INS AND OUTS OF RUSSELL'S THEORY OF TYPES

by

ALI BORA ENDERER

A thesis submitted to

The University of Birmingham

For the degree of

MA BY RESEARCH

Department of Philosophy

School of Philosophy, Theology and Religion

The University of Birmingham

April 2018

ABSTRACT

This thesis examines A. N. Whitehead and B. Russell's ramified theory of types. It consists of three parts. The first part (pp. 8–22) considers the Poincaré-Russell notion of predicativity, as well as the rationale for eliminating impredicative definitions from mathematics. It outlines the construction of the ramified hierarchy; and finally discusses the shortcomings of the ramified analysis for the formalization of the most basic parts of classical analysis.

The second part (pp. 23–33) considers two versions of the Russell antinomy. The greater portion deals with the proposition-theoretic version of the Russell antinomy, a solution to which is offered based on the construction of a hierarchy of propositions. Alternative ways out by relaxing the criterion for identity of equivalent propositions are also mentioned.

The third part (pp. 34–44) attempts to explain the requirement of predicativity as a by-product of the substitution interpretation of the second-order quantifiers; it is suggested that this explanation coheres well with a predicative conception of sets, according to which set-existence is parasitic upon identity conditions for predicatively definable sets.

ACKNOWLEDGEMENTS

I wish to express my deepest thanks to Dr Nicholas Jones of the University of Birmingham, and to Professor Ian Rumfitt of the University of Oxford, and to Professor Timothy McCarthy of the University of Illinois at Urbana-Champaign, and to Dr Sidney Felder, for valuable suggestions. I also wish to thank the examiners for some helpful revisions.

I have profited from attendance at Professor Rumfitt's Philosophy of Mathematics seminar during the Spring 2016 at the University of Birmingham, and at Professor McCarthy's Philosophy of Mathematics seminar during the Spring 2017 at the University of Illinois.

INTRODUCTION

One of Bertrand Russell's key contributions to the mathematical logic was to introduce the notions of *type* and *order* so as to provide the frame for a logical reconstruction mathematics. In what sense do the logical antinomies suggest the need for a theory of types and orders? This is the question that we shall seek to partly clarify in the present thesis.

A preliminary section (pp. 3–7) is included for the purpose of recalling the relevant notions of set theory such as 1-1 correspondence and cardinality. We also recall the proof of the diagonal lemma and Cantor's theorem, which will be relevant especially in the second part.

The first part (pp. 8–22) begins by considering the aftermath of the logical antinomies. The larger portion of this part is devoted to understanding the Poincaré-Russell notion of predicativity. It is notoriously difficult to make fully exact the philosophical notion of predicativity. We describe some prominent features of impredicative definitions, and illustrate them by classic examples. Next, we discuss the rationale for eliminating impredicative definitions from mathematics by referring to the Poincaré-Russell debate. Having discussed Russell's essential ideas leading to the ramified theory of types, we next outline the construction of the ramified hierarchy. Our attention will be directed towards foundational problems surrounding the interpretation of the ramified type structure rather than the formalization of the syntax of the ramified type theory. The first part ends with a discussion of the shortcomings of the ramified analysis for the formalization of the basic parts of classical analysis.

The second part (pp. 23–33) considers two versions of the Russell antinomy. The greater portion deals with the proposition-theoretic version of the Russell antinomy, a solution to which is offered based on the construction of a hierarchy of propositions. Alternative ways out by relaxing the criterion for identity of equivalent propositions are also mentioned.

The third part (pp. 34–44) attempts to justify the vicious-circle principle as a natural by-product of the semantic interpretation of the ramified second-order logic. The key idea to achieve the goal of coming up with an adequate interpretation of the first layer of ramified hierarchy is to switch to a non-standard way of stating the truth conditions for the second-order quantifiers in terms of substitution. The substitution interpretation of the quantifiers provides a natural way to explain the predicative restrictions brought by the vicious-circle principle. The explanation is that an assignment of truth conditions to sentences of a set-theoretic language that express propositions about sets of numbers presuppose a prior assignment of truth conditions to sentences of a number-theoretic language.

Russell has been the major source of influence to me throughout the time I have been working on this thesis. Reading his papers prior to the writing of the *Principia Mathematica* had a transformative impact on me, and led me to see better the intricacies involved in the way of finding an adequate solution to the antinomies (see RUSSELL [1956, 1973]). I do not claim to have understood the big picture yet, but I am certainly inclined to think that Russell's solution was essentially correct. This might have caused a potentially undesirable feature of this thesis in that our discussion is not wholly impartial. The fact is that Russell's philosophy of mathematics and logic has many aspects, which do sometimes conflict with each other. In this thesis, the constructive and critical side of his philosophy is emphasized over and above his logicism, at the risk of misrepresenting some of his original views in certain ways. We also sometimes suggest certain departures from his philosophy of mathematics, but his philosophy of logic is unique and absolutely impeccable. So I wish to dedicate this thesis to Russell.

PRELIMINARIES

In this preliminary section, we recall the very basics of naive set theory to facilitate our discussion in the subsequent sections. For our purposes, the important notions of naive set theory are 1-1 correspondence and cardinal number. We also recall the proof of Cantor's theorem.

As usual, by a *set* (or *class*) we mean a collection of distinct objects whose ordering, multiplicity, and mode of specification are not considered. The objects in a set are its *elements* (or *members*). The *membership relation* ($\in$) between sets and their elements is a primitive notion.

The *set-inclusion relation* ($\subseteq$) is defined in terms of the set-membership primitive: a set M is said to be a *subset* of some other set N just when each element of M is also an element of N. The set of all subsets of a set M is denoted by $S(M)$. That is, $S(M) = \{X \mid X \subseteq M\}$.

A set may be introduced either by an *enumeration* of its elements or by a *specification* as the collection of exactly those objects that fulfill a certain condition or definition. In general, a set whose elements are enumerable is also specifiable by means of a condition. The converse does not hold, since the first method of definition is in general not applicable to infinite sets. Accordingly, we think of any set as specifiable by a certain condition.

A *specification* of a set M by a condition or definition $A(x)$ is a statement to the effect that M is a set all and only objects such that $A(x)$ holds. So a specification of a set M is of the form:

$$(\forall x)[x \in M \leftrightarrow A(x)]$$

A set M is said to be *definable* by a first-order formula $A(x)$ if and only if $A(x)$ holds for all and only the elements of M. A set is called *arithmetical* if it is definable by a formula in the language of first-order number theory.

The basic principles of naive set theory are two in number.

The *axiom of comprehension* (or *specification*) states that any given condition specifies a set comprising all of those objects satisfying that condition. It is expedient to write it as a schema which yields by substitution an infinity of statements of the form:

$$(1) \qquad (\exists y)(\forall x)[x \in y \leftrightarrow A(x)].$$

We read (1) as "there is a set Y comprising all and only objects x such that $A(x)$ holds". Note that naive set theory, there no operative principle to prevent sets being elements of other sets.

The *axiom of extensionality* states that we can describe a set completely by identifying its elements: for any given sets M and N (not necessarily distinct), $M = N$ holds if and only if M comprise all and only the elements of N. In symbols, for any given sets M, N,

$$(2) \qquad M = N \leftrightarrow (\forall x)[x \in M \leftrightarrow x \in N].$$

In consequence of (2), for any condition $A(x)$, the set of all x such that $A(x)$ holds is unique, called the *extension* of $A(x)$ and denoted by $\{x \mid A(x)\}$.

We sometimes abbreviate statements of the form $(\forall x)[x \in M \to A(x)]$ by statements of the form $\forall m \in M : A(m)$. Similarly, we abbreviate $(\exists x)[x \in M \wedge A(x)]$ by $\exists m \in M : A(m)$.

Intuitively, a pairing $m \longleftrightarrow n$ is said to establish a 1-1 correspondence from a set M onto some other set N when to each element of M corresponds exactly one element of N, and each element of N is the correspondent of exactly one element of M. One writes $M \sim N$ to express the fact that the elements of M can be put into 1-1 correspondence with those of N.

Definition. $M \sim N$ holds if and only if there is a relation R that fulfills the conditions:

(3) $\forall m \in M : \forall n, n' \in N : R(m, n) \wedge R(m, n') \rightarrow n = n'$

(4) $\forall n \in N : \forall m, m' \in M : R(m, n) \wedge R(m', n) \rightarrow m = m'$

(5) $\forall m \in M : \exists n \in N : R(m, n)$

(6) $\forall n \in N : \exists m \in M : R(m, n)$

Let us recall Cantor's Diagonal Lemma, which is crucial for future reference.

Cantor's Diagonal Lemma. Let M be a set. If Σ is a collection of subsets of M such that $M \sim \Sigma$, there is at least one subset of M that fails to be an element of Σ.[1]

Proof. There are (at least) two ways to prove this lemma.

1. Let Σ be a collection of subsets of M.
2. Let $m \longleftrightarrow S_m$ be a pairing of each element m of M with exactly one subset S_m of M.
3. Let $\Sigma = \{S_m \mid m \in M\}$ so that $m \longleftrightarrow S_m$ yields a 1-1 correspondence from M onto Σ.
4. Define $D = \{m \in M \mid m \notin S_m\}$.
5. If $m \in M$ such that $S_m = D$ holds, then $m \in D$ holds if and only if $m \notin S_m = D$ holds.
6. So, by *reductio ad absurdum*, $D = S_m$ fails to hold for each $m \in M$.
7. It follows that D fails to be an element of Σ.

---

1. Suppose that Σ is a collection of subsets of M such that $M \sim \Sigma$.
2. Let $R : M \longleftrightarrow \Sigma$ be the relation defining the 1-1 correspondence.
3. Define: $D = \{m \in M \mid \exists S \in \Sigma : R(S, m) \wedge m \notin S\}$
4. Observe that D is a subset of M by construction.
5. So, if $D \notin \Sigma$ holds, we are done.

6. Otherwise, without loss of generality, let $m_0 \in M$ such that $R(D, m_0)$ holds.
7. If $m_0 \notin D$ holds, $m_0 \in D$ holds. So, by *reductio*, $m_0 \notin D$ fails to hold.
8. If $m_0 \in D$ holds, let $S_0 \in \Sigma$ such that both $R(S_0, m_0)$ and $m_0 \notin S_0$ hold.
9. It follows immediately that $S_0 \neq D$ holds.
10. But, R is 1-1 —in particular R satisfies the condition (4); whence $D \notin \Sigma$ holds.
11. So $m_0 \in D$ reduces to a contradiction.
12. So there is no such m_0, whence D fails to be an element of Σ.

Definition. A set M is said to have *cardinality n* if and only if the elements of M can be put into 1-1 correspondence with the natural numbers upto n exclusive. In symbols,

$$|M| = n \text{ if and only if } M \sim \{0, 1, \ldots, n\text{-}1\},$$

where $|M|$ denotes the cardinality of M. In general, a set M *has cardinality equal* to a set N if and only if $M \sim N$. A set M is *finite* if and only if $|M| = n$ for some natural number n; otherwise M is *infinite*. A set M is *countably infinite* if the elements of M can be put into a 1-1 correspondence with the set of all natural numbers.

Given two sets M and N, there are four logically possible circumstances:

(7) $S \sim N$ for some $S \subseteq M$, but $M \sim T$ for no $T \subseteq N$,

(8) $S \sim N$ for no $S \subseteq M$, but $M \sim T$ for some $T \subseteq N$,

(9) $S \sim N$ for some $S \subseteq M$, and $M \sim T$ for some $T \subseteq N$,

(10) $S \sim N$ for no $S \subseteq M$, and $M \sim T$ for no $T \subseteq N$.

