1 Introduction

> Admittedly, the present state of affairs where we run up against the paradoxes is intolerable. Just think, the definitions and deductive methods which everyone learns, teaches, and uses in mathematics, the paragon of truth and certitude, lead to absurdities! If mathematical thinking is defective, where are we to find truth and certitude?
>
> – Hilbert (Bencerraf and Putnam, 1983)

Philosophers have often used first-order logic to analyze mathematical and scientific claims. However, we seem to grasp a notion of logical possibility prior to, and independent from, our grasp of mathematical objects like set-theoretic models. Powerful reasons to accept this notion as an additional logical primitive have emerged (Boolos 1985; Etchemendy 1990; Gómez-Torrente 2000; Hanson 2006; Field 2008).

In this book, I'll make a case that philosophical analyses using (a natural generalization of) this notion of logical possibility can illuminate the philosophy of mathematics, metaphysics and philosophy of language. Much of this case will focus on pure mathematics and the philosophy of set theory. For example, I will show that formulating set theory in terms of logical possibility (along Potentialist lines suggested by Putnam and Hellman in response to the Burali-Forti paradox) yields a new and more appealing justification for one of the standard ZFC (Zermelo–Frankel with choice) axioms of set theory. This brings us closer to realizing the traditional hope of justifying mainstream mathematics from principles that seem clearly true.

However, we will see that using a primitive logical possibility operator can also help us develop a modestly neo-Carnapian philosophy of language. And philosophical analyses of scientific theories using the logical possibility operator can illuminatingly "factor" scientific claims into a logico-mathematical component and a remainder, in a way that reveals hidden heterogeneity in the role of mathematics in the sciences and clarifies debates over Quinean and post-Quinean indispensability arguments.

1.1 Mathematics as a Touchstone and the Centrality of Set Theory

Mathematical proofs provide a touchstone of clarity and convincingness which serves as an inspiration to philosophy and other disciplines. While it is possible to doubt the results of mainstream mathematical arguments (philosophers are capable of doubting anything), there's something striking about just how convincing mathematical proofs often are. Consider the standard argument that there are infinitely many primes. Even philosophers who deny that there are numbers (and hence think the argument as usually stated is unsound) are strongly tempted to say that we know *something like* the premises and that these proofs provide some kind of valuable amplification of this knowledge. The premises we use in informal mathematical reasoning have a combination of prima facie obviousness, power and generality, which makes them exemplary tools for expanding our knowledge and resolving disputes in cases where people's initial hunches disagree. It's no surprise that Leibniz[1] wished philosophers could resolve their disputes like mathematicians by saying "let us calculate" (or at least, "let us each look for a proof").

In many ways, set theory lies at the heart of modern mathematics, and it does powerful mathematical (not just philosophical) work as a foundation for the whole. So, one might hope that the set-theoretic foundations for mathematics would share the clarity and convincingness we hope for from mathematical arguments.

However, certain problems in the philosophical foundations of set theory raise worries. These concerns are more mathematical and specific to set theory than standard philosophical worries about, e.g., whether there are any abstract objects. And these concerns are more threatening to mathematical practice than philosophical doubts typically are, insofar as they raise doubts about whether the standard ZFC axioms of set theory are even logically consistent.

The development of set theory resolved a great many problems in analysis. It also provided a formal framework to allow interactions between various areas of mathematics – creating, as Hilbert (1926) famously observed, a kind of mathematical paradise. However, contradiction, in the form of Russell's paradox, threatened Hilbert's paradise.

This problem was almost, but not quite, solved by accepting the iterative hierarchy conception (also called the cumulative hierarchy conception) of sets and the standard ZFC first-order axioms for set theory. On the iterative hierarchy conception of set, we think about the sets as being formed in layers, with the sets at each layer containing only elements from prior layers. This lets us avoid the appearance that there should be a set of all sets that aren't members of themselves, and hence Russell's paradox. And it is hard to deny that the mathematical results which are currently stated in terms of this conception of set theory reflect genuine and important knowledge of some kind.

However, as we will see below, a few further problems arise, including a question of how to justify our acceptance of all the ZFC axioms. So, we may ask, is the price of

[1] See Chrisley and Begeer (2000: 14).

remaining in Cantor's set-theoretic paradise giving up the old ambition of founding mathematics on intrinsically obvious seeming principles?

One of the main projects of this book will be to answer this question in the negative. I will develop a unified understanding of set-theoretic talk, which vindicates our intuitive expectations regarding set theory – and demonstrates that all theorems of set theory (ZFC) are derivable from principles that seem clearly true (not to say indubitable). The approach I'll develop differs from standard Actualist set theory, in a few important ways.

