, is not uncontroversial. A number of commentators on Frege have held that Frege's conception of logic makes it impossible or incoherent for him to engage in this kind of metatheoretic reasoning.

In particular, the view in question is that Frege conceives of logic as "universal" in a sense that makes the whole idea of metatheory incoherent.

Dreben and van Heijenoort, for example, hold that:

[N]either in the tradition in logic that stemmed from Frege through Russell and Whitehead, that is, logicism, nor in the tradition that stemmed from Boole through Peirce and Schröder, that is, algebra of logic, could the question of the completeness of a formal system arise.

For Frege, and then for Russell and Whitehead, logic was universal: within each explicit formulation of logic all deductive reasoning, including all of classical analysis and much of Cantorian set theory, was to be formalized. Hence not only was pure quantification theory never at the center of their attention, but metasystematic questions as such, for example the question of completeness, could not be meaningfully raised. . . . We have no vantage point from which we can survey a given formalism as a whole, let alone look at logic whole.18

Similarly, Goldfarb holds that:

If the system constitutes the universal logical language, then there can be no external standpoint from which one may view and discuss the system. Metasystematic considerations are illegitimate rather than simply undesirable.19

The spatial metaphor, according to which there is no "vantage point," no "external standpoint," no "perspective" from which to carry out metatheoretic reasoning, is pervasive in this line of interpretation. Metatheory, on this account, requires one in some sense to "stand outside" the formal system about which one is reasoning, but, the view continues, Frege's understanding of logic or of his system as "universal" entails that it is impossible to do so.

There are a number of things one might mean by the claim that logic or a system of logic is "universal," some of which do, and some of which do not, have much to do with metatheory. We'll begin with two straightforward ways of interpreting the term and its associated line of reasoning, if only to clear the ground for the more interesting and more complex versions.

One sense in which one might be a "universalist" about logic is simply to hold that logic as a whole, i.e. that collection of principles underlying all correct inference, is universal in the sense that it applies everywhere and (hence) that it serves as the grounds of all justification and explanation. As Ricketts puts it:

Any explanation will draw on the principles of logic. In this way, logic, the maximally general science, provides a framework that embraces every science . . . . Indeed, because of logic's maximal generality, as Frege
Frege is, to be sure, a universalist in this sense. From the Fregean perspective, logic applies everywhere, and the relation of logical entailment is essential to the justificatory force of all explanations and lines of argument. From this form of universalism, it follows immediately that one can never engage in non-circular justifications of the fundamental principles of logic or of, in Dreben and van Heijenoort's (1986) term, "logic as a whole." Any such justification would presuppose the very principles it seeks to justify.

But this restriction has nothing to do with metatheory. Metatheory is never an attempt to justify logic as a whole; it is always an attempt to evaluate particular formal systems, i.e., particular codifications of those universal logical principles. At issue in metatheoretic investigations are the virtues of the codification, not of the underlying principles being codified.21

A second sense of "universalism," more in line with the first quotation above, is that in which a formal system $S$ is taken to be universal in virtue of being applicable everywhere, to all areas of inquiry. Frege was certainly a universalist in this sense; his view was that his formal systems, suitably modified by adding vocabulary as required, could serve as frameworks for presenting proofs not just about arithmetic but about any area of discourse in which rigor of proof was at a premium. But this form of universalism, again, brings with it no difficulties for metatheory: that a system is (intended to be) universally applicable in this sense is no barrier to our asking meaningful questions about it, e.g., about whether it does in fact have the expressive and deductive resources required for universal applicability, about whether it is reliable, and so on.

We turn now to conceptions of universalism that are arguably more promising in providing difficulties for metatheory.

The barrier to metatheory arises not when one takes it simply that logic in general is universal or when one takes it that one's formal system is universally applicable, but when one holds the considerably stronger thesis about that formal system that its derivations offer the only way of presenting compelling or scientifically-acceptable arguments. The idea that Frege holds such an exclusivist position about his formal system is taken to be supported by two lines of thought in Frege. The first is that the inadequacies of ordinary language, including primarily the ambiguity and unclarity of its terms, make ordinary language an unsatisfactory vehicle for the expression of compelling or scientifically-acceptable arguments. The remedy for this inadequacy is taken to be the expression of all such arguments via derivations in the formal system.

The second line of thought is that compelling or scientifically-acceptable arguments require expression within a system of "unified science." For Frege, as Ricketts reads him:

...understands justification and explanation, no other science can have justificatory or explanatory relevance to logic.20

The justification of knowledge is its systematization. Frege understands systematization through the lens of his logical work as the economical axiomatic formalization of a branch of knowledge within the framework provided by the *Begriffsschrift*.22

Similarly,

Proofs receive definitive expression only inside the framework the *Begriffsschrift* provides for all sciences.23

The idea that a formal system $S$ is universal in this sense—i.e., that it formalizes all of scientific inquiry in such a way that no scientifically-acceptable arguments can be given except via derivations in $S$—causes problems for metatheory in two ways.

