THE REVIEW OF SYMBOLIC LOGIC
Volume 10, Number 4, December 2017

WHAT RUSSELL SHOULD HAVE SAID TO BURALI–FORTI

SALVATORE FLORIO
University of Birmingham

GRAHAM LEACH-KROUSE
Kansas State University

Abstract. The paradox that appears under Burali–Forti's name in many textbooks of set theory is a clever piece of reasoning leading to an unproblematic theorem. The theorem asserts that the ordinals do not form a set. For such a set would be–absurdly–an ordinal greater than any ordinal in the set of all ordinals. In this article, we argue that the paradox of Burali–Forti is first and foremost a problem about concept formation by abstraction, not about sets. We contend, furthermore, that some hundred years after its discovery the paradox is still without any fully satisfactory resolution. A survey of the current literature reveals one key assumption of the paradox that has gone unquestioned, namely the assumption that ordinals are objects. Taking the lead from Russell's no class theory, we interpret talk of ordinals as an efficient way of conveying higher-order logical truths. The resulting theory of ordinals is formally adequate to standard intuitions about ordinals, expresses a conception of ordinal number capable of resolving Burali–Forti's paradox, and offers a novel contribution to the longstanding program of reducing mathematics to higher-order logic.

§1. Introduction. A simple, informal statement of the paradox of the greatest ordinal—the paradox commonly attributed to Cesare Burali–Forti—might be this:

> Given any well-ordered collection of objects, we may ask which position an object in the collection occupies. Indeed, since the collection is well-ordered, an object will come first (the least object in the ordering), another will come second (the least object in the ordering besides the first), yet another will come third (the least object in the ordering besides the first and the second), and so on. Other objects might come after those occupying the finite positions. So there might be the least object besides those occupying the finite positions, another object after that, and so on again. Ordinal numbers are objects representing positions in well-ordered collections of objects. Evidently, there is natural ordering of the ordinal numbers: the object representing a certain position comes before the object representing later positions (e.g., the object representing the first position comes before the one representing the second positions). And so each ordinal occupies the position it represents.
> However, we can define well-orderings that *extend* the natural ordering of the ordinal numbers. For example, we can place some other object r

Received: August 19, 2016.
2010 *Mathematics Subject Classification*: 00A30, 03A05, 01A.
Key words and phrases: Burali-Forti, Russell, abstraction, higher-order logic, logicism, no class theory, ordinals, paradox.

doi:10.1017/S1755020316000484

> immediately after all the ordinals. But now the object r is left without a position. Since this collection is well-ordered, there should be an ordinal representing the position of r. But that cannot be: the ordinals have already been 'used up' to represent their own positions. This calls into question the coherence of our intuitive talk of ordering and positions.

The form of the paradox which appears under Burali–Forti's name in many textbooks of set theory goes something like this:

> Consider the class of Von Neumann ordinals—transitive sets well-ordered by the membership relation, each one representing the ordinal giving the count of its own members ordered by $\in$. Since it is easy to verify that for any two Von Neumann ordinals, one is a member of the other, and since the restriction of $\in$ to this class is transitive and well-founded almost by definition, this class, if it were a set, would be another Von Neumann ordinal, representing the count of the whole ordinal sequence. But then it would contain itself (since it contains *all* Von Neumann ordinals) contradicting the earlier claim that it was well-ordered by $\in$. Thus, the class of Von Neumann ordinals is not a set.

Stated in this way, Burali–Forti's paradox leads to an unproblematic theorem of Zermelo-Fraenkel set theory with the Axiom of Choice (henceforth ZFC), and indeed of much weaker set theories. The ordinals constructed in set theory do not form a set. For, such a set would be—absurdly—an ordinal greater than any ordinal in the set of all ordinals.

In what follows, we shall argue that the standard set-theoretic rendering of the paradox of Burali–Forti conceals the conceptual difficulty that the paradox discloses, and that this difficulty is, some hundred years after its discovery, still without any fully satisfactory resolution. In the course of our argument, we will propose an analysis of this paradox, and show how the extant approaches to the paradox can be helpfully classified by means of this analysis. Our classification will highlight one possible resolution to the paradox—an account of the ordinals along the lines of Russell's no class theory—which seems to be both promising, and neglected by the current literature. In the final sections of the paper, we will develop and motivate a no class theory of ordinals which is formally adequate to standard intuitions about ordinals, expresses a conception of ordinal number capable of resolving Burali–Forti's paradox, and offers a novel contribution to the longstanding program of reducing mathematics to higher-order logic.

1.1. Why set theory does not resolve the paradox of Burali–Forti. To see why the paradox of Burali–Forti is not resolved by set theory alone, it may be helpful to compare it to another paradox that set theory *can* in some sense resolve. A good example of such a paradox would be Cantor's paradox:

> Consider the set of all sets, the universal set U. By Cantor's theorem, we know that the set of all the subsets of U, $\mathcal{P}(U)$ cannot be placed into one-to-one correspondence with U. Yet, there is obviously an embedding of U into $\mathcal{P}(U)$ (by mapping each set to its singleton), and of $\mathcal{P}(U)$ into U (by mapping each set to itself). So by the Schröder-Bernstein theorem, there *is* a one-to-one correspondence between U and $\mathcal{P}(U)$.

Most set theorists will say that this argument goes wrong in the very first sentence. There is no universal set.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

It seems that ZFC supports this standard diagnosis. It licenses every other part of the paradoxical reasoning above, but does not prove that there is a universal set. So by process of elimination, the claim that there is a set of all sets must be rejected—and indeed, "by Cantor's paradox", it is a theorem that there is no such set. But diagnosing the paradoxical reasoning above by appealing to ZFC only pushes the problem back to the justification of ZFC. For, why should we think that ZFC is right to say there is no universal set? There are, after all, alternative set theories in which such a set is countenanced.

The solution to this trouble is to give a *conception of set*: a picture of the universe of sets on which our axioms and much of our naïve reasoning about sets can be seen to be justified. Such a picture helps fix a clear subject matter for the axioms, and elevate them from a set of so-far-consistent assertions to a well-motivated account of the nature of that subject matter.

The conception of set traditionally favored by set theorists who are in the business of justifying their axioms is called the iterative conception. Here is how Gödel describes it:

> When theorems about all sets (or the existence of sets in general) are asserted, they can always be interpreted without any difficulty to mean that they hold for sets of integers as well as for sets of sets of integers, etc. (respectively, that there either exist sets of integers, or sets of sets of integers, or ... etc., which have the asserted property). This concept of set, according to which a set is something obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of"... has never led to any antinomy whatsoever; that is, the perfectly "naïve" and uncritical working with this concept of set has so far proved completely self-consistent. (Gödel, 1964, 180)

This is a description of an iterative hierarchy of sets that begins from the integers. ZFC is the theory of the iterative hierarchy of sets that begins from the empty set.

The iterative conception, when elaborated in sufficient detail, supports most or all of the axioms of ZFC. It also explains why there is no universal set. Such a set would have to contain itself. But a set containing itself cannot arise at any point in the accumulation process: each set is built out of previously built sets, and no set is built before it is built.

The iterative conception also explains why there is no set of all Von Neumann ordinals. The Von Neumann ordinals constructed at any given point in the accumulation process, when assembled into a set at the next stage, always form a new Von Neumann ordinal. So a stage never occurs at which *all* the Von Neumann ordinals are at previous stages. Hence, the Von Neumann ordinals are never available all at once to be assembled into a set.

Does this explanation of the fact that there is no set of all Von Neumann ordinals constitute a solution to the original paradox of Burali–Forti? We think not, at least in the sense that the paradox of Burali–Forti is not directly resolved by the iterative conception in the same way as Cantor's paradox is. We do not mean to deny that the iterative conception could be part of some solution to the paradox of Burali–Forti. But Cantor's paradox is about the properties of a certain set, whereas the paradox of Burali–Forti is about ordinal numbers. A conception of set, at least in isolation, suffices to address a paradox of the former kind but not of the latter kind. If ordinal numbers are not Von Neumann ordinals, there is a gap here that needs to be bridged before we can claim to have resolved the paradox. We need to be told why a feature of the represented structure (the progression of

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

ordinal numbers) can be explained by features of a certain representational medium (the Von Neumann ordinals). In the next section, we offer some reasons to think that ordinal numbers are not Von Neumann ordinals.

1.2. Origins of the ordinal number concept. Are the ordinals literally the Von Neumann ordinals? No, for at least two reasons. First, the notion of transfinite iteration that gives us the iterative hierarchy of sets in which the Von Neumann ordinals are situated appears to presuppose ordinals—or something akin to them—to index the stages of the process. If we revise our usage to refer only to Von Neumann ordinals, we lose sight of the ordinals that give us the iterative conception. Second, even if one denies that the ordinals are presupposed or denies the iterative conception, say by embracing an alternative conception of set, the Von Neumann ordinals are still not the ordinals as they were historically conceived. One is free to adopt a new terminology on which ordinals are just Von Neumann ordinals. But to do so would be to disown the fruitful tradition that originally gave us the concept of ordinal number. In this section, we review some features of the tradition that will be important to our analysis of the paradox of Burali–Forti.

Let us begin by recalling a definition associated with *cardinal* numbers, rather than ordinal numbers. This is Hume's principle, the principle that the number of As is the same as the number of Bs, just in case there is a one-to-one correspondence between the As and the Bs. In symbols:

$$\forall A\, \forall B\, (\#A = \#B \leftrightarrow A \approx B). \tag{HP}$$

The left-hand side expresses the literal equality of an object, the number of As (denoted by the "number of" abstraction operator # applied to the variable A), to another object, the number of Bs. The right-hand side abbreviates the statement that there is a function witnessing the equinumerosity of A and B.

While Frege sometimes gets credit for this principle—perhaps, ironically, because he was the first to attribute it to Hume (1884), perhaps because he indirectly showed how this principle, over second-order logic, has the strength of second-order Peano Arithmetic, or perhaps because he sometimes refers to it in the *Grundlagen* as "my definition"—its lineage appears to considerably predate the attribution.

It would take us much too far afield to fully explore Frege's sources and influences, although there's much to be said.[1] According to recent historical work by Paolo Mancosu, Frege here is at least partly influenced by the innovations of Grassmann's *Ausdehnungslehre*—originally published about forty years before Frege's *Grundlangen*—which includes a number of definitions by abstraction exactly analogous to Hume's principle. For example, Frege's definition of direction is as the thing that two parallel lines have in common (so that two directions are equal when two lines having them are parallel) can be found in texts by Grassmann, e.g., Grassmann, 1844 and Grassmann, 1847 (see Mancosu, 2015 for details). Mancosu's work strongly suggests that this style of definition would likely have been recognizable to a reader of Frege's time.

---

[1] When Hume's principle is introduced in the *Grundlagen*, Frege alludes the work of others, both for the idea that number should be defined in terms of a one-to-one correspondence (Frege, 1884, 73–74), and for the idea that in general, a new type of entity can be introduced by giving sense to identity statements (like "the number of As is equal to the number of Bs") concerning that type of entity. (Frege ultimately rejected the latter claim.) Frege's allusion to the work of other mathematicians using one-to-one correspondence in the elucidation of the number concept mentions Cantor, Schröder, and Kossak. The reference to Cantor is especially important to what follows.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

And, indeed, the basic idea of definition by abstraction seemed to have been clear to Leibniz.[2] He explains his own use of an abstraction-like procedure in his correspondence with Clarke:

> I have here done much like Euclid, who, not being able to make his readers understand what *ratio* is absolutely in the sense of the geometricians, defines what are *the same ratios*. Thus in like manner, in order the explain what *place* is, I have been content to define what is *the same place*. (Leibniz, 1989, 704)

Leibniz is referring to Euclid's definition 5, from the fifth book of the *Elements*:

> Magnitudes are said to **be in the same ratio**, the first to the second and the third to the fourth, when if any equimultiples whatever be taken of the first and the third, and any equimultiples whatever of the second and the fourth, the former equimultiples alike exceed, alike are equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order. (Translated in Heath, 1908, 114)

We have approximately what one might call Euclid's principle:[3]

$$\begin{aligned}\forall m_1, m_2, m_3, m_4 \big(r(m_1, m_2) = r(m_3, m_4) \leftrightarrow \\ \forall a, b\, ((a \cdot m_1 > b \cdot m_2 \;\wedge\; a \cdot m_3 > b \cdot m_4) \\ \vee (a \cdot m_1 = b \cdot m_2 \;\wedge\; a \cdot m_3 = b \cdot m_4) \\ \vee (a \cdot m_1 < b \cdot m_2 \;\wedge\; a \cdot m_3 < b \cdot m_4))\big).\end{aligned} \tag{EP}$$

So, in formulating Hume's principle, Frege had a number of influential antecedents.

It should not be too surprising, then, that in 1878, six years before the publication of Frege's *Grundlagen*, Cantor offered the following definition of "power" or cardinal number:

> If two well-defined manifolds M and N can be coordinated with each other univocally and completely, element by element (which, if possible in one way, can always happen in many others), we shall employ in the sequel the expression, that those manifolds have the same power. (Cantor, 1878; see Ferreirós, 2007, 188)

---

[2] Leibniz at least indirectly influenced Frege and Grassmann. Indeed, both defend the adequacy of their definitions by abstraction in precisely the same way: by pointing to Leibniz. Frege quotes Leibniz's dictum that "Things are the same as each other when one can be substituted for the other without loss of truth" (Frege, 1884, 76) and takes the justification of definition by abstraction to consist, as least in part, in showing that the abstracts obey this principle. Grassmann took himself to be improving on Leibniz's geometrical calculus by introducing new symbols (in hindsight, abstraction operators) that would allow one to subsume the transfer of properties between congruent geometrical figures to simple operations of substitution of identicals. As in Frege, this maneuver is justified by a Leibnizian conception of equality "in which we just set as simply equal that which we can substitute for the other in each proposition." (Grassmann, 1844, 321). Whether Leibniz's idea of definition by abstraction had a direct influence on either figure, or whether Leibniz's influence is confined to the laws of identity, remains an interesting historical question which we do not attempt to answer here.

[3] There is an exegetical issue concerning the use of the identity sign in (EP). In proposition 11, book 5, Euclid proves that it is legitimate to reason with the notion of sameness of ratio as one does with the notion of identity.

Modulo the use of "manifolds" rather than Fregean concepts, this is just Hume's principle. So Cantor too was a part of the tradition described above—indeed Cantor writes admiringly of Grassmann's style of definition (Cantor, 1883, 897).

We have now described the concept of cardinal number given in Hume's principle and the intellectual tradition from which that concept emerged. To summarize the crucial point: definition by abstraction was at least a recognized technique around the time of Cantor's greatest innovations and was part of Cantor's methodological repertoire. We wish to claim that Burali–Forti's paradox is best viewed as *a paradox within this practice of concept formation*, rather than a paradox within the concept of set. With that background in place, let us return from the concept of cardinal number to our original target: the concept of *ordinal* number.

