The Russell Paradox in the Theory of the *Grundgesetze* and Why Frege could not have done Better

Jönne Speck*

8th January 2010

Introduction

This essay examines which assumptions lead to the contradiction in the theory of Frege's *Grundgesetze* (§1) known as the 'Russell paradox', and discusses whether Frege could have avoided it (§2). It concludes that Frege could not have saved his theory.

Bertrand Russell discovered the paradox named after him when studying Cantor's set theory [Russell, 2009, p. 44]. Irritated by his finding, Russell turned to Frege's works, hoping that there he would find a solution [Burgess, 2005, p. 32]. However, he soon realized that Frege's theory contained the same inconsistent assumptions [Russell, 1967].

The present study connects with this early insight of Russell's. It introduces the paradox as a contradiction within naïve set theory (§1.1), and then shows how the theory of Frege's Grundgesetze entails naïve set theory (§1.2), and consequently the Russell-paradox (§ 1.3). Subseqently (§ 2), I ask whether Frege could have blocked the contradictory reasoning and examine three different ways he might have gone (§2.1 to §2.3). None of these, however, proves feasible due to basic commitments of Frege's.

*js784@st-andrews.ac.uk

1 Which Assumptions lead to the Russell Paradox

1.1 The Russell Paradox in Naïve Set Theory

Naïve set theory is based on classical, second order logic. In addition to the standard axioms of extensional set theory, its characteristic axiom is[1],

Naïve Comprehension $\forall F\,\exists x$ (x is a set and $\forall y(Fy$ if and only if $y \in x)$

Naïve set theory is inconsistent, by the following reasoning, known as the 'Russell paradox'. '$\neg x \in x$' is a proper predicate of set theory. This 'Russell predicate' is true of all and only those sets which do not fall into themselves.

Now, by naïve comprehension, there is a set of all sets which do not fall into themselves. Let it be called the 'Russell set' or simply 'r'. Now, r either falls into itself or not. So let us assume that it does. However, r comprises only those sets which fall into themselves. Therefore, if r falls into itself it does not fall into itself. Thus, by classical *reductio*, naïve set theory proves that r cannot fall into itself.

But r comprises all sets which do not fall into themselves. Therefore, it does fall into itself. Consequently, naïve set theory proves both that r does not fall into itself and that it does. Thus, naïve set theory proves a contradiction. It is inconsistent.

1.2 The Theory of the *Grundgesetze*

This section identifies the theory developed by Frege in the first two sections of his *Grundgesetze der Arithmetik* [Frege, 1998], and shows how it entails naïve set theory, and consequently the Russell paradox.

I begin with the logic of this theory (henceforth 'GG'). In §47 of the *Grundgesetze* Frege summarizes its axioms and in §48 its rules of inference. They comprise the propositional and predicate calculus which Frege had developed as early as in his *Begriffsschrift* [Frege, 1879][2], and which amounts to complete first order logic.

However, GG goes beyond this. In addition, its language contains free variables 'f', 'g' and 'h' which can be replaced by any formula (rule 9 of §48). In fact, all the axioms of GG are formulated by means of these latin letters. Bound second order variables ('$\mathfrak{f}$', '$\mathfrak{g}$' and '$\mathfrak{h}$') are governed by a specialization axiom (*IIb*) which has the same form as the first order specialization (*IIa*). In effect, second order quantification in GG ranges over all functions expressible in the language. Finally, GG entails full second order

[1] See [Cantor, 1932, p. 204, p. 282] and for a more contemporary exposition, [Boolos, 1998, §1]
[2] An insightful survey is found in [Sullivan, 2004].

comprehension [Boolos, 1985, p. 337]: $\exists \mathfrak{f} \forall \mathfrak{a}(f(a) \equiv \mathfrak{f}(a))$. In sum, *GG* includes classical second-order logic.

Frege extends his classical second order logic by a theory of 'value-ranges'. The value-range of a function registers which arguments it maps to which values. For example, the value-range of the natural logarithm function pairs numbers with the power to which the Euler number must be raised in order to produce the number. Frege implements this idea in *GG* as follows. The language is extended by a symbol ' $\acute{}$ ' which binds lower Greek vowels ('ε', 'α' etc.) as variables, and is governed by a formation rule such that for any function expression '$f(\xi)$', '$\acute{\varepsilon} f(\varepsilon)$' is a term. This term refers to the value-range of the function $f(\xi)$. Thus, '$\acute{\varepsilon}\Phi(\varepsilon)$' itself stands for a function, namely from functions into their value-ranges.