If (7) holds, we say that M has cardinality greater than N and write $|M| > |N|$. If (8) holds, we say that M has cardinality less than N, and write $|M| < |N|$. According to *Schröder-Bernstein theorem* in set theory, $M \sim N$ whenever M and N satisfy the condition (9). If (10) holds, we say that the cardinalities of M and N are *incomparable*. However, in set theory, the case (10) is excluded, so that any two cardinal numbers are comparable. In fact, the statement that any two cardinals are comparable turns out to be equivalent to Zermelo's axiom of choice.

We say: M has *cardinality less than or equal to N*, and write $|M| \leq |N|$, to mean that either $|M| < |N|$ or $M \sim N$ holds. Similarly, we say: *M has cardinality greater than or equal to N*, and write $|M| \geq |N|$ to mean that either $|M| > |N|$ or $M \sim N$ holds. In consequence, $M \subseteq N$ implies $|M| \leq |N|$. The converse implication does not hold, of course.

Cantor's Theorem. For any set M, $|M| < |S(M)|$.

Proof. Clearly, $|M| \leq |S(M)|$ holds, since we can put the elements of M into 1-1 correspondence with the unit subsets of M by pairing each element m of M with that unit subset of M whose sole element is m. That is, $M \sim \{\{m\} \mid m \in M\} \subseteq S(M)$. Suppose, for the sake of a contradiction, that $M \sim S(M)$. Observe that $S(M)$ is a collection of subsets of M. So, by the satisfies the conditions for the Lemma. So by the Lemma, there must be a subset S of M such that $S \notin S(M)$. This is impossible, since by definition $S(M)$ comprises all subsets of M.

Cantor's Paradox. Where V is the set of all sets, $|V| < |S(V)|$ both holds and does not hold.

Proof. Any subset of V is a set; so $S(V) \subseteq V$, whence $|S(V)| \leq |V|$, so $|S(V)| > |V|$ does not hold. In conjunction with Cantor's theorem this yields the contradiction.

---

[1] The lemma is stated almost the same way as in S. Kleene (see KLEENE [1952, §5 Lemma A]). Also compare our proofs with those in BOOLOS [1997] and BELL [2004], wherein similar arguments are used to directly prove Cantor's theorem itself by taking Σ be the set of all subsets of M. My reason for not doing so are based on concerns about finding a neutral formulation of the lemma that does not presuppose the full powerset of an arbitrary set.

I

The earliest marks of the fact that the naive conception of sets is contentious were the logical antinomies directly stemming from Cantor's transfinite arithmetic. The most prominent ones are: (a) Cantor's paradox of the greatest cardinal (1895), (b) Burali-Forti's paradox of the greatest ordinal (1897), and (c) Russell's paradox (1902). These are called *logical antinomies* because we have in each case a logically flawless argument that is used to prove a contradiction from basic principles.

Logicians were subsequently able to discover a considerable number of antinomies that involved no reference to the fundamental concepts of set theory. Among these the most prominent ones are: (d) Richard's paradox (1905), (e) König's paradox of the least indefinable ordinal (1905), (f) Berry's paradox of the least integer not nameable in fewer than nineteen syllables (1906), (g) the Nelson-Grelling paradox of the expression 'heterological' (1908) and, of course, (h) the paradox of the statement 'this statement is false'. Antinomies of this kind are called *semantical antinomies* because they deal with semantical notions such as definability and truth. RAMSEY [1926] and HILBERT-ACKERMANN [1928] suggested that these have to be distinguished from the antinomies of logic and set theory, because they thought that the derivation of semantical antinomies somehow depend upon extra-logical features, such as linguistic features, of the relevant paradoxical statements. Nonetheless, the so called semantical antinomies share important characteristics in common with the logical antinomies.

In fact, it does not seem far-fetched to suspect that the antinomies of logic and set theory also have to do with problems pertaining to the use of symbolism. We saw that underlying the naive set theory are two principles that allow the passage from any condition to infer the *existence* and *uniqueness* of the set of objects satisfying that condition. So the relation of conditions to sets is many-one, which induces an operation on the domain semantically associated with a class of expressions. Seen in this way, (1) allows us to take any well-formed formula and to assert existence of an object named by the corresponding closed expression that purports to involve reference to a set; and (2) guarantees that no other object is named by the same expression. So one way to spell out the critique of naive set theory is the following: paradoxes result from processes of concept-formation terminating with a description that refers to nothing (like Frege's 'the least rapidly converging series'). The expression 'the set of all sets' refers to nothing. The logical status of statements containing expressions of this kind would of course depend on one's theory of referential failure. At any rate, no paradox is forthcoming unless we adopt the untenable assumption that any condition $A(x)$ can be used to generate a substantive existence claim "the set of all x such that $A(x)$ holds exists". So, each of the paradoxes of set theory reduces the relevant instance of the comprehension axiom into a definite contradiction. The inconsistency of naive set theory is due to the fact that the comprehension principle is taken without any restriction on its range of application. As such, the procedure is illegitimate as a means of definition.

But this is not the whole story. The problem disclosed by set-theoretic paradoxes is a very deep one in the foundations of mathematics that directly concerns the conceptual constructions underlying *set-existence* claims made in actual mathematical practice. As E. Beth remarked, "the effect of a contradiction arising in such a context is fatal, because the inconsistency of a proposed basis for pure mathematics entails the inconsistency of the resulting deductive development of pure mathematics itself." (BETH [1966: 307]) Even after the situation caused by the discovery of the antinomies is diagnosed in some way or another, there remains a serious problem of establishing a consistent basis upon which mathematics could be developed. The minimal requirement is to somehow restrict the range of application of the comprehension principle in set theory. This task was achieved to a certain extent by the axiomatization of set theory by ZERMELO [1908] by supplanting the comprehension principle with the separation axiom [*Aussonderung*] together with substantive set-existence axioms. The separation axiom allows, for any given set M, the formation of a subset of those elements of M that satisfy a certain *definite* condition or property φ, i.e. a condition or property that for any given argument x yields a determinately true or false statement or proposition $\varphi(x)$.

A natural attempt to avoid the antinomies to scrutinize the construction inducing this state of affairs to mark out what is characteristic of a whole variety of pathological cases. This is not only supposed to provide an extensionally correct way of circumscribing what predicates that have to be rejected as failing to determine an extension, but it is also supposed to deliver an explanation of *why* they fail to do so. This second aspect of the problem is very important, as it is the *only* way one can attain the certainty that mathematics is free of contradictions.

From the standpoint of predicative foundations of mathematics, the main objectionable feature of the logical antinomies was a peculiar kind of circularity implicit in the definitions that form the bases for the arguments. The sort of circular definitions in question are called *impredicative* definitions. In general, a definition of an object (individual, set, property, proposition, etc.) is said to be impredicative if the *definiens* involves (explicit or implicit) reference to the object defined. This is the general form of circularity that one encounters in mathematics. The easiest way to grasp this notion is to look at examples that the predicativist rejects.

The Leibniz identity relation is defined as follows: x is identical to y if and only if x possesses *all* (and only) the properties of y. One property that y is certain to possess is that of being identical to y. Thus, as we need to run through the properties of y so as to ascertain whether x possesses each of them, we would have to ascertain whether x has the property of being identical to y. Since the procedure for determining whether the identity relation holds between two objects leads us back to the question with which we started off, the definition of Leibniz identity is impredicative.

The Dedekind-Frege-Russell definition of the natural numbers is the following: n is a natural number iff n possesses every inductive property, where a property F is *inductive* iff F belongs to 0, and F belongs to $n+1$ whenever F belongs to n. This definition is circular because the property of being a natural number is an inductive property. Thus, e.g. to ascertain that the number 3 is a natural number we need to see whether 3 has every inductive property; in particular, whether 3 has the property of being a natural number. (cf. CARNAP [1931]) So the procedure for ascertaining whether an object is a number leads us to a circle. Similarly, the set of all naturals ω is defined the intersection of *all* sets containing $\varnothing$ which are closed with respect to the successor operation $n \mapsto n \cup \{n\}$. So to ascertain whether something is an element of ω we would have to ascertain whether it is an element of each set in a certain collection to which ω belongs in particular. So the definition of the natural numbers ω is impredicative.

Compare our description of impredicativity with those given by S. Kleene and A. Church:

> When a set M and a particular a particular object m are so defined that on the one hand m is a member of M, and on the other hand the definition of m depends on M, we say that the procedure (or the definition of m, or the definition of M) is *impredicative*. Similarly, when a property P is possessed by an object m whose definition depends on P (here M is the set of the objects which possess the property P). An impredicative definition is circular, at least on its face, as what is defined participates in its own definition. (KLEENE [1952: 42])

> To avoid impredicativity, the essential restriction is that quantification over any domain (type) must not be allowed to add new members to the domain, as it is held that adding new

> members changes the meaning of quantification over the domain in such a way that a vicious circle results." (CHURCH [1976: 747])

In this thesis, we shall be concerned with one very basic predicative requirement: the closed singular term that purports to refer to a set of objects must *never* contain a bound set-variable. Thus, the definition of a set of objects $\{x \mid A(x)\}$ is impredicative if A contains set-quantification and $\{x \mid A(x)\}$ belongs to the domain of quantification.

So predicative mathematics emphasize semantic notions such as *definability* and places restrictions on formulas so as to distinguish those that constitute defining conditions for sets. It is crucial on this view the elements of a collection can be generated in a non-circular way; otherwise the purported definition is rejected as impredicative. An instance of the comprehension schema expresses nonsense unless the formula used to instantiate it satisfies these predicative requirements. Since instances of the comprehension axiom are all statements of set-existence, this means that, on the predicativist approach, set-existence is logically parasitic upon identity conditions for predicatively definable sets. The vicious-circle principle was put forward by Russell to eliminate impredicative definitions from mathematics.

Some of the key ideas involved in the foundations of predicative mathematics originated in the famous Poincaré-Russell debate that took place at the turn of the 20th century; reading their exchange helps to understand the reasons why impredicative definitions are worrisome. For one thing, Poincaré pointed out that in the case of an impredicative specification the domain of quantification is constantly expanded, whence such a classification is "disordered by the introduction of new elements". A classification of this kind is not rigid, while Poincaré thought that "the condition necessary for the rules of this logic to be valid... is that the classification which is adopted be immutable." (POINCARÉ [1913: 45])[1]

Russell held views similar to Poincaré although he held that "the contradictions have no essential reference to infinity" (RUSSELL [1906: 197]) Nevertheless, Russell agreed that the paradoxes result from the fact that it is illegitimate to suppose that infinite sets form a *totality*, as well as the observation that the solution to the paradoxes require the elimination of impredicative definitions. In a series of papers, Russell gave a definitive analysis of the antinomies. The most important among these papers is RUSSELL [1905]. Below are excerpted what I take to be the key points of Russell's analysis of the logical antinomies in this paper:

> a propositional function $\varphi!x$ may be perfectly definite, in the sense that, for every value of x, $\varphi!x$ is determinably true or determinably false, while yet the values of x for which $\varphi!x$ is true do not form a class. [1905: 137]

> The refutable assumption as to the nature of classes and relations is only this: that a class is always uniquely determined by a *norm* or property containing one variable, and that two norms which are not *equivalent* (i.e. such that, for any value of the variable, both are true or both false) do not determine the same class, with a similar assumption as regards relations.

> We have thus reached the conclusion that some norms (if not all) do not define classes. Norms (containing one variable) which do not define classes I propose to call *non-predicative*; those which do define classes I shall call *predicative*. [1905: 141]

> …the contradictions result from the fact that, according to the current logical assumptions, there are what we may call self-reproductive processes and classes. That is, there are some properties such that, given any class of terms all having such a property, we can always define a new term also having the property in question. Hence we can never collect all of the terms having the said property into a whole; because, whenever we hope we have them all, the collection which we have immediately proceeds to generate a new term also having the said property. [1905: 144]

The idea of the introduction of “new” objects (terms) needs clarification. For example, H. Curry said of this type of argument: “Ebensogut aber ist dieses Argument: man könnte nichts über alle Apfelsinen behaupten, weil jedes Jahr neue Apfelsinen ershaffen werden, und es also keine bestimmte Gesamtheit von Apfelsinen gäbe u.s.w.” (CURRY [1930: 517]). This amusing interpretation is based on a misunderstanding. According to Russell’s statement quoted above, the question whether the newly introduced object x has the said property F depends for its answer on a prior determination of a set of objects M possessing that property. Under these conditions, if x possesses the property F and yet fails to be an element of M, this shows that M fails to exhaust all objects with the property F. It hardly needs noting that the newly created orange does not depend for its identity on the set of previously existing oranges; on the contrary, any particular orange —and more generally any concrete spatiotemporal object— comes with its own principle of individuation. So the Poincaré-Russell notion of predicativity does not prevent the possibility of making generalized claims about concrete objects.