The definition of ordinal number we are interested in seems to first have been systematically presented in Georg Cantor's *Grundlagen* in 1883; he arrived at the concept over the month following his September 1882 meeting with Dedekind (see Ferreirós, 2007, 269).[4] The critical idea, likely stimulated by conversation with Dedekind, is the possibility of taking the order-theoretic aspects of numbers—their positions in the number sequence—as their essential or defining features, as Dedekind was inclined to do. While Cantor would continue to maintain that ordinary numbers are essentially cardinals—essentially connected with considerations of size, rather than order—he was not blind to the intelligibility of Dedekind's point of view.

To capture the concept of number connected with order (Cantor used the term *Anzahl* to express this concept), one wants a definition that establishes that order-types correspond to numbers. Roughly, this is a principle about ordinals requiring that every well-ordering R is represented by an ordinal number determined entirely by the order type of R. In effect, one introduces ordinals as invariants of ordered collections, just as one introduces cardinals as invariants of bare unstructured collections. The obvious definition is given by the following principle of ordinal abstraction:

$$\forall R_1 \forall R_2(\mathbf{ord}(R_1) = \mathbf{ord}(R_2) \;\leftrightarrow\; R_1 \cong R_2). \tag{O-AB}$$

Now the left-hand side expresses the literal equality of an object, the ordinal of R_1 (denoted by the "ordinal of" abstraction operator **ord** applied to the variable R_1), to another object, the ordinal of R_2. The right-hand side abbreviates the statement that there is a one-to-one order preserving correspondence between well-orderings (an order isomorphism). The quantifiers range over well-orderings.[5]

The obvious definition is, indeed, precisely what Cantor proposed:

> Two well-ordered sets are now said to be of the same Anzahl (with respect to their given successions) when a reciprocal one-to-one correlation of them is possible such that, if E and F are any two elements of the one set, and E_1 and F_1, are the corresponding elements of the other, then the position of E and F in the succession of the first set always

[4] But it should of course be noted that the concept of ordinal, as a counting number rather than a measure of cardinality, obviously predates Cantor, and the concept of an infinite ordinal is prefigured by Cantor's work on trigonometric series, where terms behaving like names for transfinite ordinals appear in formal manipulations as uninterpreted "symbols of infinity".

[5] If one makes the restriction of the quantifiers explicit, the result is:

$$\forall R_1 \forall R_2(\mathrm{WO}(R_1) \;\wedge\; \mathrm{WO}(R_2) \rightarrow (\mathbf{ord}(R_1) = \mathbf{ord}(R_2) \;\leftrightarrow\; R_1 \cong R_2)),$$

where $\mathrm{WO}(R)$ abbreviates the statement that R is a well-ordering (see Appendix A for a formal definition).

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

> agrees with the position of E_1, and F_1 in the succession of the second set (i.e., when E precedes F in the succession of the first set, then E_1, also precedes F_1, in the succession of the second set). (Cantor, 1883, 885)

And it is *this* definition that leads to the original Burali–Forti paradox.

§2. The paradox of Burali–Forti. In his paper on Cantor's ordinal number concept, Burali–Forti (1897) gives us the following definition of ordinal number (in Peano's notation)

$$(u, h), (v, k) \, \varepsilon \, \mathrm{Ko} \, .\supset\!: T\text{`}(u, h) \; = \; T\text{`}(u, k)$$
$$. = . \, (u, h) \sim (v, k).$$

What this means, essentially, is that, if (u, h), (v, h) are pairs each consisting of a collection and a well-ordering of that collection, then their ordinal numbers are equal so long as their orderings are isomorphic. This is, in effect, just a notational variant on our earlier principle of ordinal abstraction. And it was on this basis—not on the basis of any formalization within set theory—that Burali–Forti outlined the argument that now bears his name.

Burali–Forti's original presentation is marred by a number of difficulties. The most serious of these is that his theorem turns on a misunderstanding of Cantor's definition of a well-ordering. Burali–Forti seems to have understood this to mean "ordering with no descending ω-sequence under the predecessor relation" rather than "ordering with no descending ω-sequence under the less-than relation". So rather than claiming to have shown that ordinal abstraction is inconsistent, Burali–Forti merely claimed to have established, by his paradox, a reductio of the claim that the "ordinals" are linearly ordered. Others, however—most notably Russell and Poincaré—soon gave Burali–Forti's argument a sharper edge.[6]

Combining the findings of Burali–Forti, Russell, and Jourdain, the following paradox emerges. Our presentation of the paradoxical reasoning relies on second-order logic to capture talk about relations. This is to a certain extent a rational reconstruction. We do not strive to give a fully faithful rendering of the exact thoughts of the different historical actors. In addition to providing a simple and illuminating setting, the second-order framework has the advantage of letting us see the continuity between the origins of the paradox and contemporary work on abstraction principles.

Argument sketch. There is a natural way of comparing two ordinals. One ordinal α is greater than another β if a well-ordering R_1 represented by α is "longer" than a well-ordering R_2 represented by β (that is, if R_2 is isomorphic to a proper initial segment of R_1).

Using standard higher-order resources, one can demonstrate that there is a relation $<$ on the ordinals corresponding to this notion of greater-than, and that this relation is a well-ordering (this was proved by Jourdain; see his 1904 article). The comprehension axioms used in the argument are impredicative—these are instances in which the comprehending

[6] The first presentation of the argument as a paradox came in 1903, in Russell's *Principles of Mathematics*. There, Russell reconstructs Burali–Forti's argument, but with the correct Cantorian definition of ordinal number. Apparently unaware of the difference between Burali–Forti and Cantor's definitions, Russell points out that one can easily establish that every initial segment of the ordinal number sequence is well-ordered (and indicates that he believes Cantor's proof that the ordinals are linearly ordered is correct). Like Burali–Forti, however, he ultimately views the argument as a reductio, this time of the claim that the entire ordinal sequence is well-ordered—an assertion that Russell found intuitive, but not impossible to abandon. Jourdain's proof that the ordinals are, in fact, well-ordered would later force Russell to take more aggressive measures in his excision of the inconsistency (see Moore & Garciadiego, 1981 for details).

formula contains quantifiers over properties or relations. So one must at this stage assume that to impredicative definitions there correspond relations, and furthermore, that a single existential quantifier in a comprehension axiom can range over *all* such relations. For the relation $<$ to have the desired interpretation, it's also necessary to read the first-order quantifiers in the comprehension scheme as ranging over absolutely all objects. (These presuppositions will be important later.)

Once we have the relation $<$, for any ordinal α, we can then show (using only a more harmless form of comprehension) that there is a relation that restricts $<$ to the ordinals less than α. Call this relation $<_\alpha$. If one grants that $<_\alpha$ lies within the domain of ordinal abstraction—in particular that it is a well-ordering of *objects*, and so that ordinals are objects—then it is not difficult to prove that for any ordinal α, α is the representative of $<_\alpha$.

By an application of O-AB—again, taking the second-order quantifiers in this principle to range over *all* relations—one has that to $<$, there corresponds an ordinal, which we can call Ω. This means that Ω is the representative of $<$. But, by the previous fact, Ω is also the representative of $<_\Omega$. Since by definition $<_\Omega$ is a proper initial segment of $<$, the representative of $<_\Omega$ is less than the representative of $<$. In other words, $\Omega < \Omega$, which contradicts the fact that $<$ is a well-ordering. □

A careful formalization of the proof above—rather too long and involved to be included here, although it is outlined in Appendix B—establishes that O-AB is inconsistent over second-order logic. The model-theoretic unsatisfiability of O-AB (in a standard model) is a well-known fact (see, e.g., Hodes, 1986; Cook, 2003; Linnebo & Pettigrew, 2014). A proof of the principle's inconsistency is surprisingly intricate.[7] We hope to provide a useful reference by laying out an explicit proof, which is so far lacking in the literature.

If we are to make a proper philosophical assessment of the paradox, we must make its assumptions fully explicit. A model-theoretic proof of the unsatisfiability of O-AB in a standard model of second-order logic, however, already builds in far-reaching assumptions about quantification and second-order ontology. This is why a proof-theoretic perspective on the paradox promises to be more philosophically illuminating. A formalization of the paradox reveals its dependence on two assumptions, which we drew attention to in the proof sketch above:

> *Second-order comprehension*
> Every open formula defines a property or a relation.

> *First principle of ordinals* (O-AB)
> Every well-ordering R is represented by an ordinal determined entirely by the order type of R.

Moreover, the paradox relies on three presuppositions:

> *First-order absolute generality*
> It is possible to quantify over absolutely all ordinals.

> *Second-order absolute generality*
> It is possible to quantify over absolutely all relations.

[7] The step from unsatisfiability to inconsistency is not necessarily trivial. Hume's principle and Wright's nuisance principle (Wright, 1999) have long been known to be jointly unsatisfiable, while their joint inconsistency (given some assumptions about choice or pairing principles) has only recently been established in Walsh & Ebels-Duggan (2015), Ebels-Duggan (2015), and Walsh (in press).

Second principle of ordinals
Ordinals are objects representing well-orderings.

Each of these assumptions and presuppositions underwrites one or more crucial steps in the proof above. The first principle of ordinals motivates O-AB, which is used to establish that $<$ is a well-ordering, to prove that $\mathbf{ord}(<_\alpha) = \alpha$, and to reason about Ω. Instances of second-order comprehension are used throughout the proof that the ordinals are well-ordered. In particular, as we noted, one needs instances in which the comprehending formula contains quantifiers over properties or relations, i.e., *impredicative* instances of comprehension. It turns out that a full formalization of our proof that $<$ is a well-ordering requires what is called Σ^1_1-comprehension.[8] Moreover, the definition of $<$ requires an instance of comprehension that is not only impredicative but viciously so.

Intuitively, an instance of comprehension is *viciously impredicative* if the comprehending formula must be taken to quantify over the very relation that it defines. But this is what is required in the final stages of the paradox. The defining condition of the relation $<$ is that for any two ordinals α and β, $\alpha < \beta$ just in case there are two well-orderings R_1 and R_2, such that α is a representative of R_1, β is a representative of R_2, and R_1 is isomorphic to a proper initial segment of R_2. We conclude the paradox by moving from the fact that $<_\Omega$ is proper initial segment of $<$ to the assertion that $\Omega < \Omega$. But this inference involves taking R_1 to be $<_\Omega$ and R_2 to be $<$. So the relation being defined, $<$, must be in the range of the existential quantifier binding R_2 in the comprehending formula.

Finally, throughout the proof, the use of classical inference rules for first-order and second-order quantifiers presupposes a fixed domain for each type of variables to range over. If different occurrences of any given quantifier ranged over different domains, some rules of classical logic would have to be given up. For instance, universal instantiation would fail, since there would be no guarantee that a given universal quantifier would include in its range objects introduced by existential quantifiers occurring elsewhere. This is illustrated by a well-known response to Russell's paradox. The following sentence is at the core of the paradox:

$$\exists y\, \forall x\, (x \in y \leftrightarrow x \notin x).$$

[8] Roughly, an instance of comprehension is said to be Σ^1_1 if it involves a formula that is equivalent to one headed by a block of second-order existential quantifiers and containing no other second-order quantifier. It is Π^1_1 if the relevant formula is headed by a block of second-order universal quantifiers. An instance of comprehension is said to be Δ^1_1 if it is both Σ^1_1 and Π^1_1.

The use of Σ^1_1-comprehension shows up, for example, in Lemma 8.2 of Appendix B asserting that if every proper initial segment of one well-ordering R_1 is isomorphic to a proper initial segment of another R_2, then R_1 is isomorphic to an initial segment of R_2—this intuitively requires a union of given isomorphisms, which requires Σ^1_1-comprehension.

It's plausible that no comprehension principle weaker than Σ^1_1 will support the lemma. For one cannot prove that

for any two well-orderings, one is isomorphic to an initial segment of the other (*)

using only Δ^1_1-comprehension. But (*) is logical consequence of the lemma. The unprovability of (*) from Δ^1_1-comprehension follows from the fact that, in the language of second-order arithmetic, (*) is equivalent over RCA_0 to ATR_0 (Simpson, 2009, 198), which is strictly stronger than the theory of Δ^1_1-comprehension (Simpson, 2009, 338–345). However, the conclusion that the lemma requires Σ^1_1-comprehension is not immediate owing to the presence of O-AB. For O-AB might add enough strength to theory of Δ^1_1-comprehension to allows us to derive (*).

One might resist the reasoning that leads from this sentence to contradiction by interpreting the range of the existential quantifier as wider than the range of the universal quantifier (see, e.g., Parsons, 1974b; Parsons, 1974a; Glanzberg, 2004; Glanzberg, 2006). A more perspicuous notation would thus be:

$$\exists^{+} y\, \forall x\, (x \in y \;\leftrightarrow\; x \notin x).$$

Although in the course of an argument, one might consider a witness, r, to the existential quantifier above, such a witness could lie outside of the range of the universal quantifier. This makes the inference to

$$r \in r \;\leftrightarrow\; r \notin r$$

illegitimate, since the universally bound x cannot be be instantiated by r. First-order and second-order absolute generality secure a fixed domain for first-order and second-order quantifiers. For if all quantifiers range over the maximal domain, no disagreement about quantifier range is possible.

An intuitive picture of domain expansion that has been found appealing comes from thinking of the more inclusive domains on the model of possible worlds. In some cases, the logic of domain expansion can be thereby subsumed under standard modal logic. This idea will be relevant below.

Now that we have identified the assumptions of the paradox, a solution will begin by rejecting one or more of these assumptions. And indeed, the natural solutions of the paradox that have been proposed are cleanly classified by which assumption they reject.[9]

§3. Options. There is a rich set of available responses to the paradox. In this section, we introduce and discuss some of them. In the next section, we focus on our preferred response, which is inspired by Russell's no class theory and, more recently, by work of Harold Hodes and Kevin Klement. After that, we defend this 'eliminative' approach from a number of potential objections. We think that the existence of such a rich range of options to deal with the paradox is of great importance for a proper philosophical assessment of its solutions.

3.1. Restricting second-order comprehension. We have noted that at several points in the reasoning of the paradox we appeal to instances of the axiom scheme of comprehension. Perhaps these instances of comprehension are the crucial assumptions that drive the paradox. Russell himself observed that this is one of the conclusions that may be drawn from the paradox. He suggested two possible criteria for problematic comprehension instances. According to the first, an instance is to be rejected if it determines a property or relation whose size is *too big*. Since comprehension on the relation $<$ leads to paradox, one infers that this relation is too big to be comprehended. Since an instance of comprehension introducing $<$ is essential to the argument above, the paradox is averted (Russell, 1905, 43–44). The paradox may simply be regarded as a proof that $<$ is too big to be comprehended.

Russell, however, forcefully rejected the limitation of size approach.

> A great difficulty of [limitation of size] is that it does not tell us how far up the series of ordinals it is legitimate to go. It might happen that ω was already illegitimate: in that case all proper classes would be finite... Or it might happen that ω^2 was illegitimate, or ω^ω or ω_1 or any other ordinal

[9] There are points of similarity between our classification and that of Shapiro, 2007. This gives us some confidence that we are indeed cutting logical space at the joints.