The first difficulty is one of circularity: if the only way to present legitimate or scientifically-compelling arguments is via derivations in $S$, then all attempts to give such arguments presuppose the reliability of $S$. Hence, to try to argue in a legitimate or scientifically-compelling way for $S$'s reliability would be, by the universalist's lights, to reason in a very small and vicious circle.24

It is worth noting that this circularity-argument does not rule out all metatheory. The metatheoretic claims one might make about a formal system can be divided into two camps as follows. The first contains just those claims whose truth is presupposed when we treat derivations within the system as expressing justification-conferring proofs. Included here are most obviously the reliability of the system, its consistency, the truth-preservation of its derivations, and various forms of soundness. Into the second camp fall those claims about formal systems that are not presupposed in making such justificatory appeal to its derivations, i.e., those metatheoretic claims about a system whose falsehood regarding that system does not undermine its reliability. Included here, for example, are various claims of comprehensiveness for the system and completeness in the modern sense. That a system fails to include all of the derivations of a certain class does not mean that the derivations it does include are in any way faulty. The exclusivist position regarding a formal system $S$ immediately and obviously entails that those metatheoretic claims about $S$ that fall into the first camp—the "reliability" camp—cannot be non-circularly demonstrated. But there is no such quick route from exclusivism about $S$ to the indemonstrability of those second-camp claims (those from the "comprehensiveness" camp) about $S$. That we would have to presuppose the reliability of $S$ in order to demonstrate its comprehensiveness or its completeness—in the same way that $S$'s reliability would be presupposed in order to demonstrate anything at all, on this line of argument—does not make such demonstrations circular.25
The second difficulty posed for metatheory by this form of universality is that of semantic paradox. In order to obtain from the system of Grundgesetze a formal system $U$ sufficient not just for e.g. physics and astronomy, but also for semantics, one would need to add primitive terms for, among other things, the fundamental properties of interest in metatheory. And we know that, when the subject of the metatheoretic investigations is $U$ itself, this can't be done. If $U$ is consistent, then it must lack some of the resources (e.g., a well-behaved negation operator, referring terms for its own formulas, generous rules of sentence-formation, a well-behaved truth-predicate, and so on) that one might reasonably take to be essential for metatheory. If Frege held that the only scientifically-acceptable investigations are those that can be carried out in a formal system for unified science, then he held a view from which very little metatheory can count as scientifically acceptable.

Does Frege hold such a form of universalism? In particular, does he hold either (a) that scientifically-acceptable arguments can only be presented via derivations in a system $U$ of unified science, a system that contains the primitive vocabulary for all of the (coherent) sciences, or (b) the weaker claim that scientifically-acceptable arguments can only be presented via derivations in a rigorous formal system that incorporates essentially the logical resources of Frege’s own $G$ or $B$? Note that the global problems for metatheory arising from semantic paradox are difficulties only on the assumption of the strong thesis (a), since the weaker requirement that the metatheory of e.g. Frege's system $G$ is carried out in an appropriately-rigorous formal system is easily met by pursuing that metatheory in the system $G'$ obtained from $G$ by adding a truth-predicate for $G$. The problem cannot of course be circumvented if there's only one acceptable system $U$ within which both arithmetic and semantics must be carried out, as proposed by (a). The problem of circularity outlined above arises for metatheory on the assumption of either (a) or (b).

As to (a): Frege clearly held that the framework provided by his formal system was to have been applicable quite broadly. He remarks in 1882 that:

I did not wish to present an abstract logic in formulas, but to express a content through written symbols in a more precise and perspicuous way than is possible with words. In fact, I wished to produce, not a mere calculus ratiocinator, but a lingua characteristica in the Leibnizian sense. 26

Whether he means here that his system will be like a Leibnizian lingua characteristica simply in the sense of expressing determinate contents via its formulas, or in the further sense of (ideally) sufficing for the expression of all scientific thought, is not made clear in this text. Frege does talk in Begriffsschrift of possible later expansions of his system to include e.g. geometry and physics, and in [1882] discusses future expansion as follows:

I have attempted to supplement the formula language of arithmetic with symbols for the logical relations in order to produce—at first just for arithmetic—a conceptual notation of the kind I have presented as desirable. This does not rule out the application of my symbols to other fields. The logical relations occur everywhere, and the symbols for particular contents can be so chosen that they fit the framework of the conceptual notation. 27

The possibility of adding new primitive symbols to the language, and of thereby gaining a richer system applicable to special sciences, is one that Frege seems generally committed to. As he says, "the logical relations occur everywhere," and hence the perspicuous representation of proofs via Begriffsschrift-style derivations can be of use well outside the domain of arithmetic. 28

Frege does not, however, claim that there will be an eventual uberv-system containing the resources necessary to do all of science. More to the present point, he does not claim that scientific acceptability must await such a final expansion of his system to all of science. Quite the contrary. Frege's own system $G$ is far from being universal in the relevant sense, containing only extremely minimal expressive resources; but of course Frege takes derivations in $G$ to be very much up to the standards of scientific acceptability. (Indeed, as we'll see below, he takes them to go well beyond those standards.) It may well be that "universalism" in the strong sense of (a) was never meant by those advocating a universalist reading of Frege; in any case, this reading is a non-starter.