The above analysis immediately yields the *vicious-circle principle*, of which Russell gave several formulations. The earliest formulation runs as follows: “Whatever involves an apparent variable must not be among the possible values of that variable” [1906: 198] This was later superseded by the classic formulation which appeared in the *Principia Mathematica*:

> Whatever involves *all* of a collection must not be one of the collection”; or, conversely: “If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total. [1925: 37]

The adoption of the vicious-circle principle was criticized in GÖDEL [1944] who distinguished three formulations of the vicious-circle principle, and pointed out that they are not equivalent, provided that one grants that sets exist independently of our definitions. It seems clear, however, that these ontological assumptions about set-existence were explicitly rejected by Russell, so that it is more appropriate to say that Gödel's remark merely discloses one of the assumptions on which the justification of the vicious-circle principle rests.[2]

In the third part of the thesis, we adopt as a basic assumption that set-existence is logically parasitic upon the identity conditions for predicatively definable sets, and undertake the task of justifying the vicious-circle principle on this basis. Admittedly, this is a controversial thesis in the foundations of mathematics, but so are many theses of the classical (cantorist) viewpoint which are accepted by the majority of mathematicians. The issue of adjudicating between the classical and predicative foundations is beyond the scope of this thesis.

To sum up, the possibility of collecting objects of some type or another into a set requires that the classification or definition that specifies the collection be rigid; this in turn depends on whether the domain has a precisely delimited boundary. The fact that collections of sets, cardinals, ordinals can be iterated indefinitely show that it is impossible to conceive of totalities of them, or to conceive of sets of them as formed by selecting items from a previously given totality by reference to that totality. The range of correct modes of specification is delimited precisely by the vicious-circle principle.

The preceding considerations furnish one possible way of understanding why Russell was led to establish a new hierarchy towards developing, under the guidance of Whitehead, the system of their monumental *Principia Mathematica* (Cambridge, 1910).

The introduction of types in a logical system amounts to a classification of all its terms in such a way that types of composite terms are determined by types of other terms which are always less complex than the ones in question. The motivation for introducing type theories originally derives from the need to provide a secure foundation of classical mathematics. The fact that it is possible to codify the naive set-theoretic principles in a *type-free* logical system, which incorporate the full comprehension principle and allow unrestricted quantification, suggested the idea that higher-order logic is really in the same boat as the naive set theory. So it was an imperative to restrict the comprehension principle in higher-order logic.

The ramified theory of types and orders has two components. The first consists of a stratification of functions resp. sets into *types*, which restricts the range of admissible arguments reps. elements for the functions and sets in question. The second consists of a stratification of terms of the same type into *orders* according to their internal structure, which further restricts the range of admissible substituends for the terms in question. Stratification is a partition of the relevant hierarchy into mutually disjoint domains, which is brought in order to forestall the paradoxes. So, in general, there are two related hierarchical structures; the ramified theory combines them to yield a system of a considerable amount of complexity. We begin with the simple type hierarchy because it is simpler.

Types in our sense are also called *levels*, in view of the fact that functions of one variable or several variables lie at the same level irrespective of the number of arguments they take, whereas they are said to be of distinct types. The general construction is based on the following recursion: where $\tau_1, \ldots, \tau_n$ are types, $(\tau_1, \ldots, \tau_n)$ is the type of functions of n variables taking arguments of types $\tau_1, \ldots, \tau_n$ respectively. Having said this, we shall restrict ourselves to functions of one variable.[3] With this simplification, the intuitive picture of the hierarchy of types can be described as follows: given nonempty domain D comprising objects of a certain specific sort, the members of D, called *individuals*, are of level 0 and lie at the bottom of the hierarchy; properties of individuals resp. sets of individuals are of level 1 and lie at the first layer of the hierarchy; properties of properties of individuals resp. sets of sets of individuals are said to be of level 2 and lie at the second level; and so on up. This generates an ever increasing hierarchy without a top.

The core idea of simple type-theoretic responses to the set-theoretic antinomies is to make a type distinction between functions and their arguments in order to forestall the possibility of a function having the same level as its arguments; a corresponding distinction between sets and their elements forestalls the possibility of a set having the same level as its elements. The type distinctions are made the manner indicated in the previous paragraph. In set terms, the language of the simple type theory is characterized by the following restriction: where x_i^n and x_j^m are variables ranging over objects of levels n and m respectively, the formula $[x_i^n \in x_j^m]$ is well-formed if and only if $m = n+1$. This avoids the Russell paradox in a very obvious way, since the formula $[x \in x]$ is not well-formed in the language of simple type theory.

Even though this construction suffices to forestall the set-theoretic antinomies, it does not preclude the vicious circles that occur within each level, which we mentioned above. To avoid them, Russell implemented, for each type, a further stratification at the interior of the domain of objects of the same type. The basic idea is to make an order distinction between an object defined by a condition involving quantification and the objects forming the domain of quantification, so as to forestall the possibility that the *definiendum* itself does not belong to the domain of quantification associated with the *definiens*. Stratification of objects of the same type into orders is also called *ramification*. It is also common to use the word ‘degree’ in lieu of ‘order’ (as in, e.g., FEFERMAN [1964]). In order not to clash with the terminology in the foundations of mathematics, we will use the words ‘level’ and ‘degree’ to mean (simple) types and orders. Let us restrict attention to items of level 1, that is, properties or sets of individuals. Then, the intuitive picture is as follows: objects of level 1 and degree 0 are those which can be defined by a condition that does not involve quantification over items of level 1; objects of level 1 and degree 1 are those which can be defined by a condition that involves at most quantification over objects of level 1 and degree 0, and so forth. This doubly stratified structure is what is referred to as the *ramified hierarchy*. In consequence, one gets rid of an unstratified higher-order domain and the idea of "arbitrary property", i.e. a property of an arbitrary order.

Finally, let us adopt the following notation (similar to Chihara's) and write T_k^n to denote the *ramified type* of objects of level k and degree n. Now, we can settle the question of what the level and degree of any given object are simply by inspecting the structure of its defining condition. Let us take a simple example from number theory. Intuitively, it is easy to see that both the set of *prime numbers* and the set of *natural* (or *finite*) *numbers* should be of level 1, given the naturals as forming the domain of individuals. What is less obvious is that they are of different degrees. So, let us look at their defining conditions to see whether this is indeed the case: where x, y, z ranges over individuals, the set of primes is defined as follows:

$$(\forall x)[x \in P \leftrightarrow 1 < x \wedge (\forall y)((\exists z)[z \times y = x] \rightarrow (y = x \vee y = 1))].$$

Since the defining condition of P involves at most quantification over individuals, the set P is of ramified type T_1^0. By contrast, the naturals of degree 1 are now defined as objects that belong to every set of type T_1^0 that is closed with respect to the successor relation. Where Y^0 ranges over arithmetically definable sets, we write:

$$(\forall x)[x \in N^1 \leftrightarrow (\forall Y^0)[0 \in Y^0 \wedge (\forall z)[z \in Y^0 \rightarrow (z{+}1) \in Y^0] \rightarrow x \in Y^0].$$

As the defining condition of N^1 involves at most quantification over sets of type T_1^0, the set N^1 itself must be of type of type T_1^1. So, N^1 turns out be of the *same level but higher degree* than all sets of degree 0. In consequence, the definition of naturals of degree 1 is now predicative.

To describe the more general construction, start again with the natural numbers as forming the bottom of the ramified hierarchy. We call a condition *arithmetical* if and only if it involves at most quantification over the naturals. Then, sets definable by an arithmetical condition, or *arithmetically definable sets*, are of type T_1^0. Sets definable by a condition involving quantifiers ranging over at most arithmetically definable sets will be of $T_1{}^1$, and so forth. We proceed in the same manner to get ever higher degrees at the interior of level 1. Now, define the *degree of a formula* A to be 0 if A expresses an arithmetical condition, and otherwise $m+1$ where m is the maximum of all positive integers such that the set-variable Y^m appears bound in A. Then, the *predicative comprehension schema* reads:

(11) $$(\exists Y^n)(\forall x)[x \in Y^n \leftrightarrow A(x)].$$

where A is a formula of degree n and the exponent indicates the degree of sets that the quantifier ranges over. (See FEFERMAN [1964] for further details.) Thus we come to what is called the *ramified second-order logic*. The construction can be iterated into the transfinite in the presence of suitable ordinals taken as degrees. To get the *ramified third-order logic*, one adds (ramified) types for properties taking as arguments objects of level 1, or sets having objects of level 1 as elements, and so on. Note, however, that we are concerned only with the ramified second-order logic in this thesis.

This concludes our account of the Russell's ramified theory of types. It is sketchy but I hope adequate for the following discussion. At this point, it is almost obligatory to mention the difficulties in the way of developing mathematics in the ramified theory of types. Unfortunately, the most satisfactory attempts to develop classical mathematics in this manner turned out be too restrictive. In particular, it turns out that it is hopeless to represent real analysis in such a predicatively inspired framework.

For the sake of completeness, let us consider what happens in the *ramified analysis*. The *least upper bound principle* says: every set of real numbers that has an upper bound has a least upper bound, which is minimal with respect to all upper bounds. Suppose that we model the real numbers by Dedekind cuts. One can establish a one-one correspondence between

Dedekind cuts and sets of natural numbers, so real numbers will occupy the level 1. However, the definition of the least upper bound of a bounded set of reals is impredicative: the least upper bound of a bounded set of reals M is that real number α such that any real number that is less than α (under the natural ordering of the reals) is not an upper bound of M. In symbols:

$$(\forall x)[x < \alpha \leftrightarrow \exists m \in M : x < m]. \tag{12}$$

If this is to be avoided, the associated domain of real numbers will have to be partitioned into degrees, so there will be an "infinituplication" real numbers; consequently, one is not permitted to quantify over arbitrary real numbers, so there will have to be infinitely many least upper bound principles. However, Feferman notes, "Russell realized this would be ludicrous for the practicing analysts" (see FEFERMAN [1964: 6]). So the ramified analysis is hopeless for the formalization of even the most basic parts of classical analysis. In fact, it is highly suspect if it is sufficient for the formalization of Peano arithmetic. As shown by J. Myhill, the induction principle violates the order distinctions, and the ramified hierarchy of natural numbers of finite degrees *never* collapses at any degree whatsoever (see Myhill [1974]). This result partly answers the question of whether arithmetic is reducible to logic in the negative.[4]

There are (at least) two sorts of attitudes that we may adopt in response to this situation. We might think of as the complications introduced by the ramified theory as generating only a *pragmatic* problem, because one does not have anything corresponding to it in actual development of classical mathematics. Perhaps, the ramified theory is a philosophical reconstruction that is not intended for everyday mathematicians. One way to make the theory more feasible for the practicing analyst is to follow Whitehead and Russell and postulate strong axioms such as the reducibility principle, which states that, for any property resp. class of any level and any degree there is at least one property resp. class of the same level and degree 0. The reducibility principle is symbolized as:

(13) $$(\forall X^n)(\exists Y^0)(\forall z)[z \in X^n \leftrightarrow z \in Y^0].$$

This effects the collapse of the whole hierarchy of degrees within each level. But, it should be pointed out that introduction of this axiom lacks sufficient motivation: if we would end up developing analysis and set theory in a system incorporating the reducibility principle, what was the point of introducing the ramified hierarchy in the first place? This is one of the major criticisms raised by L. Chwistek and F. P. Ramsey against the *Principia*.[5] On the other hand, many people noted that one of the sources of motivation for ramification was to avoid the semantical antinomies mentioned at the beginning, as well as paradoxes related to intensional contexts, such as the Epimenides paradox. What difference is there between the simple type theory and the ramified type theory + reducibility from the perspective of dealing with the semantical antinomies? The above mentioned critics thought that this would reinstate the semantical antinomies within each type. To evaluate this criticism would require a lengthy analysis of the semantical antinomies, which is beyond the scope of this thesis. But, J. Myhill persuasively argued that this criticism is unfounded, because the equivalence of the simple theory and the ramified theory + reducibility does not entail that their extension by the addition of intensional and semantical notions such as *assertion* and *denotation* will remain equivalent, which in fact is not the case (MYHILL [1979]; see also URQUHART [1988, 2001]). This points certainly deserves more attention in view of the fact that the *Principia* system represents the idea that logic is universally applicable.