1.2 Actualism and Its Discontents

On standard Actualist approaches to set theory, set theory studies abstract mathematical objects called "sets," which form an iterative hierarchy as evoked above. Apparent existence claims made by set theorists (like "there is a set which has no elements") are made true by the existence of corresponding objects, just like ordinary existence claims about cities or electrons or cars. Crudely speaking, three problems arise when we look at set theory from this familiar point of view (each of which I'll describe in much greater detail in Section 2.2).

First, there's a problem about our conception of the hierarchy of sets as a mathematical structure. We don't seem to have a precise conception of the intended structure of the iterative hierarchy of sets, in the way that we do *seem* to have a conception of the natural numbers. In particular, the height of the hierarchy of sets is left vague or mysterious. As the Burali-Forti paradox (Burali-Forti 1897) dramatizes, a certain naive conception of the hierarchy of sets (as containing ordinals corresponding to all ways some objects could be well ordered by some relation) is incoherent. And once this naive paradoxical conception of the height of the hierarchy of sets is rejected, we don't seem to have a precise idea about the intended height of the hierarchy of sets left over to replace it. It appears that, for any height that the hierarchy of sets could achieve, there could be a strictly larger structure, which extra layers of "sets" on top of the original hierarchy and fits with everything in our conception of the sets equally well. But it seems arbitrary to suppose that the hierarchy of sets happens to stop at any particular point.

Second, there's a worry about generality and the role of set theory as a foundation for all of mathematics. One might hope that set theory would be able to represent any mathematical structure one might want to study. The idea that set theory has this kind of generality is prima facie quite intuitive. But Actualist set theory is prima facie unable to represent the study of mathematical structures that are "too large." Thus, Actualism makes it hard to capture the intuition that "any possible structure" should, in some sense, be fair game for mathematical study, and hence treatment within set theory.

Third, there's a problem about intuitively justifying the standard ZFC axioms of set theory. As noted above, mathematical proofs can usually be reconstructed so as to derive their conclusion from premises that are prima facie extremely plausible (if not

https://doi.org/10.1017/9781108992756.001 Published online by Cambridge University Press

indubitable or impossible to philosophically or empirically cast doubt upon). So, one might hope that (once we understand set theory correctly), every claim provable from the ZFC axioms for set theory can be shown to follow from principles that seem clearly true in the way that foundational mathematical axioms often do.

However, philosophers have noted that the axiom of Replacement (one of the standard ZFC axioms) doesn't seem clearly true and deriving it from principles that do seem clearly true has proved challenging. For example, Hilary Putnam (2000) writes, "Quite frankly, I see no intuitive basis at all for . . . the axiom of Replacement. Better put, I do not see that a notion of set on which that axiom is clearly true has ever been explained." Instead, philosophers of mathematics and mathematicians have made do with less ambitious approaches to justification. For example, some mathematicians have invoked external justifications, like the failure of mathematicians to discover any contradictions during over a century of work with ZFC. Others have shown that the axiom of Replacement follows from certain powerful and plausible (if not clearly consistent or true) principles that also imply many of the other axioms of set theory. But insofar as the latter powerful conceptions aren't clearly consistent, showing this doesn't suffice to show Replacement follows from clearly true premises. Nor does simply combining Replacement itself with the bare-bones conception of the intended width of an iterative hierarchy of sets evoked above yield something that seems obviously consistent. This state of affairs can feel unsatisfying.

1.3 Potentialism and the Justification Problem

In response to the first two problems above, philosophers like Putnam, Parsons, Hellman and Linnebo (Putnam 1967; Parsons 1977; Hellman 1994; Linnebo 2018a) have proposed that we should reject[2] Actualism about set theory in favor of a different approach: Potentialism. The key idea behind Potentialism is that, rather than taking set theory to be the study of a single hierarchy of sets which stops at some particular point (as the Actualist does), we should instead interpret set theorists as making modal claims about what hierarchy-of-sets-like structures are possible and how such structures could (in some sense) be extended.

As we will see in more detail in Chapter 2, switching to a Potentialist understanding of set theory solves the first problem for Actualism noted above. The Potentialist avoids postulating an arbitrary (or indeterminate) height for the hierarchy of sets,[3] and Potentialism also plausibly solves the second problem above, by honoring the intuition that any possible mathematical structure can be studied within set theory.

[2] Strictly speaking Putnam proposes Actualism and Potentialism are (in some sense), two perspectives on the same thing.

[3] At least Potentialists like Hellman (1996), Linnebo (2018a) and Studd (2019) avoid positing such an arbitrary stopping point for the sets. Putnam's view, on which Actualist set theory and Potentialist set theory are (somehow) two perspectives on the same thing, does not let us avoid this problem in any obvious way.