However, '$\acute{\varepsilon}\Phi(\varepsilon)$' is not defined explicitly. Instead, ' $\acute{}$ ' is taken as a primitive symbol and governed by the axiom V of *GG*.

V. $$\acute{\varepsilon} f(\varepsilon) = \acute{\varepsilon} g(\varepsilon) \equiv \forall \mathfrak{a}(f(\mathfrak{a}) \equiv g(\mathfrak{a})))$$

V says that functions have the same value-range if and only if they map the same objects to the same values. V entails, by the underlying classical second-order logic that every function has a value-range. Since this '*GG*-comprehension' will prove crucial in due course, I would like to make its derivation explicit.

1. $\acute{\varepsilon} f(\varepsilon) = \acute{\varepsilon} g(\varepsilon) \equiv \forall \mathfrak{a}(f(\mathfrak{a}) \equiv g(\mathfrak{a}))$ — axiom V
2. $\acute{\varepsilon} f(\varepsilon) = \acute{\varepsilon} f(\varepsilon) \equiv \forall \mathfrak{a}(f(\mathfrak{a}) \equiv f(\mathfrak{a}))$ — rule 9
3. $\forall \mathfrak{f} \exists \mathfrak{a}(\mathfrak{a} = \acute{\varepsilon}\mathfrak{f}(\varepsilon))$ — 2, classical second order logic

A special case of value-ranges are the extensions of concepts. On Frege's view [Frege, 1998, §3], concepts are special functions, namely those which map objects into the truth values . For example, the extension of the concept of primeness is the value-range which pairs all numbers which are prime with the True, and all the numbers which are not prime, with the False. This value-range Frege identifies with the extension of the concept.

Since all concepts are functions, their extensions are all value-ranges. *GG*, with its axiom *V*, contains a theory of value-ranges. If the latin letters in V are restricted to concepts, the result is a theory of concept-extensions. Thus, *GG* contains a theory of concept-extensions.

By means of it, finally, a relation of elementhood can be defined [Frege, 1998, § 34]. a is an element of u iff u is the extension of some function $\mathfrak{g}$ which maps a to itself. Given this definition, theory of extensions contained proves two important theorems [Zalta, 2009, §

2.4]. First, V entails that an objection falls into an extension if and only if the corresponding concept applies to it. Second, it proves that two extensions are identical if and only if they contain exactly the same elements. Elementhood among concept-extensions thus behaves just like elementhood among sets. Thus, concept-extensions can be regarded as sets, and GG-comprehension amounts to naïve set-theoretical comprehension 1.1.

In conclusion, GG contains a classical second-order theory of concept-extensions which entails naïve comprehension. Thus, GG contains naïve set theory.

1.3 The Paradox in the *Grundgesetze*

The theory of Frege's *Grundgesetze* ('GG') contains naïve set theory, as the preceding section has shown. Naïve set theory, however, allows for the contradictory reasoning of the Russell paradox, as shown before (p. 2). Consequently, GG as well is inconsistent. This section explains how the Russell paradox is engendered within GG.

To prepare later discussion, I distinguish between three steps: A, B and C.

A. Since the language of GG includes full second order quantification (p. 3), the Russell-property can be expressed in the language of the *Grundgesetze*: '$\exists \mathfrak{g}(\xi = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\xi))$'. According to Frege, every well-formed formula refers to a function. Open sentences, especially, refer to functions from objects to the truth-values, in other words, concepts. Thus, the above formula refers to the concept of being an extension which does not fall under its own concept. Thus, second order quantification generates a 'Russell-concept'.

B. Secondly the theory of concept-extensions captured by Frege's axiom V yields a corresponding extension. By rule 9, one of the latin letters in V can be replaced by the formula '$\exists \mathfrak{g}(\xi = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\xi))$', and by the reasoning from above V entails that $\exists \mathfrak{a}(\mathfrak{a} = \acute{\alpha}`\exists \mathfrak{g}(\alpha = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\alpha)))$. In prose, V entails that there is an extension of all extensions which do not fall under their concepts. It corresponds to the Russell set of naïve set theory. Let it be called 'r'.

C. Finally, GG also provides the means to complete the contradictory reasoning. Recall its first step (p. 2): If the set of all sets which do not fall into themselves, falls into itself, then it does not fall into itself. This can be derived within Frege's theory, from the following instance of the second-order substitution axiom *IIb*:

1. $\forall \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \supset \mathfrak{g}(r)) \supset (r = \acute{\alpha}(\exists \mathfrak{g}(\alpha = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\alpha))) \supset (\exists \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(r))))$

By the meaning of 'r' this amounts to

2. $\forall \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \supset \mathfrak{g}(r)) \supset \exists \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(r)))$.