> having no immediate predecessor. We need further axioms before we can tell where the series begins to be illegitimate. . . (Russell, 1905, 44)

Russell's point seems to be that one can say very little about what kinds of relations are "too big"—many different size restrictions seem to suffice to block the paradox, and there is no principled way of making a choice between them.[10]

One might try to avoid this objection by implementing the limitation of size approach in something like the following way. Perhaps a property or relation is "too big" if its extension or field, respectively, is too big to form a set. Since set theory is supposed to be an absolutely general theory of collections (individuated extensionally), one might think that it would be a good guide to what sizes of collections are possible. This, of course, presupposes a background set theory. Formally speaking, such a presupposition is unproblematic. However, from a philosophical perspective it is unsatisfactory, for the reasons we have discussed above. On this approach, the ordinals become just shadows of their set-theoretic surrogates. But the paradox of Burali–Forti, as we insisted, is first and foremost a problem about concept formation, not about sets.

More sophisticated developments of the limitation of size approach that address these difficulties can be found in Linnebo (2010), Studd (2016), and Florio & Linnebo (in progress). The idea is to develop a notion of definiteness and to hold that the axiom scheme of comprehenion must be curtailed when dealing with indefinite domains.

According to the second criterion proposed by Russell, an instance of comprehension is to be rejected if the comprehending formula is "complicated and recondite" (Russell, 1905, 38). Russell dubs this approach to the paradox the *zig-zag theory*, and adopts *zigzaginess* to denote the feature of formulas that renders them ineligible for comprehension. His original exposition leaves open just how zigzaginess is to be defined. One proposal for spelling out the precise nature of zigzaginess was given by Poincaré. He writes:

> [T]he definitions which should be regarded as not predicative [read: ineligible for comprehension] are those which contain a vicious circle. [...] Is this what Russell calls 'zigzaginess'? (Poincaré, 1912, 534)

By 'vicious circle' Poincaré means an appeal to the very notion being defined. Poincaré's example is the paradoxical class of Richard: the set of all the decimal numbers definable by a finite number of words. Since this set is countable, a new decimal not contained in the set can easily be defined in a finite number of words. Thus accepting Richard's paradoxical class leads us to paradox. The problem with the phrase that introduced Richard's class is that its denotation depends on which decimal numbers are definable. However, this in turn depends on what classes are denoted by phrases of English. Thus the phrase that introduces Richard's class depends for its denotation upon what classes are denoted by phrases of English, including itself. Until a denotation for 'Richard's class' is given, a denotation for 'Richard's class' cannot be determined. And this is a case of vicious circularity.

The vicious circle in the phrase defining Richard's class derives from the fact that the phrase contains an implicit quantifier over definitions. The solution is to ban classes introduced by phrases like this. In second-order logic, instances of comprehension can be

[10] This difficulty should be familiar to the neologicist, as it arises in connection with abstraction principles modeled on Boolos's New V. Although the formalism associated with New V puts the "too big" bottleneck further downstream than Russell proposed (by allowing that we might comprehend big concepts, but denying that abstraction operators can be applied to them in the usual way), it faces problems of arbitrariness parallel to Russell's original challenge. See the literature on the bad company problem—in particular Shapiro & Weir, 1999.

regarded as definitions of properties or relations. Hence, in this context, Poincaré's idea is usually rendered as a ban on impredicative instances of comprehension.

Allen Hazen (1986) has advocated a modified form of Poincaré's response, where one also allows *ramified* instances of comprehension—roughly one ramifies by distinguishing between different "levels" of variables and allowing comprehension schemes to introduce relations at a certain level of this hierarchy if they contain only quantifiers ranging over relations at lower levels. This generalizes a predicative approach, since a predicative approach amounts to a ramified approach that includes only level zero variables. Hazen remarks that, owing to its vicious impredicativity, $<$ cannot be defined in the ramified setting. He speculates that, as a result, the paradox can be avoided.

A predicative or ramified restriction of comprehension blocks the paradox of Burali–Forti at more than one point. As noted above, the definition of the less-than relation on ordinals ($<$) is impredicative. Of course, blocking one derivation of an inconsistency does little to guarantee that there is no derivation of an inconsistency to be found. One might, therefore, ask what can be shown in general about the safety of a predicative theory of ordinal abstraction.

The predicative approach to ensuring consistency has recently been the subject of a great deal of research in connection with neologicism. The basic results indicate that both the predicative and ramified fragment of Frege's system in *Grundgesetze* are consistent and interpret Robinson arithmetic (Heck, 1996). Moreover, given Frege's definition of number, the resulting system proves Hume's principle (Heck, 1996).[11] Ferreira and Wehmeier (2002) have shown that consistency is preserved even if we adjoin Δ^1_1-comprehension; and furthermore, Ferreira (2005) has recently shown that by adjoining an axiom of reducibility for finite concepts to the ramified predicative fragment of Frege's system, we can produce a theory which relatively interprets ACA_0 (the predicative second-order theory extending full elementary Peano arithmetic). Most recently, work by Sean Walsh on the connection between predicative abstraction and Gödel's constructible universe has yielded an extension of the theorem by Ferreira and Wehmeier: in the presence of full second-order comprehension for "pure" formulas of second-order logic together with Δ^1_1-comprehension for formulas containing nonlogical vocabulary, an arbitrary collection of abstraction principles is consistent so long as these principle are based on formulas that provably express equivalence relations on concepts.

Walsh's result establishes that a nontrivial strengthening of the predicative theory of ordinal abstraction is consistent. On the one hand, this result offers an insight. On the one hand, it presents a challenge. The insight is this: since the impredicative uses of comprehension occurring early in the paradox of Burali–Forti are pure (and therefore, by Walsh's result, are jointly consistent with O-AB), the locus of inconsistency can be seen to lie with the viciously impredicative definition of the relation $<$. If a traditional predicativist wishes, however, to take advantage of this insight and to endorse pure forms of impredicative comprehension, then they must provide a clarification of the concept of *vicious* impredicativity which motivates the different treatment of pure and impure instances of impredicative second-order comprehension.

Without such a clarification, the traditional predicativist, to avoid arbitrariness, has no choice but to renounce impredicative comprehension altogether, thus sacrificing standard theorems about well-orderings that rely on impredicative but pure comprehension instances

11 However, it does not *prove* the axioms of Robinson arithmetic with respect to Frege's definition of addition and succession of cardinal numbers (Linnebo, 2004). The interpretation of Robinson arithmetic is nonstandard for these definitions.

(see footnote 8). This should perhaps not come as a surprise. Historically, the main objection to predicativism is that it must abandon standard theorems of classical mathematics. Those who regard this sacrifice as significant have good reason to look elsewhere for a solution to the paradox of Burali–Forti.

3.2. Rejecting the first principle of ordinals. The paradox may also be blocked by giving up the first principle of ordinals, namely the principle that every well-ordering R is represented by an ordinal determined entirely by the order type of R. Let us explore how this could be done.

One could retain the principle that all well-orderings have representatives but deny that they are *entirely* determined by the order-type of the associated well-orderings. While allowing that nonisomorphic well-orderings have distinct representatives, it could be maintained that isomorphic well-orderings need not have the same ordinal as representative. It could even be allowed, perhaps, that fixing a unique well-ordering fails to determine a single ordinal. Unfortunately, this approach does not get off the ground in the presence of an appropriate principle of choice.[12]

A better option for rejecting the first principle of ordinals is to give up the assumption that *all* well-orderings have a representative. On this view, we retain second-order comprehension but exempt certain well-orderings from having a representative. If a background set theory is available, one could exempt well-orderings that are too big to form a set. Alternatively, one could take inspiration from Cantor's remarks on inconsistent multiplicities or from the literature on neologicism and allow only abstracta of properties that are not too big. For Cantor, a property is too big if it has a subproperty of the same size as the ordinals—since the ordinals are an 'inconsistent multiplicity' in his terminology. For the neologicist, a property is too big if it is equinumerous with the universe. The resulting restriction of the first-principle of ordinals has been dubbed *Size-Restricted Ordinal Abstraction Principle* in Cook (2003).

The high cost of this line of response to the paradox lies in the fact that the first principle of ordinals expresses the characterizing property of an ordinal number and thus it appears to be central to our understanding of ordinals. If one wishes to avoid doing violence to the concept of ordinal, one should not abandon this principle lightly.

3.3. Rejecting first-order absolute generality. A useful way to think of the role of first-order absolute generality in the paradox is as a principle guaranteeing a uniform reading of the first-order quantifiers. As noted above, the proof makes use of classical inference rules for the quantifiers, which fail if different quantifiers are assigned different ranges. If absolutely general quantification is possible, then the quantifiers in the paradoxical reasoning can be read as absolutely general and therefore uniform. In particular, such a reading of the quantifiers prevents a solution to the paradox which postulates a domain expansion.

Some have concluded that real lesson of Burali–Forti's paradox is that absolute generality is not possible. Geoffrey Hellman writes:

> What has emerged, however, is the point that we have a choice: either we stick with the above instance of 'absolute generality' and give up on desideratum (3) [i.e., the first principle of ordinals], or we seek to enforce

[12] If there is a well-ordering of the objects, we could recover the inconsistent principle of ordinal abstraction by defining the ordinal associated with a well-ordering to be the least object representing any well-ordering of the same order-type.

> the latter but deny [...] that it makes sense to refer to 'absolutely all ordinals', or 'absolutely all well-order relations'. (Hellman, 2011, 633)

So Hellman maintains that if we are to uphold the first principle of ordinals, then we must give up absolute generality. How does Hellman's rejection of absolute generality translate into a solution of the paradox?

Simply denying absolute generality is not enough. Absolute generality together with the other principles identified above suffice for the contradiction. But slightly weaker principles might suffice as well, provided that those principles license a uniform reading of the quantifiers. For a satisfying diagnosis of the paradox, one needs to explain where and why we have a variation in the reading of quantifiers figuring in the paradox. Hellman responds to this challenge by postulating a domain expansion captured in modal terms.

He states the first principle of ordinals in roughly the same way as we do: "any well-ordering, as a relation, should be represented by a unique ordinal" (Hellman, 2011, 632). However, he makes it clear that the principle needs to be read with care. In particular, the quantifiers should be read as modalized and, as he emphasizes, the second-order variables should be interpreted plurally.

The modalized version of first principle of ordinals becomes that principle that, necessarily, for any two well-orderings, there could be an ordinal determined entirely by the order type of those well-orderings. A formal rendering of this principle might go as follows:

$$\Box\forall R_1\forall R_2\,\Diamond(\exists y(y=\mathbf{ord}(R_1)\ \wedge\ y=\mathbf{ord}(R_2))\ \leftrightarrow\ R_1\cong R_2). \qquad (\textsf{O-AB}^{\Diamond})$$

The exact details of its proper representation will depend on the system of modal logic one favors for reasoning about domain expansions. But the above formalization does at least clarify Hellman's insistence on a plural reading of the variables. For Hellman, a well-ordering is just some ordered-pairs with the relevant properties. This has the effect of ensuring an extensional reading to the second-order variables. Pluralities, unlike properties, are thought to have their members necessarily.

Hellman needs to ensure an extensional reading of the second-order variables. An intensional reading would reinstate the paradox in the expanded domain. If R_1 and R_2 could be instantiated with the relation $<$ on the ordinals, intensionally construed, then $\textsf{O-AB}^{\Diamond}$ would entail that, in the expanded domain, the relation $<$ has an ordinal. Once it is admitted, in this way, that the well-ordering of the ordinals has an ordinal, the rest of the reasoning of the paradox is familiar.

Note that, even with respect to the specific reading of the paradox at hand, there are—perhaps inevitably—sacrifices that Hellman is forced to make. He can retain the first principle of ordinals only in the modalized sense, not in the original one. Of course, Hellman could respond by suggesting that the modalized sense is a natural one and the first principle should have been read this way all along. However, if this is true, then one wonders why a modalized reading should not be given to the other assumptions of the paradox. Specifically, Hellman cannot allow a modal reading of second-order comprehension. Such a reading isn't plausible given a plural interpretation of the second-order variables.[13] This might come as relief to Hellman, since a modalized version of second-order comprehension

[13] Consider the following modalized instance of comprhension

$$\Diamond\exists X\ \Box\forall x(Xx\ \leftrightarrow x=x).$$

While it acceptable on the intensional reading—it asserts that there could be a property tracking self-identity across worlds—it is not on the extensional reading. On this reading, it asserts that

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

in combination with O-AB$^\Diamond$ allows one to reconstruct the paradox. But it does commit Hellman to accepting that, in spite of similarity of syntax and usage, the meanings of the quantifiers in the first principle of ordinals and in second-order comprehension are radically different. Hellman might, again, reply that given his plural reading of the second-order variables, one should *expect* modalized comprehension to fail. Since the modalized reading of second-order comprehension is obviously false, the nonmodal reading dominates.

So, on Hellman's reading, we compromise in various ways on the first principle of ordinals and on second-order comprehension. Hellman could argue that, at least in the context of his particular reading, these compromises are not significant. But this raises a more fundamental problem. If we grant that Hellman has established the consistency of *a* reading of the assumptions of the paradox, this does not provide a resolution to the inconsistency found in the original practice of concept formation described above. In other words, we may wonder whether the move to a plural reading of the second-order variables and a modalized reading of the quantifiers re-directs our attention to one manifestation of the paradox while failing to get to the heart of the problem. *Ceteris paribus*, a solution to the paradox that does not require a special reading of the quantifiers and variables would be preferable.

In any case, there is one additional difficulty for Hellman's approach. It is unnecessarily concessive: it gives up more than is required. Recall the distinction between first-order and second-order absolute generality. First-order absolute generality secures the possibility of quantifying over absolutely all ordinals, while second-order absolute generality secures the possibility of quantifying over the plurality of all ordinals. Hellman rejects first-order absolute generality, since he denies that we can quantify over all ordinals. This is what lies behind his reformulation of the first principle of ordinals in terms of domain expansion. Since Hellman construes second-order variables as pluralities, he also rejects second-order absolute generality. Given that pluralities depend on their members, if we cannot quantify over all ordinals, we cannot quantify over the plurality of all ordinals.

As we will show in the next section, so long as one abandons second-order absolute generality, one can in fact retain first-order absolute generality along with the other assumptions of the paradox (see Shapiro, 2003; Shapiro, 2007). In this sense Hellman's approach gives up more than is required. Abandoning second-order absolute generality is what does the work. *Ceteris paribus*, an approach that avoids this formally unnecessary concession would be preferable.

3.4. Rejecting second-order absolute generality. An alternative, less concessive response to the paradox uses domain expansions, but only relative to the second-order domain. The intuitive idea is that, while the first-order domain is absolute general, the second-order domain does not include absolutely all second-order entities. Instead, we have an increasingly inclusive series of second-order domains accessible by domain expanding quantifiers (for alternative implementations of this line of response to the paradox, see the 'thin straw' of Shapiro, 2007). Following Hellman, we construe the second-order variables extensionally—we assume that second-order entities have their members necessarily.

It is useful to introduce this new approach by means of a model-theoretic comparison with Hellman's position. A natural way to model Hellman's modal axioms is to read the modal operators in terms of possible worlds and to take these worlds to be the logical-mathematical possibilities corresponding to the various V_κ of the iterative hierarchy of sets. The accessibility relation is given by this rule: V_α accesses V_β if and only if V_α is a

---

there could be some things encompassing all possible self-identical objects. This is in tension with Hellman's principle that there could always be more ordinals.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

subset of V_β. We should emphasize that this is only a toy model of Hellman's axioms. Not every feature of the model is part of Hellman's account. In particular, though in the model ordinals are identified with Von Neumann ordinals, Hellman need not endorse this identification.