Considerably more interesting is the idea that the inadequacies of ordinary language prompt Frege to hold that scientifically-acceptable arguments can only be given via derivations in either $G$ or some expansion thereof. Frege does indeed claim that ordinary language is the source of a certain amount of logical confusion. He holds that an argument's presentation via a derivation in his formal system is a good way of warding off some of the difficulties that beset many arguments, difficulties having to do with the assumption of hidden premises, the use of ambiguous terms, and so on. But that these difficulties beset some non-formalized arguments is never cited by Frege as a reason to take all non-formalized arguments to be in any sense suspect or deficient. He does not claim that sufficiently-careful arguments carried out in natural language are to be regarded as second-rate, and does not claim that the presentation of arguments within his Begriffsschrift system is the only way to provide compelling or scientifically-acceptable grounds for conclusions.

The kinds of arguments ruled out by the universalism in question—namely, justification-conferring arguments carried out in a language other than that of a Fregean formal system—include not just arguments concerning the metatheory of such system but also all arguments in the history of mathematics (and elsewhere) prior to Frege's own work. The bold thesis that none of these arguments is sufficient to confer justification on its conclusion is one that does not sit well with Frege's
unremarkable attitude toward the mathematics of his predecessors—namely, that while some of it is problematic in very specific ways, mathematical proof has generally served to establish its conclusions. Commenting on Dedekind's considerably less-rigorous proof procedure, Frege notes:

[H]is procedure may have been the most appropriate for his purpose. . . . The length of a proof ought not to be measured by the yard. It is easy to make a proof look short on paper by skipping over many intermediate links in the chain of inference and merely indicating large parts of it. Generally people are satisfied if every step in the proof is evidently correct, and this is permissible if one merely wishes to be persuaded that the proposition to be proved is true.29

His own extraordinarily careful procedure, as Frege sees it, goes well beyond what's required simply to establish the truth of the propositions proven. His logicist purposes, which include not merely the demonstration of the truth of arithmetical claims, but a clear exhibition of the fundamental grounds of each of the (typically already-known) premises, requires that "considerably higher demands must be placed on the conduct of proof than is customary in arithmetic."30 The rigor imposed by expressing one's proofs as derivations in a Frege-style formal system is, in short, considerably more demanding than is required of arguments sufficient for establishing the truth of their conclusions, even within the demanding field of mathematics.

Frege does not claim that the work of previous mathematicians must be re-cast as formal Begriffschrift-style derivations in order to be persuasive, and does not claim that careful arguments couched in ordinary language or in ordinary-cum-mathematical language are unscientific or otherwise illegitimate. This is as one should expect of an author whose careful philosophical arguments in Grundlagen, e.g., to the effect that numbers are objects, that statements of number are assertions about concepts, that arithmetic is not about ideas, and so on, are intended to be taken seriously. Similarly for Frege's metatheoretical arguments in Begriffschrift and Grundgesetze.31

7.2.2 Internal Tensions

A final difficulty for the exclusivist account of Frege's understanding of his formal system is that it implies an important incoherence in Frege's views. As Ricketts [1985] notes, there is a clear conflict between the exclusivist thesis and Frege's general understanding of our appreciation and application of fundamental logical principles. As Ricketts sees it, the inconsistency is simply a part of Frege's views; one might, on the other hand, view it as a further reason not to read Frege in the proposed exclusivist way.

According to Ricketts [1985], Frege identifies logic with his own formal system. To return to the spatial metaphor introduced by Dreben and van Heijenoort [1986], the idea is that there is no logic "outside" of the system, which is to say that if we reason (in a way that is up to scientific standards), then we are invoking the principles explicitly laid down in Frege's construction of the formal system. To recognize in a sufficiently careful way that a conclusion follows from premises is to appeal—and to appeal only—to those axioms and rules that Frege has written down as rules for his formal system. As above, the fact that all knowledge of logical implications must, in order to be scientifically acceptable, appeal to those axioms and rules is, on this view, the heart of the problem for metatheory.

As Ricketts notes, this account is inconsistent with Frege's views about how we recognize simple and basic relations of logical implication between thoughts. There is a strain, as he puts it, between:

the role for logic defined by Frege's underlying view of judgment and his identification of logic as the maximally general science. For Frege, the ability to reason, to determine what follows from or contradicts what, is intertwined with the understanding of language, the grasp of thoughts.32

If you don't understand that the thought that it's raining is inconsistent with the thought that it isn't raining, then you have failed to grasp at least one of these thoughts. The tension arises because, on this account, Frege holds that:

The logician aims to state principles that enable us to make these determinations in a more regular, less haphazard fashion.33