Another sort of attitude is to stick to the requirement of predicativity, and reject the rest of classical theorems that violate these requirements. If we strictly adhere to this view, but wish to avoid the compilations resulting from the ramified hierarchy, there are alternative, and arguably more feasible ways of developing of mathematics on a predicative basis, as illustrated by the development of *predicative analysis* as inaugurated by H. WEYL (1918).[6] For us, the crucial idea here is to reverse the order of explanation concerning the nature of abstract mathematical objects. Taking the natural numbers as *given*, we can build up analysis and set theory on a predicative basis. Instead of explaining the natural numbers by a set-theoretic reduction, this move allows us to have a theory of sets of natural numbers, so to speak. Formally, the key idea is to restrict the range of acceptable instances of the comprehension axiom schema to arithmetical conditions. In view of this, one cannot help but join A. Hazen & J. Davoren in asking "What could Russell have done if he had combined *his* philosophy of logic, including certain skepticism about Reducibility, with *Brouwer's* (or Weyl's) personality and willingness to advocate trimming mathematics to eliminate the philosophically suspect bits?" [2000: 552]

---

[1] See also CHIHARA [1973] for an excellent discussion of Poincaré's views.

[2] Cf. the remarks by H. Wang and W. Goldfarb on this point (see WANG [1959], GOLDFARB [1989]).

[3] A more systematic introduction to the hierarchy of types and orders will not be needed for our discussion, so we shall not bother trying to give a formalization of a full-blown type-theoretic syntax. For a more rigorous treatment, see Church's formulations of the simple and ramified theory (CHURCH [1940, 1956, 1976, 1984]). I would also like to mention here those works that I found helpful while studying the ramified theory: CHIHARA [1973], COPI [1971], COQUAND [2015], HAZEN [1989], HAZEN & DAVOREN [2000], HODES [2015], URQUHART [2001].

[4] I say partly, because the situation depends on whether or not the system incorporates the axiom of extensionality. For a discussion of this issue, see HAZEN & DAVOREN [2000].

[5] See COPI [1971] for a discussion and references. This issue received a lot of attention during the last century, so we shall not discuss it any further.

[6] See also WANG [1959] and FEFERMAN [1964, 2005] for a general statement of this viewpoint.

II

This second part of the thesis presents two versions of the Russell antinomy, neither of which is the classic set-theoretic version. In fact we shall not consider the set-theoretic version since it is more easily manageable than the versions considered here. We tried to set out the two versions of Russell antinomy and Cantor's theorem in a way that highlights their structural similarity. The proof of Cantor's theorem breaks down in the ramified *Principia* without reducibility (see WHITEHEAD & RUSSELL [1925: xiv]). So, we hope this might corroborate the view that Cantor's theorem depends on impredicative definitions.

The greater portion of this section deals with the proposition-theoretic version of the Russell antinomy. The reason is partly because there is an ongoing controversy about whether it is a genuine antinomy at all. For example, in his correspondence with Russell, the great Frege dismissed the antinomy on the basis of his *Sinn-Bedeutung* distinction (see Frege [1980]; although see Klement [2001]). We believe it is a genuine antinomy, but we mention alternative "ways out" by relaxing the criterion for identity of equivalent propositions. We have been at pains not merely to translate into our notation some existing formulation of the antinomy, but to present it in an original way that makes the connection with the classic set-theoretic version perfectly plain. So this might lend some insight into the question whether it is a genuine antinomy or not. However, it should be stressed that the virtues of our analysis of the antinomy (if any) are not to be sought in technical aspects.[1] Rather, they consist in the fact that the analysis makes clear the significance of making order-distinctions among propositions similar to type-distinctions among properties and sets.

The two versions of the Russell antinomy that we shall consider read as follows.

Theorem 1. Where U is the set of all objects, $|U| < |S(U)|$ both holds and fails to hold.

Proof. Any subset of U is an object; so $S(U) = V \subseteq U$ holds, whence $|S(U)| = |V| \leq |U|$ holds. In conjunction with Cantor's theorem this yields the contradiction. Note that a variant proof proceeds from the fact that both $U \in V$ and $V \in U$ hold (cf. GANDY [1977]).

Theorem 2. Where H is the set of all propositions, $|H| < |S(H)|$ both holds and fails to hold.

Proof. Any set of propositions is an object; so $S(H) \subseteq U$ holds, whence $|S(H)| \leq |U|$ holds. But, there is a natural correspondence $U \longrightarrow H$ that maps any object x to the proposition that $x = x$ holds, and no two distinct objects to the same proposition. So $|U| \leq |H|$ holds; by chaining them, we get: $|S(H)| \leq |U| \leq |H|$. Once again, this contradicts Cantor's theorem.

Both paradoxes were noted in Russell's *Principles of Mathematics* [1903: §348, §500]. These "theorems" directly ensued from considerations about cardinalities. Here the contradictions are not due to an abuse of definition, but they are genuine paradoxes of which, or so we contend, there is *no* solution except to reject Cantor's theorem (cf. BEHMANN [1937]). We at once remark that rejection of impredicative definitions should not entail a rejection of Cantor's Diagonal Lemma (see the preliminary section). This is because the diagonal argument does not presuppose the *totality* of subsets of a given set.

To derive the Russell antinomy of the set of all objects (theorem 1), the easiest way is to effectively specify a 1-1 correspondence $\varphi : U \longrightarrow V = S(U)$, as follows: if $x \in U$ is a class, let $\varphi : x \mapsto x$; otherwise, let $\varphi : x \mapsto \{x\}$. Notice that to each element of U corresponds exactly one element of V, and each element of V is the correspondent of exactly one element of U. Hence, $U \sim V$ holds. Finally, observe that V satisfies the pre-requisites for Cantor's Diagonal Lemma. Hence, there must be a subset of U that fails to be an element of V. But, V is the set of *all* classes, whence the contradiction.

To derive the Russell antinomy of the set of all propositions (theorem 2), the second diagonal is more helpful. We specify an effective 1-1 correspondence $\pi : S(H) \longrightarrow H$, in the following way: if $M \subseteq H$ is nonempty, let $\pi : M \mapsto \bigwedge M$; and $\pi : \varnothing \mapsto \bigwedge H$, where for each set of propositions M, $\bigwedge M$ is the logical product (i.e. possibly infinite conjunction) of all propositions in M. Note that, for each $p \in H$, $\bigwedge\{p\} = p$. Also, we assume that for each $M, N \subseteq H$, $\bigwedge M = \bigwedge N$ implies $M = N$. We shall take up the question of whether this assumption is plausible later on. Granted these assumptions, we see that to each subset of H corresponds exactly one member of H, and any member of H is the correspondent of exactly one subset of H. Hence, $H \sim S(H)$ holds. So, once again, we see that $S(H)$, being a collection of subsets of H, satisfies the conditions for Cantor's Diagonal Lemma, so there must be a subset of H that fails to be an element of $S(H)$; but this is impossible since $S(H)$ is the set of *all* subsets of H.

In the remainder of this section, we consider how to construct a hierarchy of propositions to give the desired solution of the proposition-theoretic versions of the Russell antinomy. We follow the lead of Russell in doing so, who noted that "the vicious-circle principle is not itself the solution of vicious-circle paradoxes, but merely the result which a theory must yield if it is to afford a solution of them." (RUSSELL [1906a: 205]) Note that a closely related antinomy was rediscovered by the J. Myhill within the context of Church's intensional logics. (see CHURCH [1951], MYHILL [1958], KLEMENT [2014]).[2]

Consider propositions of the form *every member of M is so-and-so*, where M is a set of objects of some kind. If m is an element of M, there is a sense in which such a proposition is (indirectly) 'about' m among other things. For example, the proposition that every ostrich is bipedal is about this or that ostrich, and so on. What Russell's paradox of propositions suggests is that this way of construing the relation of a generalized proposition to the items in the domain of quantification yields a contradiction when extended to cases in which the domain of quantification comprise propositions; i.e. where propositions turn out be *about* other propositions, an idea that is comparable to the idea of sets being elements *of* others sets. Of course, this is so unless otherwise the type of propositions itself is stratified into orders, and the domain of quantification is restricted appropriately. Here, then, is the naive idea which we will have to reject: we can explicate the idea that a proposition p is about some other proposition q, as follows: where p and q are propositions, p is *about* q if and only if, for some set S, p is the logical product of S and q is a member of S.

Let H be the set of all propositions; write $[p \ @ \ q]$ to mean p is about q. We define,

$$[p \ @ \ q] \ \equiv \ \exists S \subseteq H : p = \bigwedge S \wedge q \in S.$$

This idea could be formalized in other ways, say, by using quantification over the type of propositions instead of using the infinitary conjunction operation. The crucial point is that the definition appeals to impredicative quantification over sets of propositions. But, the common moves that allow to resolve the set-theoretic paradox do not seem equally helpful in this context. For example, it turns out that even higher-order *plural* quantification leads to some version of the antinomy; see UZQUIANO [2015] for a discussion and further references.

Now, precisely the expressive power that allowed us to define the relation @ leads to a contradiction. To see this, divide propositions into two mutually disjoint sets according as they are about themselves or not. Generalized propositions to the effect that all propositions fulfill such-and-such a condition will be of two kinds, according as those that fulfill the condition which they predicate of all propositions, and those that fail to fulfill. One sees that "ordinary" sets of propositions which can be specified by a description such as 'the set of all propositions of the form ...' fail to contain their own products as elements, so the corresponding logical products will fail to be about themselves. But, the proposition that all propositions are true would be the product of a set of which it is a member, so it would be about itself.

Let R be the set of all and only those propositions which are *not* about themselves, that is:

$$R = \{p \in H \mid \neg[p \ @ \ p]\}$$

Let $r = \bigwedge R$. So, r is about all and only those propositions which are not about themselves. Paradox results when we ask whether or not $[r \in R]$ holds. If the answer is "yes", $\neg[r \ @ \ r]$ holds and hence, familiar logical rules lead to the conclusion that $[r \notin R]$ holds. If the answer is "no", then by the law of excluded middle, $[r \ @ \ r]$ must hold, whence $[r \in R]$ holds. So from each horn of the dilemma its contradictory follows. Such is the Russell paradox of the propositions which are not about themselves. The parallel with the paradox of sets which are not members of themselves should be clear.