However, the problem of justifying the axiom of Replacement from premises that seem clearly true remains. Contemporary Potentialists can, and generally do, prove that (the Potentialist translations of) every theorem of ZFC can be derived from certain intuitive assumptions about logical possibility, or some other such modal notion. However, these proofs all use principles of modal logic that aren't (and aren't claimed to) be clearly true in the way invoked by Putnam. The existing Potentialist literature has shown that Potentialism is *no worse off than Actualism* with regard to the problem of justifying Replacement that Putnam raises.[4] However, neither Potentialists nor Actualists have put forward a solution to this problem.

1.4 Outline

In this book, I will attempt to solve the above problem of justifying the axiom of Replacement from principles that seem clearly true (or at least improve on existing solution attempts) and clarify the foundations of set theory.

In Part I, I will argue that we should indeed be Potentialists about set theory for essentially the reasons indicated above, and then review major existing formulations of Potentialism about set theory and some problems for each. I'll discuss and contrast two major existing versions of Potentialist set theory: the Putnamian approach developed by Putnam and Hellman which I will largely follow, and an alternative Parsonian approach recently explored by Linnebo and Studd, which appeals to a notion of interpretational possibility, rather than metaphysical or logical possibility.

I will develop and advocate a particular form of Potentialist set theory. Although this approach largely blends and simplifies ideas from Putnam and Hellman, it has the distinctive feature of replacing claims that "quantify-in" to the diamond of logical possibility (and thereby talk about what's logically possible *for objects*) with claims about what's logically possible *given certain structural facts*, expressed using a new piece of logical vocabulary I'll call the conditional logical possibility operator ($\Diamond_{\ldots}$). Cashing out Potentialist set theory in these terms lets us avoid certain philosophical controversies,[5] as well as practically helping us state axioms that can be easily grasped and recognized as saying something clearly true.

In Chapter 2 I will discuss Actualist approaches to set theory and expand on the worries for them noted above. In Chapter 3 I'll discuss how adopting some Potentialist approach to set theory promises to solve these worries and review existing forms of the Putnamian style of Potentialism. I will defend Hellman's use of a notion of logical possibility to cash out Potentialist set theory but note that controversies over quantified modal logic raise some problems for using his version of Potentialism in our foundational project.

[4] Existing potentialists (Hellman 1994; Linnebo 2018a; Studd 2019) have generally adopted some version of a Potentialist translation of Replacement as an axiom (schema). For while these Potentialist translations are not clearly true, they are (we will see) as attractive as corresponding instances of the Replacement schema understood actualistically.

[5] See Section 3.3.1.

In Chapter 4 I'll introduce my preferred style of Potentialist paraphrase and the key notion of conditional (structure-preserving) logical possibility I'll use to give these paraphrases. Finally, in Chapter 5 I'll contrast the above approach to Potentialist set theory with those advocated by Linnebo and Studd, major proponents of an alternate "Parsonian" school of Potentialist set theory.

In Part II I will turn to the core mathematical project of this book: justifying the ZFC axioms. I'll propose general purpose axioms for reasoning about conditional logical possibility which (I claim!) seem clearly true in the way our foundational project requires. Then I will show that these axioms justify our use of normal first-order reasoning for set-theoretic claims (i.e., claims in the first-order language of set theory) even when those claims are understood potentialistically. Specifically, if we let $\phi^\Diamond$ stand for the Potentialist translation of a set-theoretic claim ϕ, let $\vdash_{\mathrm{FOL}}$ be provability in first-order logic and $\vdash$ be provability in the formal system proposed in this book, we can show the following.

Theorem 1.1 (Logical Closure of Translation). *Suppose Φ, Ψ are sentences in the language of set theory and $\Phi \vdash_{FOL} \Psi$ then $\Phi^\Diamond \vdash \Psi^\Diamond$.*

With this theorem in mind, all that's needed to justify normal mathematical practice is to demonstrate that if ϕ is an axiom of ZFC then $\vdash \phi^\Diamond$ holds. A key idea here will be to use certain *non-interference* intuitions to justify the (Potentialist translation of) the axiom of Replacement, rather than simply taking the latter as an axiom, as current Potentialists tend to do. Putting these pieces together we can conclude that for all set theoretic sentences ϕ:

$$\mathrm{ZFC} \underset{\mathrm{FOL}}{\vdash} \phi \text{ then} \vdash \phi^\Diamond$$

That is, reasoning in ZFC as if one were talking about an Actualist hierarchy of sets is harmless. If one can prove that ϕ in ZFC then the Potentialist translation of ϕ (written $\phi^\Diamond$ above) is (true and indeed) provable in my formal system.