Hellman's axioms come out true in any world in this model. As an illustration, consider V_ω, the hereditarily finite sets. This would correspond to a possibility in which the first-order domain is V_ω and the second-order domain is $V_{\omega+1}$, i.e., the set of all subsets of V_ω. In this context, the first principle of ordinals amounts to the claim that, for any well-ordering R in $V_{\omega+1}$, there is a V_κ containing a Von Neumann ordinal isomorphic to R. Given the axiom of replacement, it is easy to verify that this claim is true. Therefore, the first principle of ordinals is validated.

On this interpretation, Hellman's rejection of modalized second-order comprehension amounts to the requirement that, at a world w, the second-order quantifiers range only over subsets of the V_κ corresponding to w. The restriction on the first-order quantifiers amounts to the requirement that, at a world w, the first-order quantifiers range only over elements of the V_κ corresponding to w. Given this constraint on the quantifiers, second-order comprehension is true at each world.

One could modify this interpretation to "open up" the first-order quantifiers, letting them range over the entire iterative hierarchy, while keeping the restriction on the second-order quantifiers. The result would be another model in which the first-order domain of every logico-mathematical possibility is absolutely general: it is the set-theoretic universe V. The second-order domain of each possibility, however, is not absolutely general. In particular, at the world corresponding to V_κ, the second-order domain consists only of the subsets of V_κ. The picture is the following: while the domain of objects remains the same, the domain of concepts expands as we move from one logico-mathematical possibility to another. Importantly, the first principle of ordinals is retained, since it is true at every world that, for each well-ordering in the second-order domain, there is a corresponding ordinal in the absolutely general first-order domain.

It might be thought that the fact that this approach does not yield the standard comprehension scheme is a disadvantage. However, the semantic reason for the failure of standard comprehension is not that we don't have all the second-order objects we would otherwise have. It's that we have *more* first-order objects—we have absolutely all the sets in V. By owning up to these, a weakness that already existed in the modal comprehension scheme is made evident. A way to recover some comprehension would be to enrich the expressive resources of the object language. For example, if one adds the set-theoretic membership relation to the object language, then the following scheme of comprehension is true at every world:

$$\forall \kappa \, \Diamond \exists P \, \forall x \, (P(x) \leftrightarrow \varphi(x) \, \wedge \, x \in V_\kappa),$$

where κ ranges over any cardinal (or ordinal) set theoretically defined, and φ does not contain any free occurrences of P. Of course, this has the disadvantage of depending on the assumption that the structure of the underlying first-order domain is isomorphic to the iterative hierarchy.

This difficulty can be overcome by revising the approach, for example, in the following way. Instead of letting the second-order domain of each world be the subsets of some V_κ, we could let it be all subsets of cardinality less than κ. This yields a purely logical form of comprehension. Note that, in first-order logic, one can write down a formula $\exists_n x \; \varphi(x)$ expressing the fact that there are n things falling under the formula $\varphi(x)$, for any finite n. Pure second-order logic affords us the resources to express more cardinality facts of this

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

kind. For instance, we one can write down a formula $\exists_\kappa x\ \varphi(x)$ expressing the fact that there are κ-many things falling under the formula $\varphi(x)$, where κ is $\aleph_n$ for any finite n. The comprehension scheme now available is the following:

$$\Diamond(\exists_\kappa x\ \varphi(x)\ \rightarrow\ \exists P\ \forall x\ (P(x)\ \leftrightarrow \varphi(x)))$$

where κ and $\varphi(x)$ are such that $\exists_\kappa x\ \varphi(x)$ is expressible in the way described just above. This amounts, in effect, to a version of the limitation of size approach originally suggested by Russell (see §3.1). It has the advantage of avoiding any dependence on assumptions about the structure of the underlying first-order domain. Furthermore, it improves on the limitation of size approach by avoiding the arbitrary choice of size that worried Russell (see again §3.1).

The two proposals just sketched are not the only ways to deny second-order absolute generality while retaining first-order absolute generality. To emphasize the key feature of this style of solution, we call it the V-V_κ approach.

The following chart provides a summary of the responses to the paradox surveyed so far.

	Full second-order Comprehension	First principle of ordinals	Second principle of ordinals	First-order AG	Second-order AG
Restricting comprehension	✗	✓	✓	✓	?
Rejecting the first principle of ordinals	✓	✗	✓	✓	✓
Hellman's modal approach	?	✗	✓	✗	✗
V-V_κ	✗	✓	✓	✓	✗

We have marked some entries with a question mark to indicate that it is contentious whether the assumption is upheld. The first question mark arises because one might insist that restricting comprehension restores consistency by preventing us from talking about properties that do exist. The second question mark indicates that Hellman loses intensional forms of comprehension. Since he retains extensional forms of intra-world comprehension, we do not count this as abandoning comprehension altogether.

The common feature of these responses is that they all uphold the second principle of ordinals, namely the principle that ordinals are objects. We now want to explore a neglected response that denies this principle.

§4. An alternative approach. The approach we want to explore rejects the principle that ordinals are objects. The key idea is that, while similar on the surface to talk about objects, talk about ordinals does not directly make assertions about objects. Rather, talk of ordinals conveys higher-order logical content in a sense to be made precise below. Three sources of inspiration for this approach to the paradox are Russell's no class theory and, more recently, Harold Hodes's views on the content of arithmetical discourse (Hodes, 1986), and Kevin Klement's work on arithmetic and abstraction principles (Klement, in press).

The function of Russell's no class theory is to eliminate apparent reference to classes by interpreting such reference in terms of a higher-order language. This is how Russell himself describes the aim of the theory:

> The theory of classes which I set forth in my *Principles* was avowedly unsatisfactory. I did not, at that time, see any way of stating the elementary propositions of Arithmetic without employing the notion of "class". I have, however, since that time discovered that it is possible to give an

> interpretation to all propositions which verbally employ classes, without assuming that there really are such things as classes at all. [...] That it is meaningless [...] to regard a class as being or not being a member of itself, must be assumed for the avoidance of a more mathematical contradiction; but I cannot see that this could be meaningless if there were such things as classes. [...] The general contention that classes are a mere *façon de parler* has, of course, been often advanced, but it has not been accompanied by an exact account of what this manner of speaking really means, or by an interpretation of arithmetic in accordance with this contention [...]. (Russell, 1910, 376)

Here Russell highlights the indispensability of the no class theory for the resolution of the paradox that bears his name. As early as 1905, he suggested that the no class theory could also be deployed against the paradox of Burali–Forti (Russell, 1905).

Russell's key idea is to reinterpret propositions that verbally employ classes. Frege, likewise, seem to have arrived at the conclusion that some reconstruction of assertions that superficially employ the number concept would be the key to understanding mathematical language. Less than two years before his death, Frege wrote in a diary:

> From our earliest education onwards we are so accustomed to using the word 'number' and the number-words that we do not regard our way of using them as something that stands in need of a justification. To the mathematicians it appears beneath their dignity to concern themselves with such childish matters. But one finds amongst them the most different and contradictory statements about number and numbers. Indeed, when one has been occupied with these questions for a long time, one comes to suspect that our way of using language is misleading, that number-words are not proper names of objects at all and words like 'number', 'square number' and the rest are not concept-words; and that consequently a sentence like 'Four is a square number' simply does not express that an object is subsumed under a concept, and so just cannot be construed like the sentence 'Sirus is a fixed star'. But how then is it to be construed? (diary entry dated March 24, 1924; see Frege, 1979, 263)

Hodes calls the last line of this passage Frege's 'deathbed question'. He answers it in the spirit of Russell's no class theory by showing how to construe arithmetical discourse as a way of coding talk about a certain type of higher-order quantifier. On this view, arithmetical discourse becomes

> a special sort of fictional discourse: numbers are fictions "created" with a special purpose, to encode numerical object-quantifiers and thereby enable us to "pull down" a fragment of third-order logic, dressing it in first-order clothing. (Hodes, 1986, 144)

The usefulness of this discourse lies in the fact that first-order logic is familiar and tractable, while third-order logic is "notationally messy and logically complex" (Hodes, 1986, 144).

Klement (in press) recognizes that something in the vein of the no class theory can be used to eliminate reference to entities introduced via abstraction principles. He proposes a general method to regard talk of abstracta as mere abbreviation. In particular, Klement shows that Hume's principle and each axiom of second-order Peano Arithmetic, with the exception of the successor axiom, can be regarded as abbreviations for validities of third-order logic.

We would like to take the no class approach back to its original Russellian application to the resolution of the paradoxes. In particular, we would like to provide an interpretation of talk of ordinals without assuming, to put it in Russell's terms, that there are such things as ordinals. On this interpretation, assertions about ordinals become ways of conveying higher-order logical content, just as for Hodes assertions about natural numbers encode higher-order information about numerical object-quantifiers. Our decoding of ordinal talk involves construing ordinals as given by abstraction principles, and then showing how to reduce ordinal abstraction and its consequences to validities of third-order logic as Klement does to Hume's Principle. Unlike Hodes and Klement, however, we face a potentially inconsistent mathematical discourse. Here the no class theory serves not only to clarify the function of the discourse in question but also to witness its consistency.

4.1. The no class conception of ordinals. The no class approach denies the second principle of ordinals. The task for the no class approach is—to paraphrase Russell—to give an interpretation to all propositions which verbally employ ordinals, without assuming that there really are such things as ordinals at all. How might this be accomplished?

Concretely, our challenge is to interpret the ordinal abstraction operator and, eventually, quantification over ordinals. It is useful to begin with the case of identity statements between ordinals presented by the abstraction operator, e.g., $\mathbf{ord}(R_1) = \mathbf{ord}(R_2)$. If we are interpreting the meaning of the abstraction operator in isolation, one idea might be to interpret each $\mathbf{ord}(R)$ as simply referring to R. However, this option immediately wreaks havoc on the intuitive properties of the ordinals. The standard progression of the ordinals, characterized by $<$, would no longer be a well-ordering. And, what's worse, we lose the right-to-left direction of O-AB.

A natural way to amend this proposal would be as follows. One first selects a canonical well-ordering for each isomorphism class. Then one interprets each $\mathbf{ord}(R)$ as simply referring to the canonical well-ordering of R's isomorphism class. This enables us to recover O-AB. But the proposal is hardly satisfactory from either a technical or a philosophical perspective. From a technical perspective, defining $\mathbf{ord}(R)$ along these lines within second-order logic would require a strong principle of choice going beyond even the existence of a global well-ordering. It would indeed require a definable third-order well-ordering of the second-order domain, or something similar. From a philosophical perspective, the problem is, if anything, more acute. This interpretation of $\mathbf{ord}(R)$ requires that there be a determinate well-ordering that the expression picks out. But there is nothing about the usage of ordinal language that might fix such a reference.[14]

However, it is possible to circumvent the problem with the second proposal by taking a page from Russell's playbook. We can let the isomorphism class of R determine the meaning of $\mathbf{ord}(R)$ if, rather than analyzing the meaning of $\mathbf{ord}(R)$ in isolation, we analyze its meaning in the context of a proposition. Instead of looking for $\mathbf{ord}(R)$'s referent we look for a paraphrase of statements in which $\mathbf{ord}(R)$ occurs. We want an identity statement between two ordinals to be true if the corresponding well-orderings are isomorphic. So, rather than assigning a canonical well-ordering to $\mathbf{ord}(R_1)$ and $\mathbf{ord}(R_2)$, we analyze an identity statement between two ordinals, for example $\mathbf{ord}(R_1) = \mathbf{ord}(R_2)$, as simply

[14] There is clearly nothing about *actual* usage that might fix reference, nor is there anything about *possible* usage that might fix reference. Someone may object that we could select a choice function and determine reference with respect to it. However, the selection of an appropriate choice function is more easily proposed than accomplished. For it faces the same problem it purports to solve. Even if we accept that appropriate choice functions exist, we have no means to pick out a particular one.

expressing that there exist well-orderings V_1 and V_2, isomorphic to R_1 and R_2 respectively, such that V_1 is identical to V_2. Identity between second-order entities is cashed out, as usual, in third-order terms.[15]

Next, we analyze atomic predication of the form $P(\mathbf{ord}(R))$, where R is a well-ordering and P is an atomic predicate applicable to ordinals. Recall that our current strategy is to deny the second principle of ordinals. So we do not assume that $\mathbf{ord}(R)$ is a first-order entity, and thus we do not assume that P expresses a property of first-order entities. This opens the way to an analysis parallel to the one articulated in the previous paragraph. Just as above, we analyze $P(\mathbf{ord}(R))$ as expressing that there is a well-ordering V isomorphic to R such that $\mathbf{P}(V)$, where $\mathbf{P}$ is a higher-order counterpart of P.

Putting these ideas together, given well-orderings R_1 and R_2, we might analyze the following statement

$$\mathbf{ord}(R_1) = \mathbf{ord}(R_2) \ \wedge \ P(\mathbf{ord}(R_1))$$

as

$$\exists V_1 \, \exists V_2 \, (V_1 \cong R_1 \wedge V_2 \cong R_2 \wedge V_1 = V_2) \ \wedge \ \exists V_3 \, (\, V_3 \cong R_1 \ \wedge \ \mathbf{P}(V_3)),$$

where the bound variables range over well-orderings.

An analysis like this can be extended to the full vocabulary of the language of ordinals. We regiment this language in line with our denial of the second principle of ordinals. Since ordinals are not objects, we reserve a second sort for them, distinct from the object sort. To these basic sorts we add the standard vocabulary of second-order logic. We call the resulting language L_Ω. Our analysis provides a translation procedure, which we denote by '$*$', mapping each formula φ of L_Ω to a formula $[\varphi]^*$ of the language L_3 of third-order logic.[16] Crucially, the intuitive properties of ordinals reduce under this translation to pure theorems of third-order logic. Among these properties are O-AB and the structural properties of the progression of ordinals. In this sense, we vindicate the idea that talk of ordinals serves to convey higher-order logical content.[17]

In the next section, we give the details of these formal languages and of the translation procedure. This discussion is rather technical. Readers more interested in the philosophical implications of our findings may skip ahead to §4.3.

4.2. Implementation of the No-Class conception. We now take up the remaining task of explaining how to extend our analysis to the entire language. We begin by introducing the language L_Ω. Then we provide the full translation procedure reducing this language to the pure language of third-order logic. After that, we present an example of a formal result that witnesses the viability of our no class approach to ordinals.

Let L_2 be the standard language of second-order logic, regarded as a multi-sorted language with object sort *obj* and, for every n, a sort S_n for n-ary relations between objects. The vocabulary of the language is as follows.

---

15 Note that the current proposal amounts essentially to analyzing $\mathbf{ord}(R_1) = \mathbf{ord}(R_2)$ as expressing that R_1 and R_2 are isomorphic well-orderings. The particular implementation we adopt has the advantage of generality, as we will see in a moment.

16 Here we diverge from Klement's approach. We do not take ordinal talk to be an abbreviation. Rather, our translation procedure shows how ordinal talk as a *sui generis* language could be used to reason about and convey higher-order logical content.

17 The reader may wonder what framework we would use to cash out this notion of conveying content. Several are available, ranging from literal interpretation to rational reconstruction. We would prefer to remain neutral. Much of what we say can be embraced by advocates of any of these perspectives.