That is to say, the logician writes down statements of logical principles such that the careful reasoner can reach the conclusions of arguments from premises just by appeal to those very statements. But if the recognition of basic relations of logical entailment and inconsistency is a necessary part of the grasp of those thoughts, then it cannot be the case that the recognition of these relations is mediated by the recognition of the truth of some further statements, those written down by the logician. Indeed, the problem is not just that the view of the logician's job as presented in the last sentence quoted is inconsistent with Frege's own view of our means of apprehending logical relations; the view itself is incoherent. For, as Ricketts notes, "no statement can formulate a logical principle."34 The difficulty is that to grasp a basic relation of implication or of inconsistency between some thoughts that one is currently entertaining is not to grasp some further thought. Hence the recognition of the most basic relations of implication and of inconsistency cannot be something that one does by simply following the instructions given by the statements of a particular formal system.
Following Ricketts [1985], it's helpful to note that this is just the difficulty pointed out by Lewis Carroll's fable of Achilles and the tortoise. In that story, the tortoise affirms a conditional and its antecedent but refuses to affirm the consequent of that conditional. The interesting point is that he also happily affirms a general statement of the rule *modus ponens*. But adding the statement of this rule to the stock of premises to which he assents makes no difference: the tortoise's position does not *thereby* become inconsistent. The tortoise is obliged to infer the consequent of the conditional not because of his affirmation of the additional premise but because of the logical connection between a conditional, its antecedent, and its consequent. If the tortoise is not already reasoning in accordance with this and other principles that link statements one to another, the addition of further statements to his set of accepted premises will never change that. In short, the recognition of simple and basic entailments and inconsistencies between thoughts cannot be a matter of recognizing the truth of further thoughts.  

Ricketts's [1985] view is that Frege is just mistaken about this important point, and that the rule-statements of his formal system are intended to serve as the court of appeal for careful reasoning. We are, on this account, supposed to be able to justify conclusions on the basis of premises just by adding to those premises an appropriate stock of Fregean rule-statements and axioms, independently of any underlying capacity to recognize logical relations between thoughts. This is the sense in which logical principles are "identified with" the rules of the formal system. And of course this cannot make sense: if we are not already capable of recognizing logical connections between thoughts, then we won't gain that capacity by adding new thoughts to our collection. Perhaps more to the point: if a conclusion does not in fact follow logically from premises, then adding to that stock of premises a statement of a logical rule will not bring about an entailment.

The picture of a formal system's rule-statements as forming the basis of scientific reasoning, to the exclusion of our ordinary capacity to recognize relations of entailment and inconsistency, is not just incoherent and inconsistent with Frege's views about our apprehension of basic entailment-relations; it is also not what Frege says. Frege does indeed take it that the invention of his formal system will bring a needed degree of preciseness and clarity to the giving of proofs. A derivation that follows the rules laid down in *Begriffsschrift* is sure to express a gap-free proof, one that demonstrates with a high degree of certainty that its conclusion follows logically from its premises. That is the whole point of the system. But the reason we can be sure that the derivations are reliable in this way is that we can see, via our ordinary and carefully-employed capacity to recognize logical entailments and inconsistencies, that the axioms and derivation-rules are well chosen. As above: the justification for including the sentential derivation rule *modus ponens* among the rules of the system rests on (1) the fact that "→" expresses the material conditional and (2) the fact that a material conditional and its antecedent together entail its consequent. This last fact is one that Frege relies on his audience to recognize, and it is only because we can and do recognize this fundamental pattern of entailment that we can see that the sentential rule is a reasonable one to include in a system whose derivations are supposed to express good proofs. The benefit of a formal system, from the Fregean point of view, is that by its use we are forced to make explicit, by making very small and clear steps of inference, exactly which inferential principles we appeal to in the course of a proof; it is not to somehow allow us to avoid the use of those principles.

To sum up the discussion to this point: Frege is clearly a universalist in a number of ways. He holds that the fundamental principles of logic apply everywhere; there is no area of inquiry in which these principles somehow "fail." He holds that his system of logic is applicable everywhere, given suitable expansions of primitive vocabulary. He arguably holds that every valid argument can in principle be shown to be valid via a derivation in (some expansion of) his formal system. But he does not hold that logical inferences can only be scientifically justified by appeal to the rules of his formal system. He takes it that not just his mathematical and scientific predecessors, but also he himself, can offer scientifically-acceptable arguments by expressing the thoughts in question in ordinary mathematical German and by employing the ordinary fundamental principles of reasoning that underlie all good argument. This is just what we see him doing in his early mathematical work, in the crucial argumentative passages of *Grundlagen*, and in the metatheoretic reasoning of *Begriffsschrift* and *Grundgesetze*.

### 7.3 Soundness, Completeness, and Consistency

There is nothing about Frege's conception of logic which rules out metatheory, where "metatheory" is understood as the systematic examination of the adequacy, in various senses, of a formal system. As we have seen, Frege engages in a certain limited amount of metatheoretic reasoning himself regarding his systems B and G, primarily to the effect that those systems are reliable.

But the amenability of Frege's general approach to broadly metatheoretic reflection does not mean that there is any natural affinity between Frege's position and what one might think of as the core of modern metatheoretic reasoning. As we've seen, Frege's conception of logical truth and logical entailment as applying to thoughts rather than to partially-interpreted sentences, and his view of logical truth and entailment in terms of provability via a core of primitive and self-evident logical principles, means that he cannot take model-theoretic truth or model-theoretic entailment as the touchstone of the logical relations. Hence, the completeness or the incompleteness of a formal
system of logic doesn't have the same significance from the Fregean point of view as it has from some post-Hilbertian viewpoints. There is no sense, from a Fregean point of view, in which a system's completeness is by itself a guarantee that the system includes all of the derivations it was intended to include, or in which a system's incompleteness is a guarantee that it "misses" some logical entailments.