We can resolve this antinomy without difficulty by constructing a hierarchy of propositions in the style of RUSSELL [1908]. At the bottom of this construction lie propositions which are *not* about propositions —the ones about ostriches will be among these; these are *first-order* propositions. Then, *second-order* propositions are those that are *about* (@) first-order propositions. And so on up. In general, propositions of order n+1 are those that are about propositions of order n. The orders are *mutually exclusive*, i.e. a set of propositions M may comprise only propositions of a certain determinate order k, and the conjunction $\bigwedge M$ of those propositions of order k must be a proposition of order k+1. The essential restriction imposed in this setting the following: where p_i^n and p_j^m are variables ranging over propositions of orders n and m, respectively, the formula $[p_i^n \ @ \ p_j^m]$ is well-formed if and only if $m = n+1$. In consequence, the formula $[p \ @ \ p]$ is not even well-formed. If we indicate the orders by adding indices, it appears that the condition $[p_i^n \ @ \ p_i^n]$ violates the order distinctions, since $n \neq n+1$. So, the condition defining the Russell set R is "meaningless", and the set R is indefinable.

We "proved" two versions of the Russell antinomy based on a simple comparison between two cardinalities. While they indicated the paradoxical state of affairs, they did not reveal the fact that the underlying constructions essentially appeal to impredicative definitions. When analyzed into logical terms, the Russell-Myhill antinomy is seen to make an essential appeal to an impredicative definition. The solution was to restrict the rules of definition so as to avoid the way to defining the Russell set *R*.

Now, ramification is certainly not the only way to meet the minimal restrictions to resolve the Russell-Myhill antinomy. Another way is Quine's *New Foundations* (see QUINE [1938]).[3] This route too was indicated by Russell in that marvelous 1905 paper. No doubt, there are other ways still. Recent research shows that even Russell's type-free substitutional theory proves free from the contradiction.[4] The common core of these various solutions seems to be to reject Cantor's theorem, so there is evidence to the fact that Cantor's theorem is problematic from various viewpoints. The problem, I think, lies in the illegitimate assumption of the *totality* of subsets of an arbitrarily given set in the proof of Cantor's theorem. At this point, I must leave this issue to experts to consider.

The basic observation about the behavior of the function π is that it is a *type-reducing* function. Now, the definition of the Russell set *R* is impredicative. If we countenance impredicative definitions, as in the simple type theory, *R* would be a definable set. For, consider what would happen in a type structure, where sets are stratified into types in the usual manner and propositions are admitted as forming a type at the bottom of the hierarchy —with sets of propositions at the first level, sets of sets of propositions at the second level, and so forth. Given a set of propositions M of type 1, the associated logical product $\bigwedge M$ would be of type 0, and hence it would *not* have been impossible in general for $\bigwedge M$ be a member of M. If so, then consulting to the definition of @, we see the formulas in which the symbol @ is flanked by two occurrences of one and the same propositional variable would have been well-formed. This is why the set R would have been definable in a simply typed theory, where propositions form a type in their own right. If these observations are correct, it seems that this antinomy could not be avoided in simple type theory, where propositions form a type in their own right. (However, this is not the case for the *extensional* fragment of the simple type theory where equivalent propositions are counted as identical; see below.) So, unlike in the case of the set-theoretic version of the Russell antinomy, there seems to be no violation of simple type-theory involved in the way of specifying the relevant Russell set.[5]

Similarly, in the framework of Zermelo-Fraenkel set theory, the derivation appeals only to the separation axiom. The reason is that the mode of specification of the Russell set does not violate the restriction involved in the separation axiom (see UZQUIANO [2015] for details).

What remains to be argued is whether the relation $p = \bigwedge S$ defines a 1-1 correspondence between the set of propositions and its powerset. If it does, the paradox results from a simple application of the Lemma, as we have seen. The question is whether it is plausible to assume:

$$\forall X, Y \subseteq H : \forall p \in H : \bigwedge X = \bigwedge Y \rightarrow X = Y. \tag{14}$$

The meaning of this axiom is clear: two sets of propositions X and Y are identical if the proposition that each member of X is a true proposition is identical to the proposition that each member of Y is a true proposition. In short, two sets of propositions X and Y are identical if they have the same logical product. It turns out that this assumption implicit in Russell's argument in Appendix B is not uncontentious, and at least one logician dismissed the antinomy in question because this assumption seemed implausible (see HODES [2015]).

What reasons do we have for maintaining the axiom (14)? Russell's justification for this claim primarily comes from his underlying concept of propositions as structured complexes.

Every proposition is represented as the value of a propositional function, and two values of one and the same function differ only insofar as it is applied to different arguments. Assuming (14), then, amounts to the claim that the identity conditions of conjunctions depend upon and determined by the identity conditions of their conjuncts. On this view, the identity conditions for propositions is much more fine-grained than logical equivalence.[6]

It is essential for an axiomatic derivation of the paradox in a deductive system, it is true, that the system does *not* include an axiom to the effect that equivalent propositions are in general identical, in which case the paradox disappears. So if the law $[(p \equiv q) \rightarrow (p = q)]$ is added to the system of CHURCH [1940] the paradox disappears. The axiom of identity of equivalent propositions is standardly included in simple type-theories. The system of CHURCH [1940] does not include this axiom, but the modern presentations of the theory often (explicitly or implicitly) restrict the type of propositions to the set $\{T, F\}$ so as to be able to consider its successive power-sets; —where T, F can be taken as standard propositions representing the truth-values, of which a variety of choices of representatives are available (see HENKIN [1956]). So, if we start off with a domain $\{T, F\}$, the consecutive iteration of powerset operation countably many times will never result in an infinite set. Therefore, this antinomy cannot be derived in a type theory with an axiom ensuring the identity of equivalent propositions.

So one of the possible routes to dispense with the proposition-theoretic antinomy is to identify logically equivalent propositions. Under the assumption that logically equivalent propositions are identical, we will have, for any given set M, $\bigwedge M = \bigwedge(M \cup \bigwedge M)$ since they are logically equivalent. The problem disappears since any given proposition p would have to be coordinated with at least two distinct sets of propositions M and N such that p is an element of one but not the other. Yet if you are someone like Russell, you would never think of identifying propositions just because they are equivalent. For, as Russell told us,

> such an escape is, in reality, impracticable, for it is quite self-evident that equivalent propositional functions are often not identical. Who will maintain, for example, that "x is an even prime other than 2" is identical with "x is one of Charles II's wise deeds or foolish sayings"? Yet these are equivalent, if a well-known epitaph is to be credited. [1903, §500]

The dual to Russell's view is just to blame the paradox on the intensionality pertaining to the notion of propositional identity. How fine-grained should be the relation of identity of equivalent propositions? An alternative to the conception of propositions as structured complexes is the modal conception of propositions, where propositions can be conceived of as sets of possible worlds. On this view, necessarily equivalent propositions are identical. Someone who favors such a coarse-grained criterion of the identity of equivalent propositions would probably consider this to be a satisfactory response to the Russell-Myhill antinomy. But, it seems to me that the moves that allow us to identify equivalent propositions have only a limited range of interest, for necessarily equivalent propositions are clearly not always identical. For this and other reasons, I do not share the view that this move affords a satisfactory resolution of the Russell-Myhill antinomy. Nonetheless, this antinomy will no doubt have as many variations in various different contexts as the set-theoretic Russell antinomy. One such variant in the context of possible-worlds semantics was pointed out by KAPLAN [1995]. So there is a structurally analogous paradox that threatens the modal conception of propositions after all.[7]

Thus we are driven to adopt one of the two courses to solve the Russell-Myhill antinomy: (i) to restrict the comprehension principle so as to block the definition of the relevant set R; or (ii) to weaken the condition for identity of propositions so as to reject (14). It was the second course that we favored in this thesis, because it accounts for the central nerve of the problem.

In conclusion, we note that the solution to the Russell antinomy based on a rejection of Cantor's theorem stands in contradiction to Zermelo's solution to the Russell antinomy, which made the Cantor theorem the basis for the subsequent development of axiomatic set theory. It also undermines the contention expressed by Ramsey and Hilbert-Ackermann, which we cited in the first part of the thesis, that the semantic antinomies are irrelevant to the foundations of mathematics. The close similarity of two versions of the Russell antinomy leaves no doubt that they must have the same solution. In this respect, I believe that Russell, Poincaré, Brouwer and Weyl were right after all. Each of these thinkers pointed out that we are dealing with essentially the same set of problems. Weyl opened up his discussion of the vicious circles in the foundations of mathematics in *Das Kontinuum* precisely by reference to the illustrative Nelson-Grelling antinomy (see WEYL [1918]).

---

[1] In particular, I acknowledge some uneasiness about the fact that the hierarchy of propositions we consider does not fit into the structure of levels and degrees considered in the first part of the thesis in any obvious way. I hope to develop a more unified line of response to this paradox elsewhere.

[2] I take inspiration and encouragement from the responses of the Russell antinomy of propositions by A. Cantini and S. Walsh (see CANTINI [2004], WALSH [2016]). Each of these authors respond to the Russell-Myhill antinomy in settings quite different from Russell's theory of types. Cantini offers two responses to the paradox: the first is a type-free theory of propositions and truth based on a combinatory algebra taken as the ground structure (cf. also ACZEL [1980] in this connection); the second is based on Quine's *New Foundations*. The set theory NF is exceptional in that it invalidates Cantor's theorem without imposing predicative restrictions; see QUINE [1937]. Walsh responds to the antinomy in the framework of Church's intensional logics and proves the consistency of a predicative fragment of Church's intensional logic by using Gödel's universe of constructible sets.

[3] See OKSANEN [1999], CANTINI [2004].

[4] See PELHAM & URQUHART [1994], LANDINI [1998], CANTINI [2004].

[5] For an axiomatic derivation of the Russell-Myhill antinomy in a simple type-theoretic framework, see CHURCH [1984]. Church formulates the relevant comprehension axioms for functional and propositional variables using a propositional identity connective taken as a primitive idea; I do not know whether the derivation would go through by appeal only to the second-order definition of Leibniz identity if instead the relevant comprehension principle were handled by using propositional *abstracts*.

[6] The conception of propositions as structured complexes is also defended by M. Dummett in connection with the Myhill antinomy (see DUMMETT [2007: 125]). See also DEMOPOULOS & CLARK [2005].

[7] Note that D. Kaplan too suggests ramification in response to his paradox. For an interesting discussion of this paradox, see BACON & HAWTHORNE & UZQUIANO [2016].

III

In this third part of the thesis, we shall attempt to justify the vicious-circle principle as a natural by-product of the semantic interpretation of the ramified second-order logic. The key to achieve the goal of coming up with an adequate interpretation of the first layer of ramified hierarchy is to switch to a non-standard way of stating the truth conditions for the second-order quantifiers in terms of substitution.[1] Substitutional interpretation of the second-order quantifiers is a natural way to explain the predicative restrictions brought by the vicious-circle principle. The explanation is simply that an assignment of truth conditions to sentences of a set-theoretic language that express propositions about sets of numbers presuppose a prior assignment of truth conditions to sentences of a number-theoretic language.

We shall do so by a simple example that illustrates how the substitution interpretation of the second-order quantifiers can be used to extend the truth definition for an antecedently given subsystem to give a truth definition for a weak second order system. One obtains a substitutional model of ramified second order logic in a well-motivated manner. For the purposes of this illustration, we shall take this subsystem to be the elementary number theory (the first-order quantification theory plus identity, the Peano axioms, and the recursive equations for addition and multiplication). So let us assume the following to be given:

1. the totality of natural numbers forming the fundamental domain D_0,
2. the language L_0 of elementary number theory interpreted in a standard way,
3. with each formula $A(x)$ of L_0 (with or without parameters) containing the variable x free an associated closed term $\{x \mid A(x)\}$ called an *set-abstract*,
4. a substitution class C comprising abstracts specifiable by arithmetical predicates of L_0.

Next, we build upon L_0 a modest set-theoretic language L containing a term t of L_0 (variable or constant) and an infinite sequence of newly introduced variables X_1, X_2, X_3, … which are not contained in L_0. The formulas of L are defined recursively as follows.

- Atomic L-formulas:

$t \in \{x \mid F(x)\}$	where F is a predicate of L_0,
$t \in X$	where X is a member of the sequence X_1, X_2, X_3, …,
$\{x \mid F(x)\} = \{x \mid G(x)\}$	where F, G are predicates of L_0,
$X = \{x \mid F(x)\}$	where F is a predicate of L_0.