(i) Infinitely many variables $x_1, x_2 \ldots$ of sort *obj*.
(ii) Infinitely many variables $P^n_1, P^n_2, P^n_3 \ldots$ of sort S_n.
(iii) Standard logical symbols, including an identity sign of sort S_2 between two variables of sort *obj*, and parentheses.

The terms of L_2 are simply the variables above. The atomic formulas consist of a symbol of sort S_n—either a variable, or, if $n = 2$, the identity sign—followed by a sequence of n variables of sort *obj*.[18] Complex formulas are then built in the usual way using logical connectives and parentheses.

Let L_Ω be an expansion of L_2 with the these new sorts: a sort for ordinals *ord* and, for every n, a sort T_n for n-ary relations over ordinals. In addition to the vocabulary of L_2, we have new symbols as follows.

(iv) Infinitely many variables $\alpha_1, \alpha_2 \ldots$ of sort *ord*.
(v) Infinitely many variables $A^n_1, A^n_2, A^n_3 \ldots$ of sort T_n.
(vi) A function symbol **ord** taking binary relations of sort S_2 as arguments and returning a value of sort *ord* (so the sort of the term $\mathbf{ord}(P^2_1)$ is *ord*).
(vii) An identity symbol of sort T_2 between two terms of sort *ord*.

In addition to the terms of L_2 and the variables above, L_Ω contains terms formed by applying the symbol **ord** to a term of sort S_2. And in addition to the atomic formulas of L_2, the atomic formulas of L_Ω consist of those obtained by applying symbols of sort T_n to a sequence of n terms of sort *ord*.

Finally, we arrive at a language L_3 of third-order logic, obtained by enriching L_2 with sorts for relations between lower-level relations over objects. In particular, for every n and every sequence of relation sorts $S_{i_1}, \ldots, S_{i_n}$, we have a distinct sort $U_{i_1,\ldots,i_n}$. As for the vocabulary of L_2, we have new symbols as follows.

(viii) For every n and every sequence of sorts $S_{i_1}, \ldots, S_{i_n}$, infinitely many variables:

$$\mathbf{X}^{i_1,\ldots,i_n}_1, \mathbf{X}^{i_1,\ldots,i_n}_2, \mathbf{X}^{i_1,\ldots,i_n}_3 \ldots$$

of sort $U_{i_1,\ldots,i_n}$.
(ix) Fresh second-order variables of sort S_2: $R^2_1, R^2_2, R^2_3, \ldots$ and $V^2_1, V^2_2, V^2_3, \ldots$.

The terms of L_3 comprise the terms of L_2 together with the variables above. The atomic formulas of L_3 are those generated by the formation rules of L_2 together with those obtained by applying symbols of sort $U_{i_1,\ldots,i_n}$ to a sequence of n terms of sorts $S_{i_1}, \ldots, S_{i_n}$ respectively.

Strictly speaking **ord** only applies to the variables of sort S_2 in L_Ω. However, to make the translation procedure more uniform, we do allow **ord** to apply to the fresh variables $R^2_1, R^2_2, R^2_3, \ldots$ of L_3. We describe the result as a pseudoterm of sort *ord* of L_Ω. Likewise, we describe a formula containing such pseudoterms as a pseudoformula of L_Ω.

We are now ready to define a translation $*$ from L_Ω to L_3 showing how talk of ordinals can be eliminated. The translation is described by its action on atomic formulas and logical symbols. In order to define this translation, we first define two auxiliary translations $\bullet$ and $\dagger$.

[18] For simplicity, we do not include constants in the language. Moreover, we treat predication as part of the syntax, and do not make explicit the usual application relation that characterizes predication in multi-sorted languages. In any case, nothing important hinges on these choices.

The translation $\bullet$ maps terms of sort *ord* to terms and pseudoterms of sort *ord*.

$$\begin{aligned} \alpha_i &\overset{\bullet}{\longmapsto} \mathbf{ord}(R_i) \\ \mathbf{ord}(P_i^2) &\overset{\bullet}{\longmapsto} \mathbf{ord}(P_i^2) \end{aligned}$$

Next, we define a translation $\dagger$ from atomic formulas and pseudoformulas of L_Ω to formulas of L_3. This eliminates uses of the *ord* operator in favor of third-order vocabulary according to the strategy outlined in the previous section. We use ζ_i as a metalinguistic variable ranging over terms P_i^2 or R_i^2, and we understand identity between such terms in the usual third-order way. We use $\zeta_i \sim \zeta_j$ to abbreviate the pure second-order formula expressing that either ζ_i and ζ_j are isomorphic or neither ζ_i nor ζ_j is well-ordered. Intuitively, the second clause here is intended to ensure that all relations that are not well-ordered are assigned the same abstract.

$$\mathbf{ord}(\zeta_i) = \mathbf{ord}(\zeta_j) \overset{\dagger}{\longmapsto} \exists V_i^2 \exists V_j^2 (V_i^2 \sim \zeta_i \wedge V_j^2 \sim \zeta_j \wedge V_i^2 = V_j^2)$$

$$A_m^n(\mathbf{ord}(\zeta_{i_1}), \ldots, \mathbf{ord}(\zeta_{i_n})) \overset{\dagger}{\longmapsto} \exists V_{i_1}^2 \ldots \exists V_{i_n}^2 (V_{i_1}^2 \sim \zeta_{i_1} \wedge \cdots \wedge V_{i_n}^2 \sim \zeta_{i_n} \wedge \mathbf{X}_m^{\overbrace{2,\ldots,2}^{n\text{-times}}}(V_{i_1}^2, \ldots, V_{i_n}^2)).$$

Note that $\dagger$ essentially generalizes the transformation on atomic sentences described in the previous section. That is, a basic predication such as $A_m(\mathbf{ord}(R))$ is taken to express that there is a well-ordering V isomorphic to R such that $\mathbf{X}_m(V)$, where $\mathbf{X}_m$ is a third-order counterpart of A_m, as signaled by the matching subscripts.

Finally, we recursively define the main translation $*$, a translation from L_Ω to L_3. We use metavariables of the form t_i and x_i for terms of sort *ord* and *obj*, respectively. We use $[\varphi]^*$, $t_i^\bullet$, and $[\psi]^\dagger$ to denote the results of applying $*$, $\bullet$, and $\dagger$ to φ, t_i, and ψ, respectively.

$$\begin{aligned} P_m^n(x_{i_1}, \ldots, x_{i_n}) &\overset{*}{\longmapsto} P_m^n(x_{i_1}, \ldots, x_{i_n}) \\ x_i = x_j &\overset{*}{\longmapsto} x_i = x_j \\ \exists x_i \varphi &\overset{*}{\longmapsto} \exists x_i [\varphi]^* \\ \exists P_m^n \varphi &\overset{*}{\longmapsto} \exists P_m^n [\varphi]^* \\ \varphi \wedge \psi &\overset{*}{\longmapsto} [\varphi]^* \wedge [\psi]^* \\ \neg\varphi &\overset{*}{\longmapsto} \neg[\varphi]^* \\ A_m^n(t_{i_1}, \ldots, t_{i_n}) &\overset{*}{\longmapsto} [A_m^n(t_{i_1}^\bullet, \ldots, t_{i_n}^\bullet)]^\dagger \\ t_i = t_j &\overset{*}{\longmapsto} [t_i^\bullet = t_j^\bullet]^\dagger \\ \exists \alpha_i \varphi &\overset{*}{\longmapsto} \exists R_i [\varphi]^* \\ \exists A_m^n \varphi &\overset{*}{\longmapsto} \exists \mathbf{X}_m^{\overbrace{2,\ldots,2}^{n\text{-times}}} [\varphi]^*. \end{aligned}$$

The first six clauses are just identities. So these transformations leave the second-order formulas of L_2 untouched. The last four act as follows. Quantification over ordinals is replaced by quantification over relations of objects. Quantification over relations of ordinals is replaced by third-order quantification. The atomic formulas are subject to the transformations $\bullet$ and $\dagger$ described above.

Now that we have presented the translation, we will adopt the convention that unsuperscripted P_1, R_i, V_i are to be read as decorated by the superscript '2'. A similar convention applies to variables $\mathbf{X}_i$, where the omitted superscript is the appropriate sequence '2',...,'2'.

Let us now look at a significant example of how the translation can be applied. Recall the principle of ordinal abstraction:

$$\forall R_1 \forall R_2(\mathbf{ord}(R_1) = \mathbf{ord}(R_2) \;\leftrightarrow\; R_1 \cong R_2), \tag{O-AB}$$

where the quantifiers are implicitly restricted to well-ordering. Let $\mathrm{WO}(\zeta_i)$ abbreviate the pure second-order formula expressing that ζ_i is a well-ordering. We can now regiment O-AB in L_Ω:

$$\forall P_i\, \forall P_j\, (\mathrm{WO}(P_i) \wedge \mathrm{WO}(P_j) \rightarrow (\mathbf{ord}(P_i) = \mathbf{ord}(P_j) \leftrightarrow P_i \cong P_j)).$$

Consider the result of applying $*$ to this formula:

$$[\forall P_i\, \forall P_j\, (\mathrm{WO}(P_i) \wedge \mathrm{WO}(P_j) \rightarrow (\mathbf{ord}(P_i) = \mathbf{ord}(P_j) \leftrightarrow P_i \cong P_j))]^*. \tag{1}$$

Because $*$ has no effect on second-order vocabulary, we have:

$$\forall P_i\, \forall P_j\, (\mathrm{WO}(P_i) \wedge \mathrm{WO}(P_j) \rightarrow ([\mathbf{ord}(P_i) = \mathbf{ord}(P_j)]^* \leftrightarrow P_i \cong P_j)). \tag{2}$$

Let us focus, for the moment, on $[\mathbf{ord}(P_i) = \mathbf{ord}(P_j)]^*$. This is $[\mathbf{ord}(P_i)^\bullet = \mathbf{ord}(P_j)^\bullet]^\dagger$. Since $\bullet$ has no effect on $\mathbf{ord}(\zeta_i)$, we only need to consider the effect of $\dagger$, which gives us:

$$\exists V_i\, \exists V_j\, (V_i \sim P_i \wedge V_j \sim P_j \wedge V_i = V_j). \tag{3}$$

Substituting (3) into (2), we obtain:

$$\forall P_i\, \forall P_j(\mathrm{WO}(P_i) \wedge \mathrm{WO}(P_j) \rightarrow (\exists V_i\, \exists V_j\, (V_i \sim P_i \wedge V_j \sim P_j \wedge V_i = V_j) \leftrightarrow P_i \cong P_j)). \tag{4}$$

Note that, since P_i and P_j are restricted to well-orderings, $\sim$ in the consequent is equivalent to $\cong$. Thus (4) is equivalent to:

$$\forall P_i\, \forall P_j(\mathrm{WO}(P_i) \wedge \mathrm{WO}(P_j) \rightarrow (\exists V_i\, \exists V_j\, (V_i \cong P_i \wedge V_j \cong P_j \wedge V_i = V_j) \leftrightarrow P_i \cong P_j)). \tag{5}$$

It is easy to verify that (5) is a theorem of third-order logic (recall that identity between second-order variables is cashed out in third-order terms). This result is significant. We have just shown that the translation of O-AB under $*$ is a logical truth. This is a prime example of how talk of ordinals can be interpreted as conveying higher-order logical content.

4.3. The No-Class theory of the ordinals. The result that the regimentation of ordinal abstraction in L_Ω is carried by our translation to a validity of third-order logic is not isolated. Instead, we find a rich theory emerging from results of this kind. Our main contention is that this theory constitutes a robust conception of the ordinals.

Denote by *No-Class* the set of sentences of L_Ω whose translation under $*$ is a theorem of third-order logic.[19] Remarkably, No-Class turns out to be closed under standard logical rules concerning the connectives and the identity symbol. So it is not simply a set

[19] By 'third-order logic' we mean the theory in the language of L_3 axiomatized by the usual logical rules for logical symbols plus the full comprehension scheme for second-order and third-order variables.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

of sentences—it is a theory which sustains ordinary classical reasoning. And, moreover, this theory includes all of the key properties of ordinals. This is where its philosophical significance lies. We can regard the properties of ordinals as arising naturally from the translation scheme rather than as given by special-purpose axioms.

Specifically, No-Class includes O-AB, as noted above, as well as the full comprehension schema for both relations over objects and relations over ordinals. In this respect, it achieves something that none of its competitors manages to do: consistently combining O-AB and full comprehension. Verifying the consistency of No-Class is easy. There are sentences of L_Ω that do not belong to No-Class. Take any sentence φ of second-order logic that is not valid. Since the translation leaves the second-order formulas untouched, the translation of φ is φ. By definition, a sentence whose translation is not valid is not included in No-Class. Hence φ is not included in No-Class, which is therefore not a trivial theory. So No-Class witnesses the consistency of the theory of ordinals axiomatized by O-AB given the appropriate sortal restrictions characteristic of L_Ω.[20] We postpone the proofs of these claims to the appendix.

What other facts about ordinals are contained in No-Class? Since this theory contains each theorem of second-order logic, it contains all theorems about well-orderings provable in pure second-order logic. In particular, it includes the key lemma in the paradox of Burali–Forti: if every proper initial segment of one well-ordering R_1 is isomorphic to a proper initial segment of another R_2, then R_1 is isomorphic to an initial segment of R_2. Here No-Class seems to do better than a predicative theory of ordinals, as the latter most likely cannot prove this lemma (see footnote 8). Moreover, contra Burali–Forti, the ordinals can be shown to be linearly ordered. Furthermore, contra Russell (1903, 323), the ordinals can be shown to be well-founded. So No-Class contains Jourdain's theorem that the ordinals are well-ordered. A corollary of this fact is that No-Class supports transfinite induction (see Appendix C).

However, No-Class leaves open certain questions about ordinals. In particular, it leaves open how many ordinals there are. It only allows us to prove that there are two ordinals. The first ordinal, if we can call it that, is associated with relations that are not well-orderings. Neologicists would call it the *bad abstract* of the principle of ordinal abstraction. The second ordinal is associated with the empty relation, whose existence is provable in pure second-order logic.

The fact that the theory does not allow us prove that there are more than two ordinals is, we contend, unproblematic. This is in keeping with the No-Class conception of the ordinals according to which talk of ordinals is a vehicle to convey purely logical information. As usually conceived, third-order logic carries minimal existential information, hence we should not expect that, without further assumptions, talk of ordinals will be ontologically profligate. Instead, as we might hope, No-Class *does* carry significant *conditional* information about ordinals. For example, if there are three objects, then four ordinals can be shown to exist—counting the bad abstract as an ordinal. Assuming the standard semantics for second-order logic, it is easy to verify that if there are countably many objects, then there are uncountably many ordinals. So adding existential assumptions about objects to No-Class will imply the existence of more ordinals. The same effect can be obtained on the basis of an ontologically richer conception of logic. If one prefers nonstandard logical

[20] It is worth noting that this consistency proof does not require the construction of a model of the axioms as a witness to consistency—since the theory is deductively closed, we need only to inspect the translation in order to verify that No-Class is nontrivial.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

axioms implying substantial existential claims, one will be able to show that the set of sentences translating into theorems of third-order logic comprises more existential claims about ordinals.[21]

No-Class recovers the intuitive properties of ordinals. But how does it diagnose the paradox of Burali–Forti? The answer is that, while No-Class proves the main lemmas of the paradoxical reasoning, the proof of the contradiction breaks down in the final stages of the argument. One cannot prove in No-Class that for every ordinal α, α is the representative of $<_\alpha$, i.e., that $\alpha = \mathbf{ord}(<_\alpha)$. While the relation $<_\alpha$ can be defined and proven to be a well-ordering, $<_\alpha$ is no longer a well-ordering of *objects*. Thus $\mathbf{ord}(<_\alpha)$ is not well-formed, since the **ord** operator applies only to relations of objects.