This conditional says: If r falls into itself then it does not fall into itself.

Moreover, the substitution of '$\neg \exists \mathfrak{g}(\xi = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\xi))$' into V provides

3. $r = \acute{\varepsilon}f(\varepsilon) \equiv \forall \mathfrak{a}(\exists \mathfrak{g}(\mathfrak{a} = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(\mathfrak{a})) \equiv f(\mathfrak{a})))$.

By classical second-order logic, this entails

4. $\exists \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(r)) \supset \forall \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \supset \mathfrak{g}(r))$.

Thus, V provides the second step of the Russell paradox: If r does not fall into itself then it falls into itself.

However, GG presumes that for any concept, everything either falls into it or not. Especially, the underlying logic derives that either r falls into itself, or it does not.

5. $\forall \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \supset \mathfrak{g}(r)) \vee \neg \forall \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \supset \mathfrak{g}(r))$

Together with 2 and 4, this entails that r falls into itself *and* does not. GG is inconsistent.

2 Could Frege have avoided the Paradox?

Having finished the first volume of his Grundgesetze, which contains the theory GG as sketched in the preceding section, Frege was confident that he had provided a sound theory [Frege, 1998, p. XXVI, my translation].

> 'I could only accept as a refutation, [...] if someone would demonstrate that my principles lead to obviously false theorems. That, however, nobody will achieve.'

Alas, Russell's paradox has shown that GG is inconsistent, and therefore in fact proves every falsehood. Russell informed Frege about the paradox on June 16, 1902 [Russell, 1967]. In this section I discuss whether Frege could saved his theory, given his logical and philosophical commitments in 1902.

2.1 Frege could not have given up Second Order Quantification

Classical second order logic is consistent[3]. The logical basis of GG alone therefore cannot be held responsible for the Russell paradox. However, the first step of the paradox (A)

[3] See, for example, [Shapiro, 2000, theorem 4.6]

was the formulation of the concept of being an extension which does not fall under its concept. Since being a concept-extension is a second order concept, it can only be expressed by means of second order quantifiers: $\exists\mathfrak{g}(\xi = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(\xi))$. Only because the language of GG provides second order quantification, therefore, the Russell concept could be denoted in the first place. Although second order logic is not sufficient, it thus proves necessary for the paradox. If Frege's theory had not been formulated in a second order language, therefore, step A (p. 4) could not have been taken and the Russell paradox had not even taken off the ground.

This hypothesis finds strong support in a result due to Terence Parsons [Parsons, 1995]. He develops a weaker variant of GG by restricting the underlying logic to a complete first order calculus with identity and replacing the axiom V by a first order schema [Parsons, 1995, p. 424]. This may be called the 'first order portion' of the theory of the *Grundgesetze*. Parsons then construes a model [Parsons, 1995, pp. 425ff][4] for this first order theory and thus proves it consistent. However, this solution was not available to Frege. He was committed to GG being a second order theory.

Frege's life-long project, which the *Grundgesetze* were supposed to complete, was the derivation of arithmetic and analysis from logic. Thus, every arithmetical truth is supposed to be derivable from GG. Certainly, this presupposes that all these truths can be expressed in the language of GG. Many mathematical propositions, however, are about properties and relations, in other words, involve second order quantifiers[5]. In fact, on Frege's view the commitment of mathematics to second order quantification is fundamental. Frege's definition of the natural numbers famously relies on the relation of equinumerosity. The definition of 'ξ is equinumerous to ζ', however, involves second-order quantification [Frege, 1998, §§ 38 - 40]. Consequently, Frege could not have given up on second order quantification, since this would have meant to loose the basis of his foundationalist project.

2.2 Frege could not have given up Basic Law V

The second step of the paradox (B) was possible because GG contains the theory of value-ranges, and especially of concept-extensions, captured by the axiom V. Together with the underlying second order logic, V entails GG-comprehension (p. 1.2). By it, Frege's theory necessitates the extension $\acute{\alpha}(\exists\mathfrak{g}(\alpha = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(\alpha))$ which is just the Russell set r of all sets which do not fall into themselves, in Frege's terms.

[4] See also [Boolos, 1993, p. 226].