This doesn't mean that a Fregean must be unable to make sense of, to prove, or otherwise to appreciate the implications of modern soundness, completeness, and incompleteness results. Given a language with a deductive and a model-theoretic apparatus, it is a straightforward question whether the extension of the model-theoretic entailment relation for that language includes or is included by the derivability-relation for that language, and a question whose demonstration requires nothing problematic from the Fregean point of view. It's just that the modern point of view from which the adequacy of a deductive system is a matter of its agreement with the model-theoretic entailment relation is one that a Fregean has no reason to share. In particular, the completeness of the first-order fragment of the systems $B$ and $G$, and the incompleteness of the full systems, are results that a Fregean can find significant and interesting; but they are results that provide, by themselves, reason neither to recommend nor to fault those systems.

A modern soundness result is of more interest from a Fregean point of view. A system falling soundness is one whose derivability-relation outstrips its model-theoretic entailment relation, so that for some $\Sigma$ and $\phi$, $\phi$ is derivable from $\Sigma$ despite the fact that some structure verifies each member of $\Sigma$ while falsifying $\phi$. If readings and structures for the language correspond in the way described in Chapter 6, then there will in such a case be an acceptable reading $\tau$ of the language such that each member of $\tau(\Sigma)$ is true and $\tau(\phi)$ is false. Assuming that the language is a Fregean one, in the sense that it comes with an intended assignment $\tau$ of thought to sentences, one can ask in this case two questions: (1) whether the truth of each member of $\tau(\Sigma)$ and the falsehood of $\tau(\phi)$ imply that the intended set of thoughts $\tau(\Sigma)$ fails to logically entail the intended thought $\tau(\phi)$; and (2) whether there is in such a case something problematic about the derivability of $\phi$ from $\Sigma$. As to the first: because for Frege the logical-entailment relation between thoughts has to do not just with the syntactic form of the sentences expressing them but also with the contents of the non-logical terms, there is no way to answer question (1) in full generality. In some cases, the answer will be affirmative, and in some cases negative: that the thoughts expressed by $\Sigma$ under one reading don't logically entail the thought expressed by $\phi$ under that reading doesn't imply that the thoughts expressed by $\Sigma$ under a different reading fail to logically entail the thought expressed by $\phi$ under that second reading. But things are clearer with respect to the second question. Though the non-logical terms in $\Sigma$ and in $\phi$ may have semantic values in virtue of which $\tau(\Sigma)$ logically entails $\tau(\phi)$, it will nevertheless be simply a mistake if the deductive system, which itself appeals only to the syntactic form of the sentences, enables one to derive $\phi$ from $\Sigma$. Because the deductive system is not sensitive to the contents of simple non-logical terms, but only to the syntax of sentences, a condition on its reliability is that it contain the derivation of $\phi$ from $\Sigma$ only if under every acceptable reading $\tau$, $\tau(\Sigma)$ logically entails $\tau(\phi)$. And this condition entails that under no acceptable reading is it the case that each member of $\tau(\Sigma)$ is true while $\tau(\phi)$ is false. In short, if for each structure there is a corresponding acceptable reading, as is typically the case, then the reliability of the deductive system requires that it satisfy a modern, model-theoretic soundness result. While soundness, relying as it does on a notion of "truth on a structure" that has little independent interest for Frege, would not have been a natural result for him to investigate, it is nonetheless a result in which a Fregean has a legitimate interest, given its status as a necessary condition for reliability.

Frege does of course intend that the formal system $G$ will provide a comprehensive set of axioms for arithmetic, and hence for various arithmetical sub-theories. Consider for example the theory of the natural numbers. Let $A_0$ be that collection of sentences of $G$'s language that express truths about the natural numbers. Because of his realism, Frege will take $A_0$ to be complete, in the sense of including for each sentence of the relevant fragment $L'$ of the language either that sentence or its negation. So the Fregean requirement is simply straightforward theory-completeness as restricted to $L'$: For every sentence $\phi$ of $L'$, either $G \vdash \phi$ or $G \vdash \neg \phi$.

The modern result most closely related to Frege's interests is a proof of consistency for a formal system. Consistency, i.e. the fact that no formula and its negation are both theorems of the system, is an obvious necessary condition for reliability in any language containing negation. A consistency-demonstration wouldn't have been a natural exercise for Frege to have included in Begriffsschrift or Grundgesetze, just because the stronger reliability result was, Frege thought, easily seen. But for a system not antecedently known to be reliable, a proof of consistency would count, from the Fregean point of view, as an important first step in ensuring reliability. The modern proofs that the first-order fragment of Frege's system is consistent, and indeed that some interesting second-order fragments are as well, are of real significance from the Fregean point of view.