- The class of L-formulas is the smallest class comprising atomic L-formulas that is closed with respect to rules of forming connective combinations and generalizations.

So we want to contrast two ways of interpreting this set-theoretic syntax.

In classical second order logic, it makes sense to think of the second-order variables as ranging over arbitrary (possibly infinite) sets of numbers, and hence second-order quantifications as involving reference to the set of all possible sets of numbers. From a predicative standpoint, the concept of the totality of sets of numbers is not acceptable, so it makes more sense to take the variables range over those sets of numbers that can be specified by arithmetical conditions expressible in the language L_0.

Using the approach of the set-theoretic model theory of second-order number theory, it is natural to define truth for L via Tarski's notion of satisfaction of a formula by an assignment of values to its constituent variables. To implement this idea requires that we specify a new domain D comprising subsets of D_0 as the range of the second-order quantifiers. A model of second order number theory is defined as a structure of the form $\langle D_0, D, I\rangle$, where the number variables x, y, z, … range over D_0, the set variables X, Y, Z, … range over D, and all the axioms of second-order number theory holds when $\in$ receives its usual meaning; I is the interpretation function such that $I(0)$ is the initial object of the domain D_0, $I(s)$, $I(+)$ and $I(\times)$ are the standard operations on D_0. If the new domain D is chosen to be the set of all subsets of D_0, one obtains the intended model of second-order number theory, which is unique up to isomorphism (thanks to the categoricity of full second-order logic). Otherwise, we have a general model in L, which is the usual way of interpreting second-order number theory.[2]

The *satisfaction condition* for the existential quantifier reads: where A is a formula of L containing a distinguished free variable X, and s is an assignment of values from D to variables (i.e. a sequence of sets of numbers), then s satisfies $(\exists X)A$ in $\langle D_0, D, I\rangle$ if and only if there is at least one sequence s' differing from s at most in what it assigns to the variable X such that s' satisfies A in $\langle D_0, D, I\rangle$.

And, the *truth condition* for the existential quantifier reads: where A is a formula of L containing a distinguished free variable X, $(\exists X)A$ is true in $\langle D_0, D, I\rangle$ if and only if $(\exists X)A$ is satisfied in $\langle D_0, D, I\rangle$ by all assignments of values from D to variables.

So, the truth condition for an existential generalization $(\exists X)A$ in L is given by the fact that $(\exists X)A$ is satisfied by all assignments of values from D to variables. Now, an assignment of values to variables involves reference to sets to no less a degree than the interpretations of constants. Under the satisfaction interpretation[3], then, we are treating variables as names of *arbitrary* sets of natural numbers in order to interpret truth-values of second-order statements; so a standardly interpreted generalization involves reference to the *totality* of sets of natural numbers. Since we want to avoid *that*, we cannot interpret the set quantifiers in this way. Fortunately, a formal truth definition for the language L can be given in a less inflationary way that does without reference to sets, via the substitution interpretation of the set quantifiers.

Under the substitution interpretation, open formulas of L are assigned no semantic value. This allows us to define the notion of truth without appealing to the notion of satisfaction, which is the major difference from the satisfaction interpretation (cf. KRIPKE [1976: 330]). So in explaining the meaning of the substitutional quantifiers we appeal to nothing except the idea of "true in L" and the syntactic idea of substitution. The quantified variables are not assumed as ranging over a domain of sets of numbers. Instead, they are associated with the substitution class C comprising abstracts that may be substituted within the scope of the quantifier to yield a closed substitution instance of the quantified formula.

Accordingly, the truth condition for the existential quantifier reads: where A is a formula of L with a distinguished free variable X, $(\exists X)A$ is true in L if and only if there is an abstract θ from C such that the closed substitution instance $A(\theta / X)$ is true in L.

However, interpreting quantification substitutionally does not prevent the truth condition of a quantification in L from reintroducing reference to sets via the truth conditions of its closed substitution instances in L_0. For, when we unwind the contribution that the connectives and the quantifiers make so as to determine the truth conditions of L-statements, we would reach to atomic L-statements. Let A be an atomic statement of the form '$t \in \{x \mid —x\ldots\}$'. Under the substitution interpretation, A is true just when there is a numeral which could be substituted for t in A to get a true sentence. If we were to construe the result of substitution as making the claim that a certain number fulfills the condition that $t \in \{x \mid —x\ldots\}$ holds, and explicate this in accordance with the satisfaction interpretation, A would be true just when the number in question stands in the membership relation to the set of all x such that $—x\ldots$ holds. So this would have reintroduced reference to the set-membership relation and to sets all the way up, in which case the difference between the standard and the substitution interpretations would not have been significant whatsoever.

What other way is there to interpret these atomic statements without assigning denotations to sets? Recall that Russell's theory of denoting was designed to explain the semantic value of a denoting phrase relative to the context in which it figures, as the contribution which it makes to determining the truth or falsity of that context. Since the closed abstracts form a species of denoting phrases, the key idea of Russell's no-class theory was to apply this analysis to yield an interpretation of statements involving the set-membership primitive without reference to sets. Because, however, Russell's no-class theory is far from modern standards, we shall adopt and adapt W. V. Quine's treatment in his book *Set Theory and its Logic* (Harvard, 1963), who developed Russell's idea of eliminating reference to sets within the range of first-order validities by the *theory of virtual classes*. Quine raised certain doubts concerning the possibility of extending the theory of virtual classes to incorporate set quantifiers, but we shall see that the substitution interpretation of the quantifiers allows us to solve that problem.[4]

The problem that Quine tackled was roughly the following: can we state set-membership claims in a first-order language without the set-membership primitive? The solution which suggested itself was to rewrite all statements of the form '$t \in \{x \mid \text{—}x\ldots\}$' by those of the form '$\text{—}t\ldots$'. On this approach, the symbol $\in$ does not stand for the set-membership relation; it is just an auxiliary symbol that plays no semantic role. Considered as a translation scheme from a weak fragment of a set-theoretic language into a first-order language, this commutes with each of the boolean operations (see QUINE [1963: 18]). In this manner, we can recover any boolean combination by a first-order formula containing the connectives and the quantifiers. So we adopt Quine's theory of virtual classes as a manual for translating atomic statements of L into corresponding statements of L_0.

This allows us to reduce the truth conditions of *atomic statements* in L to the truth conditions of arithmetical statements of L_0 as follows: Where F, G are predicates of L_0,

- $t \in \{x \mid F(x)\}$ is true in L if and only if $F(t)$ is true in L_0.
- $\{x \mid F(x)\} = \{x \mid G(x)\}$ is true in L if and only if $(\forall x)[F(x) \leftrightarrow G(x)]$ is true in L_0.

So, the interpretation of the statements of L_0 induces the interpretation of atomic statements of L without having to assign a denotation to their constituent expressions at all. As we can interpret atomic set-membership claims in L, we can also interpret their generalizations by simulating quantification over sets of natural numbers by substitutional quantification over arithmetical conditions expressible in L_0. Truth conditions for generalizations involving the set-membership primitive are explained in terms of the availability of abstracts satisfying formulas of L_0. The abstracts in L have the status of incomplete symbols, which are explained only in context, in the same way as descriptions are explained in Russell's theory of denoting.

By contrast, Quine's theory of virtual classes would not allow to generalize set-membership claims. The reason for this is simply that in this theory the quantifiers are interpreted *a la* Tarski, where the truth condition a generalization involves reference to all possible values of the bound variable at once. So the satisfaction interpretation seems to block any reasonable way to specify a translation scheme for second-order quantification without reference to sets. For this reason, Quine's theory of virtual classes could not have accommodated sentences containing second-order quantifiers. As such, there would have been no way to state in this theory the truth conditions of generalizations in L. So what makes the substitution interpretation of the quantifiers a viable choice is that it makes possible to extend this translation scheme to higher order languages containing various set quantifiers. Now let us see how we can get around this problem by interpreting the set quantifiers substitutionally.

The basic observation about the substitution interpretation of the quantifiers is that, unlike the satisfaction interpretation of the quantifiers, this move makes a existential (resp. universal) generalization truth conditionally equivalent to the possibly infinite disjunction (resp. disjunction) of its closed substitution instances (see KRIPKE [1976: 335]). So, in the case of our language L, when its second-order quantifiers are substitutionally interpreted, the truth condition of an existential (resp. universal) generalization in L is the same as the truth condition of the corresponding disjunction (resp. conjunction) of its closed substitution instances in L_0 each of which expresses an *arithmetical* condition. (Recall that this means that each of the conjuncts involves at most first-order quantification over the natural numbers D_0.) Since we are considering only definable sets specified by conditions expressible in L_0, there are no more instances of a quantified L-formula than there are predicates in L_0. But the number of predicates in L_0 is countably infinite; hence a quantified L-formula should have at most a countable infinity of instances. So to ascertain the truth of a generalization in L, it is sufficient to ascertain the truth of at most a countable infinity of substitution instances in L_0.

This suggests the idea of a mixed theory that combines the satisfaction and the substitution interpretations. On this approach, a substitutional model of L is just a structure in the sense of the model theory of first-order number theory (where it is crucial that, for each element of the structure, the language has a closed term naming it). The substitution interpretation of the set quantifiers reduces the truth condition for an existential resp. universal generalization reduces to the truth condition for the (possibly) infinite disjunction resp. conjunction of its closed substitution instances. When we get to the bottom of this construction, we have a disjunction resp. conjunction whose atomic disjuncts resp. conjuncts involve at most first-order quantification over D_0, which will be interpreted in Tarski's way.

The *satisfaction condition* for the existential quantifier now reads: where A is an formula of L containing a distinguished free variable X, and s is an assignment of values from D_0 to variables (i.e. a sequence of natural numbers), s satisfies $(\exists X)A$ in $\langle D_0, I\rangle$ if and only if, for some abstract θ from C, there is a sequence s' differing from s at most in what it assigns to the bound variables of A (if any) such that s' satisfies $A(\theta / X)$ in $\langle D_0, I\rangle$.

And, the *truth condition* for the existential quantifier reads: where A is a formula of L containing a distinguished free variable X, $(\exists X)A$ is true in $\langle D_0, I\rangle$ if and only if $(\exists X)A$ is satisfied in $\langle D_0, I\rangle$ by all assignments of values from D_0 to variables.

Clearly, the construction outlined so far hinges on the possibility of treating existential (resp. universal) generalizations as infinite disjunctions (resp. conjunctions). The language L_0 allows the formation only of finite disjunctions and conjunctions, however. This generates a residual problem concerning the interpretation of statements about infinite subsets of D_0. Technically, this outcome can be avoided without any difficulty.[5] For there are logical systems such as the infinitary logic $\mathbf{L}(\omega_1, \omega)$ that allows the formation of countably infinite conjunctions and disjunctions.[6] So we could equip the language L_0 with an infinitary logic $\mathbf{L}(\omega_1, \omega)$, where we allow quantification with respect to L_0 in the standard way, but allow countably infinite disjunctions and disjunctions. Call the extended language $\mathbf{L}(\omega_1, \omega)[L_0]$. The idea is then to translate substitutionally interpreted existential (resp. universal) generalizations in L into countably infinite disjunctions (resp. conjunctions) in $\mathbf{L}(\omega_1, \omega)[L_0]$.

Definition. The class of $\mathbf{L}(\omega_1, \omega)[L_0]$-formulas is the smallest class containing all L_0-formulas that is closed under operations of forming countably infinite conjunctions and disjunctions; thus: if $A_1, A_2, \ldots$ are formulas of $\mathbf{L}(\omega_1, \omega)[L_0]$ then so are $\bigwedge_{n<\omega} A_n$ and $\bigvee_{n<\omega} A_n$.