[5] The paradigm example is the principle of number-theoretical induction: For every property F, if 0 is F, and whenever some number is F, its successor is F, too, then all natural numbers are F.

If GG had not contained the theory of concept-extensions delivered by the axiom V, step B could not have been taken, and the paradox would be blocked. In fact, when Frege, in considered the paradox in his *Grundgesetze*, his response was to replace V [Frege, 1998, pp. 262ff.]. The correction, however, is minimal. Frege weakens merely the left-to-right direction of V. Not for all objects now two functions have to take the same value if their extensions are identical. This does not any longer hold of these very extensions.

$$V' \qquad \acute{\varepsilon}f(\varepsilon) = \acute{\varepsilon}g(\varepsilon) \equiv \forall \mathfrak{a}(\mathfrak{a} \neq \acute{\varepsilon}f(\varepsilon) \supset f(\mathfrak{a}) \equiv g(\mathfrak{a})))$$

Frege was confident that if one replaces V by the weaker V', the resulting theory on one hand does not allow for the contradictory reasoning of the paradox but one the other hand, still suffices to derive arithmetic.

However, these hopes were bound to be frustrated. The resulting theory fails to prove even that there is more than one object, as Leśniewski showed[6], although not until after Frege's death. Frege may well have discovered this by himself, or a similarly result, but kept understandable silence about this failure of his solution. At any rate, he gave up entirely on the theory of his *Grundgesetze*.

2.3 Frege could not have given up sharp concepts

However, it may seem as if Frege despaired too quickly as if there was still an option for him to avoid the paradox.

The paradox could be completed (C) since it had to be presumed within GG that any concept either applies to a given object or not. Put more metaphorically, Fregean concepts have sharp boundaries. Therefore, it could be assumed that r either is an extension which does not fall under its concept, or not. This validated the assumption 5, which completed the contradiction.

The contradictory reasoning therefore can be blocked in its final step if one gives up the assumption that r either falls into itself or not. However, 5 does nothing but reformulate the assumption that the concept of not falling into oneself has a sharp boundary. Giving it up, would mean to accept that some concept-extensions do not have sharp boundaries. This, however, was not available to Frege.

As outlined above (p. 3), Frege took concepts to be a special kind of functions, those which map objects[7] to the truth-values the True or the False. About functions

[6] Leśniewskihimself never published the result, but it is reported by Sobociński [Sobociński, 1949, pp. 220ff], also compare [Geach, 1956].

[7] Respectively, concepts of nth order, if they are of order $n+1$ themselves.

generally, again, he held that each of them must be defined for every object[8] whatsoever [Frege, 1960a, p. 33]. Frege thought that every function must be total. Especially, any concept must map any object to some truth-value.

For the present case this means that on Frege's view, the Russell concept $\exists\mathfrak{g}(\xi = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(\xi))$ must map any object either to the True or to the False. By comprehension, one object is the extension of that very concept, $\acute{\alpha}\exists\mathfrak{g}(\alpha = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(\alpha))$, i.e. the Russell set r. Now, the Russell concept maps r either to the True or to the False. Hence, either $\exists\mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(r))$ or $\neg\exists\mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(r))$. Therefore, 5 is an immediate entailment of Frege's presumption that every function is defined for every object. It may be asked whether Frege could have given up on this presumption, and allowed for partially defined concepts. He could not have done so, or so I shall argue, due to a fundamental principle of his philosophy of language.

According to the semantics of the Grundgesetze [Frege, 1998, §2], sentences are terms[9]. The objects to which they refer are the truth-values. The sentence '7 is prime' for example refers to the True only if the concept of primeness maps 7 to the True. If Frege had allowed for partial concepts, however, some sentences would be neither true nor false. To solve the Russell paradox, for example, he would have to say that '$\exists\mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg\mathfrak{g}(r))$' neither refers to the True nor to the False. Thus, Frege would have been forced to accept that for some terms, there is nothing to which they refer.

However, this would have been incompatible with the philosophy behind the *Grundgesetze*. Frege was convinced that [Frege, 1960b, p. 58]

> If words are used in the ordinary way, what one intends to speak of is their reference.

Accordingly, there is nothing a term without reference is about, and sentences without truth-value fail to say anything.

Of course, Frege did not deny that in ordinary discourse, we often use terms without referents. One example he himself used occasionally [Frege, 1960b, p. 62], [Frege, 1979, p. 191] is 'Odysseus'. However, this was only possible in contexts, such as fiction or poetry, where words are not used 'in the ordinary way' and what 'one intends to speak of' is their sense.