§5. Objections. Our solution to the paradox of Burali–Forti turns on a type distinction. From the point of view of L_Ω, ordinals are not objects but *sui generis* entities conveying information about logical constructions. One might object that this treatment does violence to the concept of ordinal. Linguistically, we do not seem to distinguish between talk of objects and talk of ordinals. The same predicates can be applied to ordinary noun phrases and noun phrases referring to ordinals. This—the objection goes—is at least *prima facie* reason to believe that ordinals just are objects.

To this objection we have a number of replies. First of all, it is not clear that ordinals were always conceived this way. For instance, Zermelo seemed to have conceived them as *sui generis* entities in his influential 1930 paper. Second, syntactic intuitions are notoriously defeasible. For example, we often tend to nominalize expressions of various syntactic categories, ranging from adjectives to verbs and sentences, but this is consistent with the view that the best formalization of some of these expressions will assign them to a logical category other than the category of ordinary logical subjects. A case in point is talk of relations: even though we commonly refer to them by means of noun phrases, mere consistency often forces us to represent them as entities of higher types. This brings us to a third, general point: respecting syntactic intuitions opens the door to a host of paradoxes. While the objector might have in mind a different solution to the paradox of Burali–Forti, they will presumably have to violate naïve syntactic distinctions somewhere—if not in the solution of Burali–Forti's paradox, at least in the solution of paradoxes arising from conceiving of relations as objects (e.g., Russell, 1908, 222–223).

There is another objection in a similar vein. Perhaps one could complain that while our sortal distinctions are not conceptually problematic, they prevent us from considering order types that evidently can be considered. Because in our framework the **ord** operator cannot be applied to relations among ordinals, we cannot refer to the order type of such relations. But—the objection continues—one obviously can refer to such order types. In fact, we just did. Thus the proposal illegitimately restricts our ability to refer to ordinals.

This objection neglects one key feature of our proposal. We do not set out to describe a preexisting domain of ordinals. Rather, we aim to describe how ordinal talk emerges naturally from higher-order reasoning about a preexisting domain of objects. In principle, after ordinals have been introduced in this way, there is no obstacle to taking them to

[21] Strictly speaking, one could relativize No-Class to an arbitrary theory T in L_3. So No-Class would become the set of sentences in L_Ω that translate to consequences of T. On this approach, how many ordinals there are depends on the existential commitments of T. Our understanding of No-Class sets T equal to a standard set of axioms of third-order logic. We prefer this minimalistic approach, since it is striking and philosophically significant that third-order logic alone suffices to capture the core facts about the ordinals.

form a new domain of objects. New talk of ordinals would then emerge from higher-order reasoning about this new domain. So the initial restriction on the order types one can consider can be overcome by iterating our construction. Gödel makes essentially this point about Russell's no class theory:

> [T]he restrictions involved do not appear as ad hoc hypotheses for avoiding the paradoxes, but as unavoidable consequences of the thesis that classes, concepts, and quantified propositions do not exist as real objects. It is not as if the universe of things were divided into orders and then one were prohibited to speak of all orders; but, on the contrary, it is possible to speak of all existing things; only, classes and concepts are not among them; and if they are introduced as a *façon de parler*, it turns out that this very extension of the symbolism gives rise to the possibility of introducing them in a more comprehensive way, and so on indefinitely. (Gödel, 1944, 133)

This iterated construction thus leads to superordinals, super-superordinals, and so on. For example, we are free to introduce a new abstraction operator applying to relations over ordinals (our sort T_2 from §4.2). This move will yield superordinals that regiment talk about order types of ordinals. Further repetitions of this procedure allow us to regiment "iterated" ordinal talk, such as talk of order types of order types of ordinals.

The objector might, at this point, chime in that this sounds like a hierarchy of proper classes. But proper classes are often thought to signal a failure of absolute generality (see, e.g., Parsons, 1974b; Boolos, 1998). This objection might seem particularly serious, since the ability to capture absolute generality was an important virtue of the no class approach.

However, this version of the objection overlooks a disanalogy between our conception of ordinals and the conception of proper classes that threatens absolute generality. Proper classes are conceived as set-like objects lying outside the domain of quantification operative in ordinary set theory. But, as we noted above, we do not aim to describe a preexisting domain of ordinals. Rather, we interpret talk of ordinals as conveying logical information about a preexisting domain of objects. Quantification over objects remains absolutely general: we can talk about absolutely *all* the objects that exist. The possibility of introducing more ordinals is not a sign that there are more ordinals that lie beyond our initial domain of quantification. As Gödel observes, it is instead a sign that every extension of our symbolism gives rise to the possibility of introducing a more comprehensive symbolism. A better analogy is perhaps one between our conception of the ordinals and the interpretation of proper classes in terms of plural quantification, which is consistent with absolute generality (Uzquiano, 2003; Burgess, 2004). Just like the plural interpretation can serve to, for example, make sense of talk of proper classes, our approach could be put to use in making sense of talk of order-types longer that the order-type of the Von Neumann ordinals; talk of this kind is occasionally found in the practice of set theory (see Shapiro, 2003 as well as Shapiro & Wright, 2006 for discussion).

§6. Conclusion. We surveyed a number of options for dealing with the paradox of Burali–Forti conceived as a paradox about ordinal numbers rather than sets. As we emphasized, all available options sacrifice one or more of the intuitive principles behind the paradox. In the spirit of Russell's no class theory, our proposal rejects the principle that ordinals are objects. Rejecting this one principle allows us to provide a consistent account of talk of ordinals. The main features of our account are well motivated by a no class conception of the ordinals. On that conception, talk of ordinals can be seen as a vehicle for conveying third-order logical content. Our work shows that a translation procedure suggested by this conception yields a theory of ordinals capturing their key structural properties. Adding this theory into the mix, the situation appears to be the following:

	Full second-order Comprehension	First principle of ordinals	Second principle of ordinals	First-order AG	Second-order AG
Restricting comprehension	✗	✓	✓	✓	?
Rejecting the first principle of ordinals	✓	✗	✓	✓	✓
Hellman's modal approach	?	✗	✓	✗	✗
V-V_κ	✗	✓	✓	✓	✗
No-Class	✓	✓	✗	✓	✓

One might ask whether our translation procedure could provide a more general account of mathematical entities obtained by abstraction. We regard this as an open question. Much of the plausibility of our no class theory of the ordinals derives from the fact that the main theorems about ordinals can be recovered in L_Ω. The same might not be true for other kinds of abstracta. But, then again, it might be. Perhaps there is a general theorem showing that the main features of abstracta will be recoverable from a corresponding no class theory. Such a theorem would be a major contribution to the longstanding program of reducing mathematics to higher-order logic. The fact that this can be done for the ordinals gives grounds for cautious optimism.[22]

§7. Appendix A: Definitions. Here we provide some basic definitions for the higher-order languages introduced above. A *field* of a two-place relation R is any monadic property M such that for any x, $M(x)$ if and only if there is y such that $R(x, y)$ or $R(y, x)$. The relation R is a *well-ordering* if R is (i) irreflexive, (ii) transitive, and (iii) any subproperty P of a field of R has an R-least element.

An initial segment of a relation R_1 with respect to x is any relation R_2 such that, for any y and z, $R_2(y, z)$ if and only if $R_1(y, z)$, $R_1(y, z)$, and $R_1(z, x)$. We usually denote the restriction of R with respect to x as R_x. A relation R_2 is an initial segment of R_1 if there is x such that R_2 is an initial segment of R_1 with respect to x. An inverse R^{-1} of a relation R is defined standardly.

A relation F is a function from a monadic property P_1 to a monadic property P_2 if it meets the usual conditions. The notions of domain and range are then defined as usual. We adopt the standard abuse of notation in using $F(x)$ to denote the unique y such that $F(x, y)$, i.e., the image of x under F.

Two relations R_1 and R_2 are said to be isomorphic if there is a function F from a field of R_1 to a field of R_2 such that if $R_1(x, y)$, then $R_2(F(x), F(y))$, and if $R_2(x, y)$, then $R_1(F^{-1}(x), F^{-1}(y))$. We use $R_1 \cong_F R_2$ to indicate that F witnesses that R_1 is isomorphic to R_2. The notation $R_1 \cong R_2$ represents that $R_1 \cong_F R_2$ for some F.

[22] For helpful comments and discussion, we would like to thank Andrew Arana, Patricia Blanchette, Francesca Boccuni, Luca Incurvati, Nick Jones, Ansten Klev, Øystein Linnebo, Richard Pettigrew, Agustín Rayo, Ian Rumfitt, Andrea Sereni, Stewart Shapiro, Sean Walsh, and an anonymous reviewer. We also wish to extend heartfelt thanks to audiences at the following venues: Logica 2015, University of Bristol, University of Oslo, and FilMat 2016 at the University of Chieti-Pescara. Their feedback helped shape the development of this article.

§8. Appendix B: The paradox of Burali–Forti. We now outline the reasoning of the paradox. Most of the proof is straightforward but labor-intensive. So we only provide its most illuminating structural features. We work within a standard language of second-order logic, parallel to L_2, with one additional function symbol '**ord**' from elements of sort S_2 to elements of the object sort *obj*.

The first lemma guarantees that there is at most one way to extend an isomorphism defined on an initial segment of a well-ordering.

LEMMA 8.1. *Let R and R' be well-orderings. Suppose there are functions F and G such that for some x_1, y_1, x_2, y_2, $R_{x_1} \cong_F R'_{x_2}$ and $R_{y_1} \cong_G R'_{y_2}$. Then if $R(y_1, x_1)$, the function F contains the function G, i.e., for all z_1 and z_2, $G(z_1, z_2)$ only if $F(z_1, z_2)$.*

The second lemma, against the backdrop of the first, guarantees that there is at least one way to extend an isomorphism defined on an initial segment of a well-ordering. This is established via impredicative comprehension.

LEMMA 8.2. *Let R and R' be well-orderings. Suppose that for every y in the field of R, there is an initial segment W of R' and an isomorphism F such that $R_y \cong_F W$. Then R is either isomorphic to R' or to one of its initial segments.*

Let us define an ordinal to be any object x such that, for some binary relation R, $x = \mathbf{ord}(R)$. By second-order comprehension, there is a relation $<$ between ordinals characterized by the following condition: for every x and y, $x < y$ if and only if there are relations R_1 and R_2 such that $x = \mathbf{ord}(R_1)$ and $y = \mathbf{ord}(R_2)$ and either (i) R_1 and R_2 are well-orderings and R_1 is isomorphic to a proper initial segment of R_2, or (ii) R_2 is a well-ordering and R_1 is not.

Recall that the first principle of ordinals, which was formalized as the principle of ordinal abstraction O-AB, states that isomorphic well-orderings are always represented by the same ordinal and non–isomorphic well-orderings are always represented by different ordinals. The leaves open the behavior of the **ord** operator when applied to relations that are not well-ordered. For simplicity, we assume that relations that are not well-ordered are all represented by the same *bad abstract*. Thus two relations have the same ordinal if and only if they are isomorphic well-orderings or neither of them is a well-ordering. Recall that we used the notation $R_1 \sim R_2$ to denote exactly this relation. So this modified principle of abstraction, which we dub O-AB*, can be stated as:

$$\forall R_1 \forall R_2(\mathbf{ord}(R_1) = \mathbf{ord}(R_2) \leftrightarrow R_1 \sim R_2). \qquad \text{(O-AB}^\star\text{)}$$

Note that, given O-AB*, the first ordinal in the ordering $<$, call it 0, is the bad abstract. The next ordinal, call it 1, is the representative of empty relations. Next, we have 2, the representative of well-orderings with two-element fields. And so on. The progression is a well-ordering with the interesting feature that any ordinal is the representative of the ordering of its predecessors. For example, the portion of $<$ before 2, i.e., $<_2$, is a well-ordering with a two-element field and hence is represented by 2. This is captured by the following lemmas.

LEMMA 8.3. *If O-AB*, then the relation $<$ is a well-ordering of the ordinals.*

LEMMA 8.4. *Let R be a well-ordering. Given O-AB*, if $x = \mathbf{ord}(R)$, then there is F such that $<_x \cong_F R$.*

LEMMA 8.5. *If O-AB*, then for every ordinal x except the bad abstract, $x = \mathbf{ord}(<_x)$.*

The paradox of Burali–Forti corresponds to the reasoning leading to the following theorem.

THEOREM 8.6. O-AB* *is inconsistent.*

§9. Appendix C: The No-Class theory. Recall the definitions of §4.2. We have the languages L_2, L_Ω, and L_3. Moreover, we have a translation $*$ from L_Ω to L_3. We defined No-Class as the set of L_Ω-sentences such that their translation under $*$ is a theorem of third-order logic.

The first result about the No-Class theory is that it is indeed a theory. That is, the No-Class theory is closed under ordinary rules of inference. To demonstrate that this is so, we need to specify a system of deduction for L_Ω. The most convenient system for our purposes will be the Hilbert system whose sole rule of inference is modus ponens, and whose logical axioms are the universal closures of instances of the following schemes 1–8.

1. $\varphi \to (\varphi \to \psi)$.
2. $\varphi \to (\psi \to \chi) \to ((\varphi \to \psi) \to (\varphi \to \chi))$.
3. $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$.

In schemes 4–6, the schematic variable x ranges over variables of each sort in L_Ω (i.e., *obj*, *ord*, S_n, and T_n).

4. $\forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$.
5. $\varphi \to \forall x\varphi$, where x does not occur freely in φ.
6. $\forall x\varphi \to \varphi(t/x)$ where the term t is free for x in φ.

In schemes 7–8, the schematic variables s, t range over terms of sort *obj* and *ord*.

7. $s = s$.
8. $s = t \to (\varphi(t/x) \to \varphi(s/x))$ where s, t are free for x in φ.

We say that φ implies ψ ($\varphi \vdash \psi$) if there exists a sequence of L_Ω-formulas

$$\varphi, \chi_1, \dots, \chi_{n-1}, \psi$$

such that each formula is either the result of applying modus ponens to a pair of earlier formulas, or the universal closure of an instance of one of the schemes above.

The first result is given by the following theorem.

THEOREM 9.1 (Fundamental translation theorem). *If φ is in No-Class and $\varphi \vdash \psi$, then ψ is in No-Class*

Proof. It is clear that if φ and $\varphi \to \psi$ are in No-Class, then ψ is in No-Class. For, if each of these is in No-Class, then $[\varphi]^*$, and $[\varphi \to \psi]^*$, which is the same as $[\varphi]^* \to [\psi]^*$, are theorems of third-order logic. So, evidently, $[\psi]^*$ is a theorem of third-order logic as well. Thus No-Class is closed under modus ponens. To verify that No-Class is closed under $\vdash$, it is then enough to show that the closure of each of the schemes above is a member of No-Class.