7.4 Categoricity

Frege's defense of logicism was to have consisted in the proof not, of course, of all of the infinitely many truths of arithmetic, but of a handful of "basic" such truths from which, as he intended, the rest would in turn be provable. A particularly interesting collection of these basic truths is the collection governing the arithmetic of the finite cardinals, i.e. of the collection of objects bearing the ancestral of the successor-relation to 0. Here the basic truths are the claims that
the successor relation is a function, that no member of that collection bears the ancestral of the successor relation to itself, and that every member of that collection has a successor.

As Richard Heck points out, Frege establishes in his proof of Grundgesetze's theorem 263 that, as one might put it, these fundamental truths provide a categorical characterization of the natural numbers. That is, Frege establishes that, in Heck's words:

... [For any relation Q ∈ η and any object a] if Q ∈ η is functional, if the G's are the members of the Q-series beginning with a, if no member of the Q-series beginning with a follows itself in the Q-series, and if each G is related, by Q ∈ η to some object, then the number of G's is Endlos because the G's, ordered by [<_o], are isomorphic to the natural numbers, ordered by [<_pred], that is, by less-than. 41

This crucial result is one that might reasonably be called "metatheoretical," and it is worth asking to what extent it is so from a Fregean point of view.

To begin with, the usual modern way of framing such a categoricity result is as a claim about structures: taking some of the terms of the language as interpreted, we ask, as did Hilbert in the case of geometry, which coordinated assignments of meanings or extensions to those terms will satisfy the set of sentences. The result in this case is that any pair of assignments to "0" and to "Pred" that satisfies the basic sentences will respectively be an object o and a relation R such that the collection of objects to which o bears the ancestral of R is isomorphic (under <) to the natural numbers (under <).

Frege of course doesn't put it quite this way, given his emphasis not on the satisfaction of partially-interpreted sentences by structures but on the thoughts and the relations expressed by (parts of) fully-interpreted sentences. From Frege's point of view, the point is that of the truth, and indeed of the provability in G, of a single universally-quantified thought, one that says essentially that, for any pair co, R; that satisfies the relation defined by the basic sentences, the closure of o under R is isomorphic (under <) to the natural numbers (under <).

We can put the difference schematically as follows. Let BL(O, Pred) be the conjunction of the basic sentences that Frege derives concerning the natural numbers. The terms "0" and "Pred" are shorthand for longer expressions referring to (Frege's versions of) zero and the predecessor relation. BL(O, Pred) expresses a true thought about the natural numbers, one that's equivalent to the conjunction of the second-level Dedekind–Peano axioms. A modern version of the categoricity result is then the claim that every reinterpretation of "0" and "Pred" that satisfies BL(O, Pred) is isomorphic to every other, and in particular is isomorphic to the "real" <0, Pred>. The Fregean version is the claim (proven in G) that ∀x∀y[RBL(x, R) → ISO (<x, R>, <0, Pred>)).

From a mathematical point of view, the difference between the two versions is insignificant: Frege's axioms are categorical in one sense iff they're categorical in the other, and for essentially the same reasons. For our purposes, the differences worth noting are as follows. First, Frege's version is not in any sense "metatheoretical." It's simply a derivation within G of a sentence expressing a universal claim about all objects and relations; it makes no mention of sentences or interpretations. Second, there are some in-principle differences between the modern and the Fregean approach to such results that don't show up in this example but are worth noting:

1. Because the Fregean method requires that the result to be proven be expressed as a conditional whose antecedent expresses the conditions laid down in the statements whose categoricity is in question (here, our "basic sentences"), the Fregean approach works only when such a conditional sentence can in fact be formed. If, for example, the basic sentences had been infinite in number, there would be no way of framing the categoricity result in G (or in any other finitary language). The modern method has no such restriction.

2. If one is working in a metatheory that's proof-theoretically stronger than the theory whose sentences are in question, it is possible for the modern result to be provable while the Fregean result is merely true but not provable.

3. If the metatheory has a richer domain than the object theory, the modern claim may be true while the Fregean claim is false. (Consider the case in which the object-language quantifies over a domain lacking the function whose existence is asserted in the isomorphism claim.)

That Frege's version of the theorem isn't metatheoretical is as one should expect, given Frege's interests and his view of the importance of the theorem. For Frege, the question of whether the basic laws he has proven will suffice for the rest of the arithmetic of the natural numbers would have been importantly bound up with the question of whether those laws sufficed to fix the overall structural features of that collection of numbers. The failure of the "isomorphism" result would have meant that some essential features of the collection of natural numbers had not yet been characterized by the basic laws proven to this point. But this isn't, from Frege's point of view, naturally viewed as a point about other interpretations of the terms "0" and "Pred." It's rather a (very similar) point expressible by a sentence that explicitly quantifies over the positions held by those terms. This is just the difference between Hilbert's understanding of his consistency-proofs as involving reinterpretation and Frege's recasting of the point of those proofs as involving universally-quantified statements with fixed meanings. 42

There is nothing to prevent Frege from stepping back and asking the modern isomorphism question. That he doesn't do so has to do, again, with (1) his historical
location, and (2) the fact that he does not think of the sentences themselves as the objects of interest. Taking sentences to be of interest only in virtue of their expression of determinate thoughts, questions about the models of a set of sentences are not natural ones to ask. Considerably more natural from this perspective are the roughly-equivalent Freganean questions of the truth of those universally-quantified thoughts whose antecedents are expressed by the conjunction of the sentences in question.