Let us consider an explicit translation scheme $[\ldots]^*$ from L to the language $\mathbf{L}(\omega_1, \omega)[L_0]$.

- For atomic L-statements:

  $[t \in \{x \mid F(x)\}]^* = F(t)$ where F is a predicate of L_0,

  $[\{x \mid F(x)\} = \{x \mid G(x)\}]^* = (\forall x)[F(x) \leftrightarrow G(x)]$ where F, G are predicates of L_0.

- For quantified L-statements:

  $[(\exists X)A]^* = \bigvee_{n<\omega} [A(\theta_n / X)]^*$ where $\theta_0, \theta_1, \theta_2, \ldots$ is an enumeration of C,

  $[(\forall X)A]^* = \bigwedge_{n<\omega} [A(\theta_n / X)]^*$ where $\theta_0, \theta_1, \theta_2, \ldots$ is an enumeration of C.

From a logical viewpoint, this construction represents an interpretation of the ramified second-order logic. The fact that the truth conditions of L-statements can be given in terms of the truth conditions of L_0-statements shows that the semantics of the weak second-order language L can be generated from the semantics of the language L_0 of first-order number theory. We can generate the whole of the ramified type structure by iterating this construction both over orders and over types. In fact, it can be iterated into the transfinite in the presence of suitable ordinals (see Chihara [1973], Hazen [1989]). The point was to explain the motivation for the doubly stratified structure of Russell's theory of types. I think that this construction does it in a satisfactory way.

We saw one reason why we cannot mix up types and orders by a purely semantical criterion. The stratification of the syntax of an extended language is explained in terms of the fact that its semantics can be generated out of the underlying ground language. In this manner, we construct a whole hierarchy of languages such that, at each layer of the hierarchy, the requirement of predicativity is met by the semantics of that language. We see that the basic requirement of predicativity for recursive definability is that the *definiens* be less complex than the *definiendum*.[7] So we get an explanation of the vicious-circle principle as a by-product of the substitution interpretation of the quantifiers. The explanation is that an assignment of truth conditions to sentences of L presupposes a prior assignment of truth conditions to sentences of L_0. This is a way of simulating set-existence, and reflects the idea that existence conditions for sets are parasitic upon prior identity conditions for predicatively definable sets. It was *this* idea that was difficult to make sense of in the original picture that Russell offered.

To conclude this part, I should like to mention some issues concerning the basic objects we started off with, namely, the natural numbers. The motivation for predicativity is tied to questions from the ontology of mathematics: what is out there from the perspective of the mathematician? What is indispensable for mathematical practice? What are the ontological commitments of mathematical theories? What sorts of objects lie within the range of the quantified variables of the sentences expressing mathematical propositions? Admittedly, our attitude towards the natural numbers was much less critical than it was to sets of naturals. This is clear from the fact that we posited the totality of naturals as forming the fundamental domain D_0, and explained quantification in the ground language L_0 by the satisfaction interpretation. The presupposition was that the natural numbers are 'out there' so to speak. This presupposition is repugnant to many philosophers and it is even explicitly rejected by some mathematicians in predicative foundations, namely, the so called "strict" predicativists.

Assuming the naturals as given seem problematic in at least respects. The first is that such an ontological commitment to the natural numbers is incompatible with the causal constraints on human knowledge. In general, the incompatibility between the ontological commitments of mathematical practice and the causal constraints on the knowledge of mathematical truths, is known as *Benacerraf's problem*. The substitution interpretation of the quantifiers is of no help to avoid ontological commitment to the integers, since at the level of the ground language all one can hope to do is to reduce quantification over integers to quantification over canonical names for those numbers. Yet Gödel showed us how to enumerate formulas and sequences of formulas with the help of integers. So, we can coordinate with each set of formulas and every relation between formulas a set of naturals and a relation between naturals. For this and other reasons, a reduction to quantification over formal expressions would not have brought an ontological freedom if the relevant language is the very language of number theory itself.

The second objection comes from observations about predicative mathematics itself. Russell and Poincaré saw the source of the logical antinomies precisely in impredicative definitions, whereas C. Parsons, building on M. Dummett's suggestion, argued that the concept of a natural number is impredicative independently of irrespective of the fact that it is characterized by the Dedekind-Frege-Russell definition. As Parsons claims, this makes it implausible to think that the vicious-circle paradoxes arise whenever predicative restrictions are violated (see PARSONS [1992]). While I do not think that this observation indicates a lack of robustness to the solution to the antinomies in accordance with the vicious-circle principle, it indicates that impredicative definitions do not always lead to vicious circles.

I share both kinds of uneasiness. But, there is a qualitative difference between the sort of commitments we have tried to eliminate here and the sort of commitment to the integers, so that these worries do not directly undermine the theses put forward in this thesis, or so I hope. But, they are certainly not the kinds of problems that we can pass over in silence.

---

[1] The substitution interpretation of the ramified theory is explored by PARSONS [1971a, 1971b, 1974], CHIHARA [1973], HAZEN [1989], HAZEN & DAVOREN [2000] and HODES [2015], and my presentation owes much to what I learned from these texts. As a general introduction to the substitution interpretation of the quantifiers, the papers DUNN & BELNAP [1968], MARCUS [1962, 1972], KRIPKE [1976] and UZQUIANO [2014] were also helpful.

[2] See LEIVANT [1994] and SHAPIRO [2005] for further details.

[3] In the literature, the Tarskian approach is often called *referential*, or *objectual*, interpretation of the quantifiers, but I call it the *satisfaction* interpretation, as this name suggests more clearly the difference between two techniques of interpreting quantification. It is the notion of satisfaction that is lacking in the substitution interpretation of the quantifiers after all.

[4] The idea was suggested by C. Parsons (PARSONS [1971a], [1971b]). See also (QUINE [1973: 140]).

[5] Philosophically, it is not clear if this is in conformity with Russell's ideas. So it might be incorrect to to interpret Russell's way of explaining quantifiers as infinite conjunctions. Experts on Russell claimed that "it is (on Russell's view) a crucial property of universal quantification that it allows us to express infinitely many facts by finite means." (see PELHAM & URQUHART [1994: 308]).

[6] From a meta-theoretic perspective, there are strong parallels between predicative logics and infinitary logics with *countable* languages with finite blocks of quantifiers; each of which are different than the usual first-order quantification theory (e.g. in the infinitary logic $\mathbf{L}(\omega_1, \omega)[L_0]$, completeness holds for but compactness fails; only the non-constructive version of the Löwenheim-Skølem theorem holds). Likewise, the infinitary logics with uncountable languages agree in various meta-theoretic properties with full second-order logics. A concise but excellent introduction to these ideas can be found in BELL [2016]; for a more elaborate treatment of infinitary logic, see KEISLER [1971]. Two inaugurating papers SCOTT & TARSKI [1958] and TARSKI [1958] should also be mentioned in this connection.

[7] This predicativity requirement was pointed out by C. Parsons and R.B. Marcus; S. Kripke imposed a similar requirement for the construction of a hierarchy of languages; imposing predicative requirements for a substitutional set theory goes back to V. W. Quine (see PARSONS [1971b], MARCUS [1972], QUINE [1973], KRIPKE [1976]). The fact that the ramified *Principia Mathematica* (2nd edition) admits a substitution interpretation was hinted by GÖDEL [1944] and MARCUS [1962]. This fact is formally proved by A. Hazen & J. Davoren (see HAZEN & DAVOREN [2000]).

CONCLUSION

We considered Russell's theory of types from several angles, according to which the basic requirement of predicativity is given by the vicious-circle principle. We discussed the relative advantages and disadvantages of this theory. Our discussion was not wholly impartial; part of the reason for this is that we were very influenced by the essential ideas guiding its principles; another is that while its disadvantages are widely recognized, some of its features that we discussed do not seem to have received the attention they deserve.

The two main goals of the thesis was to understand the notions of type and order and to justify the vicious-circle principle on this basis. In conclusion I shall make some suggestions concerning the vicious-circle principle.

The attempt to "deduce" the type-distinctions from the vicious-circle principle had the disastrous effect that it led one to thinking of it as a logical principle. This conception faced an insurmountable difficulty because, when treated as a logical principle concerning the objects of thought (propositions, classes, relations) that one typically deals with in logic, the vicious-circle principle itself violates the type-distinctions, which it was supposed to entail.

Similar considerations were raised long ago by the likes of F. Kaufmann and H. Behmann. Kaufmann pointed out, "the vicious-circle principle can no more be regarded as a restriction on legitimate thinking than for instance the principle of contradiction; if nevertheless one regards it as a norm, it can be taken be only as a norm for the structure of a logical unobjectionable symbolism (language)" (KAUFMANN [1978: 160]). And Behmann noted, "the very essence of the problem of paradoxes is no more nor less than the problem how to state and to apply symbolic definitions correctly, more generally, how to decide whether a given expression can be symbolically substituted in a given expression" (BEHMANN [1937: 220])

We have tried to show that the substitution interpretation of the quantifiers provide a setting in which it is possible to frame a criterion that determines whether a set of objects is or is not predicative. Given this insight it is crucial for the Russellian to reconsider the meaning and the role of types and orders, and to discard the limitations involved in Russell's original theory.

We gave an interpretation of ramified type theory based on substitutional quantification, where we employed the method of hierarchy of languages. In this framework we explained the vicious-circle principle as the requirement that the possible substituends for a variable in the relevant language may not involve that very variable itself. The interpretation of the ramified type structure made it plain that the requirements of predicativity are met at each layer of the hierarchy, since it is impossible to find in this setting impredicatively defined sets. So this whole picture seems to cohere very well with the predicative concept of set, according to which the existence condition for sets are logically parasitic upon the identity conditions on predicatively definable sets.

Perhaps most importantly, we reiterate the well-known fact that Russell was led into the ramified theory quite unwillingly, even though he saw the possibility of making type / order distinctions as early as in the *Principles of Mathematics*. We believe that these distinctions are quite natural. The point I wish to raise is that the type / order distinctions can be motivated independently of the paradoxes. It is unfortunate that the ramified theory of types has often been regarded as violent to mathematical thought. We hope that its limitations will be overcome by future work, not simply discarding the type / order distinctions, as it were they are fictitious notions to which no theoretical role can be attached, but considering alternative, more manageable, ways of formally implementing these distinctions.

BIBLIOGRAPHY

ACZEL, PETER

1980. *Frege structures and the notions of proposition, truth and set*. **The Kleene Symposium 14**(2): 217–34., edited by Jon Barwise, H. Jerome Keisler and Kenneth Kunen, pp. 31–59. Amsterdam: North-Holland.

BACON, ANDREW & HAWTHORNE, JOHN & UZQUIANO, GABRIEL

2016. *Higher-order free logic and the Prior-Kaplan paradox*. **Canadian Jour. of Phil. 46**(4-5): 493–541.

BELL, JOHN L.

2004. *Russell's paradox and diagonalization in a constructive context*. **One Hundred Years of Russell's Paradox**, edited by Godehard Link, pp. 221–225. Berlin: Walter de Gruyter.

2016. *Infinitary logic*. **The Stanford Encyclopedia of Philosophy** (Winter 2016), edited by Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2016/entries/logic-infinitary/>.

BEHMANN, HEINRICH

1931. *Zu den Widersprüchen der Logik und Mengenlehre*. **Jahr. der Deut. Math. Ver. 40**: 37–48.

1937. *The paradoxes of logic*. **Mind 46**(182): 218–221.

BETH, EVERT W.

1959. **The Foundations of Mathematics.** Reprint. New York: Harper Torchbooks, 1966.

1966. *Remarks on the paradoxes of logic and set theory*. **Essays on the Foundations of Mathematics**, dedicated to A. A. Fraenkel, edited by Yehoshua Bar–Hillel, pp. 307–311. Jerusalem: Magnes Press.

BOOLOS, GEORGE

1997. *Constructing Cantorian counterexamples*. **Jour. of Phil. Logic 26**(3): 237–239.