We will argue directly that the translation of each scheme is provable using the corresponding closures of 1–8 for each sort in the language of third-order logic. This implies that the closure of the translated scheme is provable as well, since if a formula with a free variable is provable using those axioms, then its universal closure is provable from them as well. Thus the translation of the closure of each scheme is provable, as required.

For the tautologically valid schemes 1–3, this is immediate, since the translation $*$ commutes with truth functional connectives.

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

For axiom 4, we have four possible cases:

$$\forall x_i(\varphi \to \psi) \to (\forall x_i \varphi \to \forall x_i \psi) \overset{*}{\longmapsto} \forall x_i([\varphi]^* \to [\psi]^*) \to (\forall x_i[\varphi]^* \to \forall x_i[\psi]^*)$$
$$\forall \alpha_i(\varphi \to \psi) \to (\forall \alpha_i \varphi \to \forall \alpha_i \psi) \overset{*}{\longmapsto} \forall R_i([\varphi]^* \to [\psi]^*) \to (\forall R_i[\varphi]^* \to \forall R_i[\psi]^*)$$
$$\forall P^n_m(\varphi \to \psi) \to (\forall P^n_m \varphi \to \forall P^n_m \psi) \overset{*}{\longmapsto} \forall P^n_m([\varphi]^* \to [\psi]^*) \to (\forall P^n_m[\varphi]^* \to \forall P^n_m[\psi]^*)$$
$$\forall A^n_m(\varphi \to \psi) \to (\forall A^n_m[\varphi]^* \to \forall A^n_m[\psi]^*) \overset{*}{\longmapsto} \forall \mathbf{X}_m([\varphi]^* \to [\psi]^*) \to (\forall \mathbf{X}_m[\varphi]^* \to \forall \mathbf{X}_m[\psi]^*).$$

It is easy to verify that all four translations are tautologies of third-order logic.

For the scheme 5, we again have four possible cases:

$$\varphi \to \forall x_i \varphi \overset{*}{\longmapsto} [\varphi]^* \to \forall x_i[\varphi]^*$$
$$\varphi \to \forall \alpha_i \varphi \overset{*}{\longmapsto} [\varphi]^* \to \forall R_i[\varphi]^*$$
$$\varphi \to \forall P^n_m \varphi \overset{*}{\longmapsto} [\varphi]^* \to \forall P^n_m[\varphi]^*$$
$$\varphi \to \forall A^n_m \varphi \overset{*}{\longmapsto} [\varphi]^* \to \forall \mathbf{X}_m[\varphi]^*$$

In the first and third case, the fact that the listed variable does not occur freely in φ ensures that it does not occur freely in $[\varphi]^*$, since new variables of this sort are not introduced by $*$. In the second and fourth, the fact that the first listed variable does not occur freely in φ ensures that the second listed variable does not occur freely in $[\varphi]^*$, since it will only occur in $[\varphi]^*$ if φ contains a free variable of the relevant sort with a matching subscript. Hence, the resulting translations are provable in third-order logic as well.

For the scheme 6, i.e.,

$$\forall x \varphi \to \varphi(t/x)$$

we can consider four cases: either x is x_i, α_i, P^n_m, or A^n_m. If it is x_i, P^n_m, or A^n_m, then t must also be a variable of that sort, since those sorts have no nonvariable terms in L_Ω. If x is α_i, then t must be of sort *ord*, i.e., either α_j or $\mathbf{ord}(P_j)$ for some relation P_j of sort S_2.

An easy induction confirms that a first-order or second-order object variable which is freely substitutable for x remains freely substitutable after the translation. If x is a second-order ordinal variable, then it is replaced by a corresponding third-order variable, which is again, by induction, freely substitutable for the third-order variable corresponding to x. These observations handle the cases where x is x_i, P^n_m, or A^n_m.

In the final case, namely where x is an ordinal variable α_i, $[\forall \alpha_i \varphi]^*$ is $\forall R_i[\varphi(\mathbf{ord}(R_i)/\alpha_i)]^*$. Now, t is either another ordinal variable, or is $\mathbf{ord}(P_j)$ for some P_j not bound by any quantifier having scope over α_i.

We handle each of these cases in turn. For this argument, we will assume that $*$ has been extended to act on complex pseudoformulas of L_Ω (i.e., formulas containing pseudoterms) by commuting with connectives according to the same rules used to define $*$ initially.

In the first case, when t is an ordinal variable α_j, $[\varphi(\alpha_j/\alpha_i)]^*$ is $[\varphi(\mathbf{ord}(R_j)/\alpha_i)]^*$. By the correspondence of subscripts and the fact that α_j is freely substitutable for α_i in φ, we have that R_j is freely substitutable for R_i in $[\varphi(\mathbf{ord}(R_i)/\alpha_i)]^*$. So, in this case, the translation of our initial scheme is:

$$\forall R_i[\varphi(\mathbf{ord}(R_i)/\alpha_i)]^* \to [\varphi(\mathbf{ord}(R_i)/\alpha_i)]^*(R_j/R_i).$$

And this is an instance of the axiom 6, and hence a theorem of third-order logic.

The second case, when t is $\mathbf{ord}(P_i)$ for some P_i not bound by any quantifier having scope over α_i, is handled similarly. We conclude that instances of 6 are in No-Class.

The next axiom scheme, 7, is clear for variables of sort *obj*. For terms of sort *ord* we have:

$$\alpha_i = \alpha_i \overset{*}{\longmapsto} \exists V_i \exists V_i (V_i \sim R_i \wedge V_i \sim R_i \wedge V_i = V_i)$$
$$\mathbf{ord}(P_i) = \mathbf{ord}(P_i) \overset{*}{\longmapsto} \exists V_i \exists V_i (V_i \sim P_i \wedge V_i \sim P_i \wedge V_i = V_i)$$

and in each case the translation is an obvious theorem of third-order logic.

Finally, 8 is clear when we are not dealing with ordinal terms. When we are dealing with ordinal terms, 8 becomes one of the following:

$$\alpha_i = \alpha_j \to (\varphi(\alpha_i/\alpha_k) \to \varphi(\alpha_j/\alpha_k))$$
$$\mathbf{ord}(P_i) = \alpha_j \to (\varphi(\mathbf{ord}(P_i)/\alpha_k) \to \varphi(\alpha_j/\alpha_k))$$
$$\mathbf{ord}(P_i) = \mathbf{ord}(P_j) \to (\varphi(\mathbf{ord}(P_i)/\alpha_k) \to \varphi(\mathbf{ord}(P_j)/\alpha_k))$$

along with the obvious reversal of the second case. These schemes translate, respectively, as:

$$\exists V_i \exists V_j (R_i \sim V_i \wedge R_j \sim V_j \wedge V_i = V_j) \to ([\varphi(\alpha_i/\alpha_k)]^* \to [\varphi(\alpha_j/\alpha_k)]^*)$$
$$\exists V_i \exists V_j (P_i \sim V_i \wedge R_j \sim V_j \wedge V_i = V_j) \to ([\varphi(\mathbf{ord}(P_i)/\alpha_k)]^* \to [\varphi(\alpha_j/\alpha_k)]^*)$$
$$\exists V_i \exists V_j (P_i \sim V_i \wedge P_j \sim V_j \wedge V_i = V_j) \to ([\varphi(\mathbf{ord}(P_i)/\alpha_k)]^* \to [\varphi(\mathbf{ord}(P_j)/\alpha_k)]^*).$$

We argue that, in every case, the consequent will be provable, under the assumption of the antecedent, by induction on the complexity of φ. If φ is atomic, it either fails to contain α_k (in which case the conditional is tautological) or, if it contains α_k, it is either an equality or it is the predication of a second-order variable of a sequence of ordinal variables. In each case, verifying that the antecedent above implies the atomic consequent is an exercise in natural deduction, and confirms that the closure of the atomic instances of 8 are in No-Class.

If φ is $\psi_1 \wedge \psi_2$, we have:

$$\exists V_i \exists V_j (R_i \sim V_i \wedge R_j \sim V_j \wedge V_i = V_j) \to$$
$$([\psi_1(\alpha_i/t)]^* \wedge [\psi_2(\alpha_i/\alpha_k)]^* \to [\psi_1(\alpha_j/t)]^* \wedge [\psi_2(\alpha_j/\alpha_k)]^*).$$

By inductive hypothesis, the formulas

$$\exists V_i \exists V_j (R_i \sim V_i \wedge R_j \sim V_j \wedge V_i = V_j) \to ([\psi_1(\alpha_i/\alpha_k)]^* \to [\psi_1(\alpha_j/\alpha_k)]^*)$$

$$\exists V_i \exists V_j (R_i \sim V_i \wedge R_j \sim V_j \wedge V_i = V_j) \to ([\psi_2(\alpha_i/\alpha_k)]^* \to [\psi_2(\alpha_j/\alpha_k)]^*)$$

are theorems of third-order logic. So it is straightforward to derive the corresponding conditional for φ.

The remaining case, where φ is $\neg\psi$ or φ is $\forall x \psi$, are similarly straightforward, requiring only a mechanical application of the translation procedure and modicum of natural deduction. □

Let us now turn to second-order comprehension. Consider the parameter-free comprehension scheme, which is expressed in the vocabulary of L_Ω thus:

$$\exists P^n_m \forall x_1 \dots x_n (P^n_m(x_1 \dots x_n) \leftrightarrow \varphi). \tag{a}$$

Since $*$ commutes with $\forall P^n_m, \forall x_i$, and with the biconditional, and $*$ is the identity on $P^n_m(x_1 \dots x_n)$, we have the following translation of (a):

$$\exists P^n_m \forall x_1 \dots x_n (P^n_m(x_1 \dots x_n) \leftrightarrow [\varphi]^*). \tag{b}$$

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

But (b) is just an instance of the second-order comprehension scheme now expressed in L_3, and so it is a theorem of third-order logic.

The result extends to instances of comprehension with parameters. These will be universal closures of formulas like (a), where φ contains one or more free variables of any sort available in L_Ω. Translating such formulas will yield closures of formulas like (b), since each quantifier of the initial universal block will be carried by $*$ to a universal quantifier (although perhaps binding a variable of a different sort). Thus we have the following.

REMARK 9.2 (Second-order comprehension). *Every universal closure of the following scheme is part of No-Class*:

$$\exists P^n_m \forall x_1 \dots x_n (P^n_m(x_1 \dots x_n) \leftrightarrow \varphi(x_1 \dots x_n)),$$

where P^n_m does not occur in φ.

Since No-Class contains second-order comprehension and is closed under $\vdash$, it also contains any theorem of second order logic provable from the comprehension scheme. In particular, it contains the Lemmas 8.1 and 8.2 mentioned in Appendix B, since those can be proven without any resources going beyond comprehension.

By the argument given in §4.2, we also have that the principle of ordinal abstraction, O-AB, is part of No-Class.

THEOREM 9.3 (O-AB). *The following statement belongs to No-Class.*

$$\forall P_i, P_j (WO(P_i) \wedge WO(P_j) \rightarrow (\mathbf{ord}(P_i) = \mathbf{ord}(P_j) \leftrightarrow P_i \cong P_j)).$$

Moreover, by a similar argument, we even have that the stronger principle O-AB*, discussed in Appendix B, is part of No-Class.

THEOREM 9.4 (O-AB*). *The following statement belongs to No-Class.*

$$\forall P_i, P_j (\mathbf{ord}(P_i) = \mathbf{ord}(P_j) \leftrightarrow P_i \sim P_j)$$

COROLLARY 9.5. *The claim that there is exactly one "bad ordinal" is in No-Class.*

Since No-Class includes a sort for relations over ordinals, one might naturally ask whether these relations support comprehension. It turns out that they do. But the proof requires a bit more effort than the demonstration of the Remark above that relations over objects support comprehension. We begin with a lemma.

LEMMA 9.6 (Substitution lemma). *If $[\varphi(\mathbf{ord}(P_{i_1}) \dots \mathbf{ord}(P_{i_n}))]^*$ is a translation of a formula in L_Ω, then the closure of*

$$\forall P_{i_1} \dots P_{i_n}, P_{j_1} \dots P_{j_n} (P_{i_1} \sim P_{j_1} \wedge \dots \wedge P_{i_1} \sim P_{j_1} \rightarrow ([\varphi(\mathbf{ord}(P_{i_1}), \dots, \mathbf{ord}(P_{i_n}))]^* \leftrightarrow [\varphi(\mathbf{ord}(P_{j_1}), \dots, \mathbf{ord}(P_{j_n}))]^*))$$

is a theorem of third-order logic.

Proof. By induction on the complexity of φ. This is clear for φ atomic, and the induction cases are straightforward. □

Now we are ready to prove the target theorem.

THEOREM 9.7 (Ordinal comprehension). *Every closure of the following scheme is part of No-Class*:

$$\exists A^n_m \forall \alpha_{i_1} \dots \alpha_{i_n} (A^n_m(\alpha_{i_1} \dots \alpha_{i_n}) \leftrightarrow \varphi),$$

where A^n_m does not occur in φ.

Proof. Without loss of generality, we may assume that φ does not contain any free ordinal variables other than $\alpha_{i_1} \dots \alpha_{i_n}$. Otherwise, by replacing the stray ordinal variables with appropriate terms of the form $\mathbf{ord}(P_i)$, we can find another formula φ_1 without any free ordinal variables other than $\alpha_{i_1} \dots \alpha_{i_n}$ such that

$$[\exists A^n_m \forall \alpha_{i_1} \dots \alpha_{i_n} (A^n_m(\alpha_{i_1} \dots \alpha_{i_n}) \leftrightarrow \varphi)]^*$$

is logically equivalent to

$$[\exists A^n_m \forall \alpha_{i_1} \dots \alpha_{i_n} (A^n_m(\alpha_{i_1} \dots \alpha_{i_n}) \leftrightarrow \varphi_1)]^*$$

A partial translation of the ordinal comprehension scheme will be a universal closure of the following

$$\exists \mathbf{X}_m \forall R_{i_1} \dots R_{i_n} \Big([A^n_m(\mathbf{ord}(R_{i_1}), \dots, \mathbf{ord}(R_{i_n}))]^* \leftrightarrow [\varphi(\mathbf{ord}(R_{i_1}) \dots \mathbf{ord}(R_{i_n}))]^* \Big). \quad \text{(a)}$$

We can further unpack the left-hand side of the embedded biconditional to produce

$$\exists V_{i_1} \dots \exists V_{i_n} (V_{i_1} \sim R_{i_1} \wedge \dots \wedge V_{i_n} \sim R_{i_n} \wedge \mathbf{X}_m(V_{i_1}, \dots, V_{i_n})).$$

Now, apply third-order comprehension to the formula:

$$\exists R_{j_1} \dots \exists R_{j_n} \left(V_{j_1} \sim R_{j_1} \wedge \dots \wedge V_{j_n} \sim R_{j_n} \wedge [\varphi(\mathbf{ord}(R_{j_1}) \dots \mathbf{ord}(R_{j_k}))]^* \right)$$

which we may abbreviate as $\psi(V_{j_1}, \dots, V_{j_n})$. Thus:

$$\exists \mathbf{X}_k \; \forall V_{j_1} \dots \forall V_{j_n} [\mathbf{X}_k(V_{j_1} \dots V_{j_n}) \leftrightarrow \psi(V_{j_1} \dots V_{j_n})]. \quad \text{(b)}$$

Pick a particular witness $\mathbf{X}_k$ to (b). We claim that this witness is also a witness to (a) above.