7.5 Conclusion

The metatheoretic issues with which Frege is explicitly concerned all have to do with the reliability of his formal system, and include (1) the reliability of individual rules and axioms, (2) the uniquely-referring nature of the sentences and subsentential pieces of language, and (3) the eliminability of defined terms. This is as we should expect. The formal system, from Frege’s point of view, is a tool with a very specific purpose: to serve as a means for expressing proofs, a means that is guaranteed to present proofs in such a way that no step can “sneak in unnoticed” in the chain of inferences from premises to conclusion. A proof expressed by a derivation in such a system will demonstrate, with as much certainty as is had by our conviction of the reliability of the system, that its conclusion follows logically from the premises explicitly laid out in that proof. A reliable such system of proof is just what Frege needs in order to demonstrate the logicist thesis. That Frege’s only real metatheoretical concern is that of the reliability of his system is unsurprising: without the reliability of the formal system, the proof-theoretic project is bankrupt, and with it, the system meets the most important criterion for the use to which Frege wants to put it.

The other crucial criterion is that the system be sufficiently rich. That is, it is essential to Frege’s logicist reduction that his formal system incorporate enough logical principles to carry out the necessary proofs. The concern with richness in this sense is again what underlies the move from B to G. If Frege had been right, first, that the whole of the arithmetic of the natural numbers was provable from his version of the Peano-Dedekind axioms and, second, that the system of Grundgesetze sufficed for the proof of those axioms, then the system G would have been sufficiently rich for the logicist project as applied to the theory of the natural numbers.

A stronger requirement is that of the comprehensiveness of a system, i.e. the condition that for each of its sentences φ and sets Σ of sentences, φ is derivable from Σ if the thought expressed by φ follows logically from the set of thoughts expressed by the members of Σ. From Frege’s perspective, there is no systematic way to test whether a formal system is comprehensive, i.e. whether it incorporates all of the principles of valid logical inference applicable to the thoughts expressed in its language. For there is no general testable condition met by all and only those pairs <Σ, φ> of a set of formulas and a formula, respectively, such that the first expresses a set of premises that logically entail the conclusion expressed by the second. No condition on such pairs plays for Frege the role that model-theoretic entailment plays for Tarski.48 And, as we’ve seen, this is due not just to Frege’s historical location, but also to philosophical principle: that a pair <Σ, φ> is (not) truth-preserving under every interpretation is not, for Frege, a guarantee that τ(Σ) logically entails (fails to entail) τ(φ).

That Frege has no non-circular way of circumscribing the formulas that express truths of logic, or the pairs <Σ, φ> that express logically-true arguments, means that he must adopt what’s been called an “experimental” approach to the comprehensiveness of his system.49 If we come across a principle of logical inference that isn’t yet reflected in the formal system we’re using, we simply add the relevant axiom or rule of inference to that system. Or, at least, we do so if we’re aiming for comprehensiveness. It’s worth recalling that the comprehensiveness of his formal system is by no means a requirement for the use to which Frege puts it. What’s required for the logicist project is that the formal system reflect in its axioms and rules enough principles of logical inference to prove the fundamental truths of arithmetic. If there are further basic principles that aren’t needed for arithmetic, their absence from the Fregean system won’t be an impediment to the central use he makes of it.

To summarize: of the now-standard metatheoretic questions that Frege does not raise, some, most notably completeness, would have been of little interest from his point of view. Others, most notably soundness and consistency, would have been of considerable interest. Yet others, notably categoricity results, are close cousins to results that were of clear importance to Frege. None of these modern results would have been of any sense incoherent, and none would have been flawed in the way that, from Frege’s point of view, Hilbert-style consistency-proofs for theories are flawed.

Notes

1. Begriffsschrift (hereafter BfF) Preface p vii/1–2. This passage is followed immediately by the sadly prophetic:

A dispute can arise, so far as I can see, only with regard to my Basic Law concerning courses-of-values (V), which logicians perhaps have not yet expressly enunciated, and yet it is what people have in mind, for example, when they speak of the extensions of concepts. I hold that it is a law of pure logic. In any event the place is pointed out where the decision must be made.

2. Grundgesetze (hereafter Gg) 1 Introduction p vii/3–4. This passage is followed immediately by the sadly prophetic:

A dispute can arise, so far as I can see, only with regard to my Basic Law concerning courses-of-values (V), which logicians perhaps have not yet expressly enunciated, and yet it is what people have in mind, for example, when they speak of the extensions of concepts. I hold that it is a law of pure logic. In any event the place is pointed out where the decision must be made.