CANTINI, ANDREA

2004. *On a Russellian paradox about propositions and truth*. **One Hundred Years of Russell's Paradox**, edited by Godehard Link, pp. 259–284. Berlin: Walter de Gruyter.

2009. *Paradoxes, self–reference and truth in the 20th century*. **The Handbook of the History of Logic vol. 5**, edited by Dov Gabbay and John Woods, pp. 875–1013. Amsterdam: Elsevier.

CARNAP, RUDOLF.

1931. *Die logizistische Grundlegung der Mathematik*. **Erkenntnis (1930-1938) 2**: 91–105.

CHIHARA, CHARLES S.

1973. **Ontology and the Vicious-circle Principle**. Ithaca: Cornell University Press.

CHURCH, ALONZO

1940. *A formulation of of the simple theory of types*. **The Jour. of Sym. Logic 5**(2): 56–68.

1951. *A formulation of the logic of sense and denotation*. **Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer**, edited by P. Henle *et alia*, pp. 3–24: New York: Liberal Arts Press.

1956. **Introduction to Mathematical Logic**. New Jersey: Princeton University Press.

1976. *Comparison of Russell's resolution of the semantical antinomies with that of Tarski*. **JSL 41**(4): 747–60.

1984. *Russell’s theory of identity of propositions*. **Philosophia Naturalis 21**(1): 513–22.

COPI, IRVING M.

1971. **The Theory of Logical Types**. London: Routledge and Kegan Paul.

COQUAND, THIERRY.

2015. *Type theory*. **The Stanford Encyclopedia of Philosophy** (Summer 2015), edited by Edward N. Zalta. URL = <https://plato.stanford.edu/archives/sum2015/entries/type–theory/>.

CURRY, HASKELL.

1930. *Grundlagen der Kombinatorischen Logik*. **American Jour. of Math. 52**(3): 509–536.

DUMMETT, MICHAEL

2007: *Reply to Jan Dejnožka.* **The Philosophy of Michael Dummett.** The Library of Living Philosophers, edited by R. E. Auxier and L. E. Hahn, pp. 114–127. Chicago: Open Court.

DUNN, MICHAEL J. & BELNAP, NUEL D.

1968. *The substitution interpretation of the quantifiers.* **Noûs 2**(2): 177–185.

FEFERMAN, SOLOMON

1964. *Systems of predicative analysis*. **The Jour. of Sym. Logic 29**(1): 1–30.

2005. *Predicativity*. **Oxford Handbook of the Philosophy of Mathematics and Logic**, edited by Stewart Shapiro, pp. 590–624. Oxford University Press.

FREGE, GOTTLOB

1980. **Philosophical and Mathematical Correspondence**, edited by Gottfried Gabriel *et alia*, translated by Hans Kaal. Chicago: University of Chicago Press.

GANDY, ROBIN O.

1977. *The simple theory of types*. **Logic Colloquium 76**, edited by R. O. Gandy and M. Hyland, pp. 173–181. Amsterdam: North-Holland.

GÖDEL, KURT

1944. *Russell's mathematical logic*. **Collected Works**. **vol II. Publications 1938–1974**, edited by Solomon Feferman, with an introduction by Charles Parsons, pp. 102–141. Oxford University Press, 1990.

GOLDFARB, WARREN

1989. *Russell's reasons for ramification*. **Rereading Russell**: **Essays on Bertrand Russell's Metaphysics and Epistemology**, edited by C.W. Savage and C.A. Anderson, pp. 24–40. Minneapolis: U. Minnesota Press.

HAZEN, ALLEN P.

1989. *Predicative logics*. **Handbook of Philosophical Logic vol. 1**, edited by Dov M. Gabbay and Franz Guenthner, pp. 331–408. Dordrecht: Kluwer Academic Publishers.

HAZEN, ALLEN P. & DAVOREN, JENNIFER M.

2000. *Russell's 1925 logic*. **Australasian Jour. of Phil. 78**(4): 534–556.

HENKIN, LEON

1963. *A Theory of Propositional Types*. **Fundamenta Mathematicae 52**(1): 323–34.

HILBERT, DAVID & ACKERMANN, WILHELM

1928. **Grundzüge der Theoretischen Logik** (4. Auflage). Berlin: Springer, 1956.

HODES, HAROLD

2015. *Why ramify?* **Notre Dame Jour. of Formal Logic 56**(2): 379–415.

KAPLAN, DAVID

1995. *A problem in possible-world semantics.* **Modality, Morality and Belief**: **Essays in Honor of Ruth Barcan Marcus**, edited by W. Sinott-Armstrong, D. Raffman, N. Asher, pp. 41-52. Cambridge U. Press.

KAUFMANN, FELIX

1930. **Das Unendliche in der Mathematik und seine Ausschaltung.** Leibzig und Wien: Franz Deuticke. Translated by Paul Foulkes in *The Infinite in Mathematics: logico-mathematical writings*, edited by Brian McGuinness, with an introduction by and Ernst Nagel. Dordrecht: D. Reidel Publishing Co., 1978.

KEISLER, H. JEROME

1971. **Model Theory for Infinitary Logic**. Amsterdam: North-Holland.

KLEENE, STEPHEN C.

1952. **Introduction to Metamathematics**. Groningen: Noordhoff.

KLEMENT, KEVIN C.

2001. *Russell's paradox in Appendix B of the Principles of Mathematics: was Frege's response adequate?* **History and Philosophy of Logic 22**: 13–28.

2014. *Russell-Myhill paradox.* **The Internet Encyclopedia of Philosophy**, edited by J. Fieser and B. Dowden. URL = <http://www.iepp.utm.edu/>.

KRIPKE, SAUL A.

1976. *Is there a problem about substitutional quantification?* **Truth and Meaning**, edited by Gareth Evans and John McDowell, pp. 324-419. Oxford University Press.

LANDINI, GREGORY

1998. **Russell's Hidden Substitutional Theory.** Oxford University Press.

LEIVANT, DANIEL

1994. *Higher order logic*. **Handbook of Logic in Artificial Intelligence and Logic Programming**, pp. 229–321. Oxford University Press.

LINSKY, LEONARD

1988. *Terms and propositions in Russell's Principles of Mathematics*. **Jour. History of Phil. 26**(4): 621–642.

MARCUS, RUTH BARCAN

1962. *Interpreting quantification*. **Inquiry 4**(1-4): 252–259.

1972. *Quantification and ontology*. **Noûs 6**(3): 240–250.

MYHILL, JOHN R.

1958. *Problems Arising in the Formalization of Intensional Logic*. **Logique et Analyse 1**: 78–83.

1974. *The Undefinability of the Set of Natural Numbers in the Ramified Principia*. **Bertrand Russell's Philosophy**, edited by G. Nakhnikian, pp. 19–27. London: Duckworth.

1979. *A refutation of an unjustified attack on the axiom of reducibility*. **Bertrand Russell Memorial Volume**, edited by G. Roberts, pp. 81–90. London: Allen and Unwin.

OKSANEN, MIKA

1999. *The Russell-Kaplan paradox and other modal paradoxes; a new solution*. **Nordic Journal of Philosophical Logic 4**: 73–93.

PARSONS, CHARLES

1971a. *Ontology and mathematics*. **The Philosophical Review 80**(2): 151–176.

1971b. *A plea for substitutional quantification*. **The Journal of Philosophy 68**(8): 231–237.

1974. *Sets and classes*. **Noûs 8**(1): 1–12.

1992. *The impredicativity of induction*. **Proof, Logic and Formalization**, edited by Michael Detlefsen, pp. 139–161. London: Routledge.

PELHAM, JUDY & URQUHART, ALASDAIR

1994. *Russellian propositions*. **Logic, Methodology and Philosophy of Science**, edited by Dag Prawitz, Brian Skyrms and Dag Westerståhl, pp. 307–326. Amsterdam: Elsevier.

POINCARÉ, JULES HENRI

1910. *Über transfinite Zahlen*. **Sechs Vorträge über ausgewählte Gegenstände aus der reinen Mathematik und math. Physik**. Math. Vorlesungen an der Universität Göttingen, pp. 44-48. Leipzig/Berlin: Teubner.

1913. *The logic of infinity*. **Mathematics and Science: Last Essays**, translated by J. W. Bolduc, pp. 45–64. New York: Dover Publications, 1963.

QUINE, WILLARD V. O.

1937. *New Foundations for Mathematical Logic*. **American Mathematical Monthly 44**: 70–80.

1938. *On the theory of types*. **The Journal of Symbolic Logic 3**(4): 125–139.

1963. **Set Theory and Its Logic**. Cambridge, MA: Belknap Press.

1973. **The Roots of Reference**. Le Salle, IL: Open Court Press.

RAMSEY, FRANK P.

1931: **The Foundations of Mathematics and other Logical Essays**, edited by R. B. Braithwaite, with an introduction by G. E. Moore, pp. 1–61. Reprint. Paterson, NJ: Littlefield, Adams & Co, 1960.

RUSSELL, BERTRAND

1903. **Principles of Mathematics**, 2nd ed. New York: W.W. Norton.

1905. *On some difficulties in the theory of transfinite numbers and order types*. In RUSSELL (1973), pp. 135.

1906. *On 'insolubilia' and their solution by symbolic logic*. In RUSSELL (1973), pp. 190–214.

1908. *Mathematical logic based on the theory of types.* In RUSSELL (1956), pp. 57–102.

1910. *The theory of logical types*. In RUSSELL (1973), pp. 215–252.

1919. **Introduction to Mathematical Philosophy.** Reprint. Edinburgh: Neill & Co. Ltd., 1950.

1956. **Logic and Knowledge: Essays 1901-1950**, edited by R. C. Marsh. London: George Allen & Unwin Ltd.

1973. **Essays in Analysis**, edited by Douglas Lackey. New York: George Braziller.

SCOTT, DANA & TARSKI, ALFRED

1958. *The sentential calculus with infinitely long expressions*. **Colloquium Mathematicum 16**: 166–170.

SHAPIRO, STEWART

2005. *Higher-order logic*. **Oxford Handbook of the Philosophy of Mathematics and Logic**, edited by Stewart Shapiro, pp. 751–780. Oxford University Press.

TARSKI, ALFRED

1958. *Remarks on predicate logic with infinitely long expressions*. **Colloquium Mathematicum 16:** 171–176.

URQUHART, ALASDAIR

1988. *Russell's zigzag path to the ramified theory of types*. **Russell 8**(1): 82–91.

2001. *The theory of types*. **The Cambridge Companion to Bertrand Russell**, edited by Nicholas Griffin, pp. 286–309. Cambridge University Press.

UZQUIANO, GABRIEL

2014. *Quantifiers and quantification*. **The Stanford Encyclopedia of Philosophy** (Winter 2016), edited by Edward N. Zalta, URL = <https://plato.stanford.edu/archives/win2016/entries/quantification/>.

2015. *A neglected resolution of Russell’s paradox of propositions*. **Review of Symbolic Logic 8**: 328–344.

WALSH, SEAN

2016. *Predicativity, the Russell-Myhill paradox, and Church’s Intensional Logic*. **J. Phil. Logic 45**: 277–326.

WANG, HAO

1959. *Ordinal numbers and predicative set theory.* **Zeitschr. f. math. Logik und Grund. d. Math 5**: 216–239.

WEYL, HERMANN

1918. **Das Kontinuum**. *Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Veit.

1921. *Über die neue Grundlagenkrise der Mathematik*. **Mathematische Zeitschrift 10**: 37–79.

WHITEHEAD, ALFRED N. & RUSSELL, BERTRAND

1910. **Principia Mathematica** (3 volumes, first edition). Cambridge University Press.

1925. **Principia Mathematica** (3 volumes, second edition). Cambridge University Press.

ZERMELO, ERNST

1908. *Untersuchungen über die Grundlagen der Mengenlehre I.* **Math. Annalen 65**(2): 261–281.