For, let $R_{i_1} \dots R_{i_n}$ be arbitrary, and suppose that $\mathbf{X}_k$ together with these relations satisfies the left-hand side of (a). We need to derive the right-hand side of (a). By the unpacking of the left-hand side, we then have that there exist $V_{i_1} \dots V_{i_n}$ falling under $\mathbf{X}_k$, such that $V_{i_1} \sim R_{i_1} \dots V_{i_n} \sim R_{i_n}$. Since $V_{i_1} \dots V_{i_n}$ fall under $\mathbf{X}_k$, we know—by the comprehension scheme (b) that introduced $\mathbf{X}_k$—that there are $R_{j_1} \dots R_{j_n}$ such that $[\varphi(\mathbf{ord}(R_{j_1}) \dots \mathbf{ord}(R_{j_n}))]^*$ and such that $R_{j_1} \sim V_{i_1}, \dots, R_{j_n} \sim V_{i_n}$. Hence, by transitivity of $\sim$, we have $R_{j_1} \sim R_{i_1} \dots R_{j_n} \sim R_{i_n}$. But then by Lemma 9.6, we see that we have the right-hand side of (a), $[\varphi(\mathbf{ord}(R_{i_1}) \dots \mathbf{ord}(R_{i_n}))]^*$, as required.

Now we prove the other direction. Suppose that $[\varphi(\mathbf{ord}(R_{i_1}) \dots \mathbf{ord}(R_{i_n}))]^*$. We then have $\psi(R_{i_1}, \dots, R_{i_n})$, taking the $R_{i_1} \dots R_{i_n}$ themselves as witnesses to the existential quantifiers of $\psi(R_{i_1}, \dots, R_{i_n})$. Hence $\mathbf{X}_k(R_{i_1}, \dots R_{i_n})$. But then, once again taking the $R_{i_1} \dots R_{i_n}$ themselves as witnesses, we see that the $R_{i_1} \dots R_{i_n}$ together with $\mathbf{X}_k$ satisfy the unpacking of the left-hand side of the translation of ordinal comprehension, as required. □

We have now shown not only that No-Class is a theory, but also that it is a natural theory of the ordinals as regimented in the language L_Ω. On the one hand, it includes comprehension principles for each of the concept sorts in this language. On the other hand, it includes intuitive principles concerning the notion of ordinal, chiefly that of ordinal abstraction. We conclude by showing that these principles suffice to establish two additional intuitive facts about the ordinals. The first is that the ordinals are well-ordered. The second is that the ordinals support transfinite induction.

THEOREM 9.8. *Let $<$ be the ordering of ordinals defined in Appendix B. Then it is a theorem of No-Class that $<$ is well-ordered.*

Proof. Working in No-Class, we need to establish that $<$ is:

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

(a) transitive,
(b) total,
(c) well-founded.

One proves (a) and (b) by proving in second-order logic the universal closures of the following statements, where F, G, R, S, T can be any of the variables P_i.

$$R \cong_G S_{x_1} \ \wedge\ S \cong_F T_{x_2} \ \rightarrow\ R \cong_{F \circ G} T_{F(x_1)}$$

$$\mathrm{WO}(R) \wedge \mathrm{WO}(S) \ \rightarrow\ (R \cong S \ \vee\ \exists x\ (R_x \cong S \ \vee\ S_x \cong R))$$

These facts about well-orderings then straightforwardly imply the relevant facts about the corresponding ordinals.

The proof that the statement asserting the well-foundedness of $<$ belongs to No-Class is a formalization of the following reasoning. For legibility, we use variables A and A' as abbreviations for unary ordinal relations, the variable P as an abbreviation for a unary object relation, variables α, β, γ as abbreviations for ordinal variables, and variables x and R with various subscripts as abbreviations for object variables and binary object relation variables, respectively.

Suppose there were A contained in the field of $<$ with no least element. Let α be some ordinal in A, and let A' be given by the following instance of ordinal comprehension:

$$\forall \beta (A'(\beta) \ \leftrightarrow\ \beta < \alpha \ \wedge\ A(\beta)).$$

We claim that A' has no least element. If it had one, say β, then this ordinal would not be least in A, since A is assumed not to have a least element. So there would be some $\gamma < \beta < \alpha$ with $A(\gamma)$. But then $A'(\gamma)$, contradicting the fact that β is least in A'. So A' has no least element.

We know that α represents at least one well-ordering. Let R be some such well-ordering, so that $\mathbf{ord}(R) = \alpha$. For each β such that $A'(\beta)$, there exist R' and x such that $\mathbf{ord}(R') = \beta$, and $R_x \cong R'$. Let P be a concept contained the field of R satisfying

$$\forall x (P(x) \ \leftrightarrow\ \exists \beta\ (A'(\beta) \ \wedge\ \mathbf{ord}(R_x) = \beta)).$$

Consider an arbitrary element y such that $P(y)$, and β such that $\mathbf{ord}(R_y) = \beta$. It follows from the definition of P that $A'(\beta)$. Since A' has no least element and $A'(\beta)$, we know that there is some γ such that $A'(\gamma)$ with $\gamma < \beta$. Since $A'(\gamma)$, $\gamma < \alpha$, thus there exists z such that $\mathbf{ord}(R_z) = \gamma$. Hence $P(z)$. Given that $\gamma < \beta$, $\mathbf{ord}(R_z) < \mathbf{ord}(R_y)$. That implies that $R(z, y)$.

We have shown that, for an arbitrary y such that $P(y)$, there is a smaller z such that $P(z)$. So we have confirmed that P has no least element. But this contradicts the well-foundedness of R. So our initial assumption—that there is A contained in the field of $<$ with no least element—must be rejected. Thus $<$ is well-founded. □

COROLLARY 9.9 (Transfinite induction). *The universal closure of each instance of the following scheme is in No-Class.*

$$\forall \alpha\ (\forall \beta\ (\beta < \alpha \ \rightarrow\ \varphi(\beta)) \ \rightarrow\ \varphi(\alpha)) \rightarrow \ \forall \alpha\ \varphi(\alpha).$$

Proof. Suppose that we have the antecedent of the claim above. Towards a contradiction, suppose that $\exists \alpha \neg \varphi(\alpha)$. By ordinal comprehension (Theorem 9.7), there is a nonempty property A of ordinals satisfying:

$$\forall \alpha (A(\alpha) \ \leftrightarrow\ \neg \varphi(\alpha)).$$

https://doi.org/10.1017/S1755020316000484 Published online by Cambridge University Press

Since $<$ is total, A is a subproperty of a field of $<$. Moreover, since $<$ is well-founded, there is a least ordinal in A. Call this minimal ordinal α, so that $\neg\varphi(\alpha)$ and (by minimality) $\forall\beta\ (\beta < \alpha \rightarrow \varphi(\beta))$. Hence, by the assumed antecedent, $\varphi(\alpha)$ obtains, contradicting our claim that $\neg\varphi(\alpha)$. □

BIBLIOGRAPHY

Boolos, G. (1998). Reply to Charles Parsons' "Sets and Classes". In Jeffery, R., editor. *Logic, Logic, and Logic*. Cambridge, MA: Harvard University Press, pp. 30–36.

Burali-Forti, C. (1897). Una questione sui numeri transfiniti. *Rendiconti del circolo matematico di Palermo*, **11**, 154–164.

Burgess, J. (2004). E pluribus unum: Plural logic and set theory. *Philosophia Mathematica*, **12**, 193–221.

Cantor, G. (1878). Ein Beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik*, **84**, 242–258.

Cantor, G. (1883). *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen*. Teubner, Leipzig. Translated as "Foundations of a general theory of manifolds: A mathematico-philosophical investigation into the theory of the infinite". In Ewald, W., editor. *From Kant to Hilber: A Source Book in the Foundations of Mathematics*, Vol. 2. Oxford University Press, 1996, pp. 878–919.

Cook, R. T. (2003). Iteration one more time. *Notre Dame Journal of Formal Logic*, **44**, 63–92.

Ebels-Duggan, S. (2015). The nuisance principle in infinite settings. *Thought*, **4**, 263–268.

Ferreira, F. (2005). Amending Frege's Grundgesetze der Arithmetik. *Synthese*, **147**, 3–19.

Ferreira, F. & Wehmeier, K. F. (2002). On the consistency of the Δ^1_1-CA fragment of Frege's *Grundgesetze*. *Journal of Philosophical Logic*, **31**, 301–311.

Ferreirós, J. (2007). *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics*. Basel: Birkhäuser.

Florio, S. & Linnebo, Ø. (in progress). *The Many and the One: A Philosophical Study*.

Frege, G. (1884). *Grundlagen der Arithmetik*. Translated by J. L. Austin as *The Foundations of Arithmetic* (second revised edition). Blackwell, 1974.

Frege, G. (1979). Posthumous writings. In Hermes, H., Kambartel, F., and Kaulbach, F., editors. *Posthumous Writings*. Oxford: Basil Blackwell.

Glanzberg, M. (2004). Quantification and realism. *Philosophy and Phenomenological Research*, **69**, 541–571.

Glanzberg, M. (2006). Context and unrestricted quantification. In Rayo, A. and Uzquiano, G., editors. *Absolute Generality*. Oxford: Oxford University Press, pp. 20–44.

Gödel, K. (1944). Russell's mathematical logic. In Feferman, S., editor. *Collected Works, Volume II, Publications 1938–1974*. Oxford: Oxford University Press, 1990, pp. 119–143.

Gödel, K. (1964). What is Cantor's continuum problem? In Feferman, S., editor. *Collected Works, Volume II, Publications 1938–1974*. Oxford: Oxford University Press, 1990, pp. 176–188.

Grassmann, H. (1844). *Die lineale Ausdehnungslehre: ein neuer Zweig der Mathematik*. Otto Wigand. Translated by Lloyd C. Kannenberg in *A New Branch of Mathematics. The Ausdehnungslehre of 1844 and Other Works*. Chicago, Open Court: 1995.

Grassmann, H. 1847. Geometrische Analyse geknüpft an die von Leibniz erfundene geometrische Charakteristik. Gekrönte Preisschrift, Leipzig: Wiedmann. Translated by Lloyd C. Kannenberg in *A New Branch of Mathematics*. The Ausdehnungslehre of 1844 and Other Works, Open Court, 1995.

Hazen, A. P. (1986). Logical objects and the paradox of Burali-Forti. *Erkenntnis*, **24**, 283–291.

Heath, T. L. (editor) (1908). *The Thirteen Books of the Elements*, Vol. 2. Cambridge, UK: Cambridge University Press.

Heck, R. (1996). The consistency of predicative fragments of Frege's *Grundgesetze der Arithmetik*. *History and Philosophy of Logic*, **17**, 209–220.

Hellman, G. (2011). On the significance of the Burali-Forti paradox. *Analysis*, **71**, 631–637.

Hodes, H. (1986). Logicism and the ontological commitments of arithmetic. *Journal of Philosophy*, **81**, 123–149.

Jourdain, P. E. B. (1904). On the transfinite cardinal numbers of well-ordered aggregates. *Philosophical Magazine*, **7**, 61–75.

Klement, K. C. (in press). A generic Russellian elimination of abstract objects. *Philosophia Mathematica*.

Leibniz, G. W. (1989). *Philosophical Papers and Letters*. Dordrecht: Kluwer Academic Publishers.

Linnebo, Ø. (2004). Predicative fragments of Frege Arithmetic. *Bulletin of Symbolic Logic*, **10**, 153–174.

Linnebo, Ø. (2010). Pluralities and sets. *Journal of Philosophy*, **107**, 144–164.

Linnebo, Ø. & Pettigrew, R. (2014). Two types of abstraction for structuralism. *Philosophical Quarterly*, **64**, 267–283.

Mancosu, P. (2015). *Grundlagen*, section 64: Freges discussion of definitions by abstraction in historical context. *History and Philosophy of Logic*, **36**, 62–89.

Moore, G. H. & Garciadiego, A. (1981). Burali-Forti's paradox: A reappraisal of its origins. *Historia Mathematica*, **8**, 319–350.

Parsons, C. (1974a). The liar paradox. *Journal of Philosophical Logic*, **3**, 381–412.

Parsons, C. (1974b). Sets and classes. *Noûs*, **8**, 1–12.

Poincaré, H. (1912). The latest efforts of the logisticians. *The Monist*, **22**, 524–539.

Russell, B. (1903). *The Principles of Mathematics* (second edition). New York: W.W. Norton and Company.

Russell, B. (1905). On some difficulties in the theory of transfinite numbers and order types. *Proceedings of the London Mathematical Society*, **4**, 29–35.

Russell, B. (1908). Mathematical logic as based on a theory of types. *American Journal of Mathematics*, **30**, 222–262.

Russell, B. (1910). Some explanations in reply to Mr. Bradley. *Mind*, **19**, 373–378.

Shapiro, S. (2003). All sets great and small: And I do mean ALL. *Philosophical Perspectives*, **17**, 467–490.

Shapiro, S. (2007). Burali-Forti's revenge. In Beall, J., editor. *Revenge of the Liar: New Essays on the Paradox*. Oxford: Oxford University Press, pp. 320–344.

Shapiro, S. & Weir, A. (1999). New V, ZF and abstraction. *Philosophia Mathematica*, **7**, 293–321.

Shapiro, S. & Wright, C. (2006). All things indefinitely extensible. In Rayo, A. and Uzquiano, G., editors. *Absolute Generality*. Oxford: Oxford University Press, pp. 253–304.

Simpson, S. (2009). *Subsystems of Second Order Arithmetic*. Cambridge, UK: Cambridge University Press.

Studd, J. P. (2016). Abstraction reconceived. *British Journal for the Philosophy of Science*, **67**, 579–615.

Uzquiano, G. (2003). Plural quantification and classes. *Philosophia Mathematica*, **11**, 67–81.

Walsh, S. (in press). Fragments of Frege's *Grundgesetze* and Gödels constructible universe. *Journal of Symbolic Logic*.

Walsh, S. & Ebels-Duggan, S. (2015). Relative categoricity and abstraction principles. *Review of Symbolic Logic*, **8**, 572–606.

Wright, C. (1999). Is Hume's Principle analytic? *Notre Dame Journal of Formal Logic*, **40**, 6–30.

Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, **16**, 29–47. Translated as "On Boundary Numbers and Domains of Sets: New Investigations in the Foundations of Set Theory". In Ewald, W., editor. *From Kant to Hilber: A Source Book in the Foundations of Mathematics*, Vol. 2. Oxford University Press, 1996, pp. 1219–1233.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF BIRMINGHAM
BIRMINGHAM, UK
E-mail: s.florio@bham.ac.uk

DEPARTMENT OF PHILOSOPHY
KANSAS STATE UNIVERSITY
MANHATTAN, KS, USA
E-mail: gleachkr@ksu.edu