The *Grundgesetze*, however, were certainly not meant to be fiction. Instead, it formulates a theory, *GG* (§1.2 above), to which Frege ascribes a significant scientific value. However, for *GG* to provide a logicist foundation of arithmetic, it is necessary that its

[8] or nth order concept

[9] In Frege's jargon, 'proper names'

language is 'used in the ordinary way'. Therefore, it is a basic requirement of Frege's project that for any term which can be formed within the language of the Grundgesetze, there is something to which it refers. Consequently, the sentence '$\exists \mathfrak{g}(r = \acute{\varepsilon}\mathfrak{g}(\varepsilon) \wedge \neg \mathfrak{g}(r))$' needs to be either true or false. Thus, the assumption 5 from the derivation of the paradox is a basic commitment of Frege's philosophy and the project of the *Grundgesetze*. It could not have been given up, and therefore Frege could not have blocked the final step of paradox, either.

Conclusion

This essay has explained how the theory of Frege's Grundgesetze is proved inconsistent by the Russell paradox (§1). It has surveyed possible solutions (§2), none of which turned out to be compatible with the project of the *Grundgesetze* and Frege's philosophical commitments.

First, (§2.1) I examined whether Frege could have prevented the Russell concept from being formulated in the first place, by giving up on second order quantification. This turned out to be incompatible with Frege's project of deriving arithmetic. Secondly (§2.2), I briefly considered the solution offered by Frege himself in the appendix to the *Grundgesetze* and remarked on its failure. Finally (§2.3), I argued that Frege could not have allowed that the Russell concept is neither true nor false of the Russell set.

References

[Boolos, 1985] Boolos, G. (1985). Reading Begriffsschrift. *Mind*, 94(375), 331 – 344.

[Boolos, 1993] Boolos, G. (1993). Whence the Contradiction. *Proceedings of the Aristotelian Society*, 67, 213 – 233.

[Boolos, 1998] Boolos, G. (1998). The Iterative Conception of Set. In *Logic, Logic and Logic* (pp. 13 – 28). Harvard University Press.

[Burgess, 2005] Burgess, J. P. (2005). *Fixing Frege*. Princeton, New Jersey: Princeton University Press.

[Cantor, 1932] Cantor, G. (1932). *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Springer.

[Frege, 1879] Frege, G. (1879). *Begriffsschrift*. Louis Nebert.

[Frege, 1960a] Frege, G. (1960a). Function and Concept. In P. Geach & M. Black (Eds.), *Philosophical Writings* (pp. 21 – 41). Oxford: Basil Blackwell.

[Frege, 1960b] Frege, G. (1960b). On Sense and Reference. In P. Geach & M. Black (Eds.), *Philosophical Writings* (pp. 56 – 78). Oxford: Basil Blackwell.

[Frege, 1979] Frege, G. (1979). Introduction to Logic. In H. Hermes, F. Kambartel, & F. Kaulbach (Eds.), *Posthumous Writings*. Oxford: Basil Blackwell.

[Frege, 1998] Frege, G. (1998). *Grundgesetze der Arithmetik: Begriffsschriftlich abgeleitet*. Georg Olms.

[Geach, 1956] Geach, P. T. (1956). On Frege's Way Out. 65(259), 408 – 409.

[Parsons, 1995] Parsons, T. (1995). On the Consistency of the First-Order Portion of Frege's Logical System. In *Frege's Philosophy of Mathematics* (pp. 422 – 431). Harvard University Press.

[Russell, 1967] Russell, B. (1967). Letter to Frege. In J. van Heijenoort (Ed.), *From Frege to Gödel* (pp. 124 –125). Cambridge, Mass.: Harvard University Press.

[Russell, 2009] Russell, B. (2009). My Mental Development. In R. Egner & L. Dennon (Eds.), *The Basic Writings of Bertrand Russell, 1903 - 1959* (pp. 37 – 50). London: Routledge.

[Shapiro, 2000] Shapiro, S. (2000). *Foundations without Foundationalism*.

[Sobociński, 1949] Sobociński, B. (1949). L'analyse de l'antinomie russellienne par Lesniewski. In *Methodos*, volume 1.

[Sullivan, 2004] Sullivan, P. (2004). Frege's logic. In *Handbook of the History and Philosophy of Logic*, volume 3 (pp. 671 – 762). Elsevier BV.

[Zalta, 2009] Zalta, E. N. (2009). Fregeś logic, theorem, and foundations for arithmetic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Summer 2009 edition.