3. See BfF §24.

1. Begriffsschrift (hereafter BfF) Preface p vii/1–2. This passage is followed immediately by the sadly prophetic:

A dispute can arise, so far as I can see, only with regard to my Basic Law concerning courses-of-values (V), which logicians perhaps have not yet expressly enunciated, and yet it is what people have in mind, for example, when they speak of the extensions of concepts. I hold that it is a law of pure logic. In any event the place is pointed out where the decision must be made.
4. A fortress §§5-6, emphasis added.
5. A fortress §14, emphasis added.
6. A fortress §14. Note the typographical error in the antepenultimate formula cited as it appears in van Heijenoort [1967b].
7. Gg 1 Introduction p xii/15.
8. To say that an individual derivation-rule is reliable is to say that it sanctions the inference of a sentence $\phi$ from a set $\Sigma$ of sentences only if the thought expressed by $\phi$ follows logically from those expressed by the members of $\Sigma$, similarly, an axiom-sentence is reliable if the thought it expresses is a truth of logic.
9. Gg 1 Introduction p xii/9. See also Gg 1 §5 footnote: "Denotational names must not occur in the Begriffsschrift."
10. This is the argument whose conclusion is, "Thus without contradicting our setting $\xi \Phi(x) = \xi \Psi(x)$ it is always possible to stipulate that an arbitrary course-of-values is to be the True and another the False" (Gg §10).
11. This is a slightly misleading way of putting it, since the term-forming operator of Frege's system refers to a first-level function; its ordinary use will be to operate on the course of values of a concept, yielding the single object falling under that concept (if there is one) and that course of values itself otherwise. (As applied to an object other than a course of values, it will revert to that object as value.)
12. This argument, as has been often remarked (including by Frege himself), is shown by Russell's paradox to be flawed, for the paradox demonstrates the existence of course-of-value names that, though well-formed by Frege's lights, do not achieve a (unique) denotation via his stipulations. For Frege's reaction to this effect, see the letter to Russell, June 22, 1902.
13. "Namely, by our stipulations it is determined under what conditions the name denotes the True. The sense of this name—the thought—is the thought that these conditions are fulfilled" (Gg §32).
14. Gg §13, emphasis added. As Frege notes here, this is just one case of such interchangeability, each of which ought (in principle to be justified separately, but: `So as not to become tied up in excessive complexity, I have wished to assume this interchangeability generally granted, and to make use of it in future without further explicit mention.' Ibid.
15. Gg §13.
16. Gg §15. I have suppressed Frege's indexes $(\alpha)$ and $(\beta)$, respectively, on the first two propositions quoted. The rule is called the `second' method of inference despite the previous introduction of both the interchangeability of subcomponents and modus ponens because Frege does not honor the first of these with the title `method of inference'.
17. Gg §14.
21. For a nice discussion of this issue, see Sullivan [2003].
23. Ibid., p 177.
24. As Ricketts puts it (ibid.):
A formalized proof of the soundness of Frege's codification of logic would thus presuppose the truth of the axioms and the soundness of the inference rules of that codification. Given Frege's view of justifications as explanation within unified science, such a proof would be scientifically pointless.
25. That is, the quick argument just outlined that gets one from exclusivist about S to the circularity of demonstrations of S's reliability is not available in the case of demonstrations of S's comprehensiveness, etc. This does not mean that there is no argument from this premise to the conclusion in question: one might well wonder whether the exclusivist
Many of Frege's best ideas have survived the conceptual and technical developments of the last 100 years. His idea that we can so structure the expression of content that good steps of logical inference are encoded in syntactic transformation rules has turned out to be more significant than he could have imagined. He was by no means unique in pursuing the syntactic expression of logically-important properties of and relations between contents, but his combination of this general idea with his innovative analysis of content in terms of quantification, together with his beautifully-designed notation for the expression of content so analysed, made his work a watershed in the development of logic. So too, Frege's analysis of the fundamental notions underlying a theory of, as he puts it, "following in a series" and of a theory of cardinality have survived to become entirely commonplace.

Some of the central Fregean ideas that haven't survived have perished for good reason. Frege's three-part combination of (1) logicism, (2) realism about arithmetical truth, and (3) the view that the principles of logic form a surveyable collection, together form a view that despite its initial plausibility we now know via the first incompleteness theorem to be clearly false. Whether Frege's logicist thesis itself—i.e. the thesis that the truths of arithmetic are each provable from purely-logical premises via purely-logical steps of inference—is itself plausible is a further question, and one not so easily settled. But one thing we know, again via incompleteness, is that Frege's means of demonstrating the thesis with a proof-theoretic reduction of the axioms of arithmetic can't work: there is no such manageable collection of axioms for arithmetic, again assuming roughly Fregean realism about arithmetic. If the truths of arithmetic form a collection that Frege would have recognized, then logicism in his sense is plausible only if the truths of logic form a much less-manageable collection than he seems to have supposed; the collection must in particular not be recursively enumerable.

Frege's view that there are extensions—i.e. objects identifiable via the equivalence between sentences of the form "\( \forall x (Fx \equiv Gx) \)" and "the extension of \( F \) = the extension of \( G \)—is decisively falsified by Russell's paradox. Also falsified along with it, arguably, is the very liberal kind of "re-carving" of content that was to have been