Note that since I choose axioms of reasoning about conditional logical possibility which are attractive for general use rather than ones that directly mirror Actualist ZFC set theory (as other Potentialists have done in proving versions of the theorem above), it's not at all obvious whether the reverse direction of the above conditional, i.e., "If $\vdash \phi^\Diamond$ then ZFC $\vdash_{FOL} \phi$" is true. In principle, there is some hope that the modal axioms I propose (or, more plausibly, further principles about conditional logical possibility that seem equally clearly true) will let one vindicate new axioms for set theory, going beyond the ZFC axioms.

Finally, in Part III of the book, I'll turn to larger philosophical questions. In Chapter 10, I consider two ways my story about set theory can fit into a larger philosophical picture of mathematics and its applications: a Nominalist approach and the weakly neo-Carnapian approach I ultimately favor.

In Chapters 11–14, I'll discuss the Nominalist approach to non-set theoretic mathematical objects and Indispensability arguments. I'll argue that adding some cheap

https://doi.org/10.1017/9781108992756.001 Published online by Cambridge University Press

tricks to the above paraphrase strategy lets the Nominalist answer certain classic Quinean and Explanatory indispensability arguments. However, I'll suggest that the mathematical Nominalist *may* face serious and under-discussed worries about reference and grounding.

In Chapters 15 and 16, I'll explain the weakly neo-Carnapian approach to non-set theoretic mathematical objects I favor, and argue that adopting it helps solve or avoid these reference and grounding problems and has certain other advantages (while retaining many benefits of Nominalism). The resulting view is a kind of neo-Carnapianism realism about mathematical objects, which drops Carnap's radical anti-metaphysical ambitions but keeps mathematicians' freedom to talk in terms of arbitrary logically coherent pure mathematical structures.

Finally, in Chapters 17–19, I'll discuss how the overall picture of mathematics developed in this book relates to traditional questions about Logicism, Structuralism and human access to facts about objective proof-transcendent mathematical facts.

1.5 Caveats and Clarifications

Let me finish this introduction with some quick caveats about the nature and aim of my project.

First, I don't claim set theorists should literally rewrite set theory textbooks in Potentialist terms. Mathematicians' current practice of (making arguments which can be reconstructed as) proving things in first-order logic from the ZFC axioms is fine. And doing something like logical deduction from purely first-order axioms may be unavoidably easier (for minds like ours) than thinking about the elaborate modal claims that figure in Potentialist set theory. If one thinks about typical set theoretic talk as abbreviating Potentialist claims, then the main result of Part II shows that it's unnecessary to unpack this abbreviation in most mathematical contexts.

However, I *am* suggesting Potentialism reflects what people should say when we think about set theory in many philosophical and foundational mathematical contexts. Replacing Actualist set-theoretic claims with their Potentialist paraphrases solves various intuitive puzzles and makes sense of things that we normally want to say about set theory.[6]

[6] In this proposal I somewhat mirror Hellman's response (Hellman 1998) to Burgess and Rosen's dilemma (Burgess and Rosen 1997). Burgess and Rosen argue that nominalistic paraphrases must be intended as either a hermeneutic theory of what scientists mean or a revolutionary theory of what they should say, but typical Nominalist paraphrases don't seem adequately supported by scientific motivations for either as they wouldn't be accepted to linguistics or physics journals.

One response to this would be to say that nominalistic paraphrases reflect what we should say in *philosophical* contexts, and this differs from what we should or do say in any scientific context. Hellman points out that one can appeal to useful divisions of labor within the sciences to motivate such a distinction. For example, a physicist who hypothesizes that heat is molecular motion (and regiments physical theories accordingly) isn't thereby making a revolutionary proposal about what higher-level scientists (biologists or ecologists) should say or a hermeneutic proposal about what they currently mean. So the untroubled friend of metaphysics can think about ontology as its own discipline, with its own level of analysis and corresponding explanatory work this analysis is invoked to perform. A Nominalist of this stripe might say: metaphysics is to physics as physics is to biology and ecology. That's why good proposals about what we

Arguably, this book's project of developing Potentialist foundations for set theory is analogous to the familiar project of providing a set-theoretic foundation for analysis. Our naive reasoning about certain concepts (limits in one case, the height of an iterative hierarchy of sets that "goes all the way up" in the other) turns out to lead to paradox. So, it is desirable to find a different way of thinking about relevant mathematical claims which will let us capture their intuitive significance and interest, while blocking paradoxical inferences.

Second, the Potentialist understanding of pure set theory advocated in Parts I and II of this book is compatible with a range of different views about how to understand other areas of mathematics. I hope my version of Potentialism will be compelling even to readers who find both Nominalism and the neo-Carnapian realism about mathematical objects (outside set theory) I advocate in Part III unacceptable.

Third, I aim to provide a foundation for Potentialist set theory which rests entirely on intuitively compelling principles that are subject matter neutral and constrain the behavior of all objects (c.f., Frege's characterization of logic in Frege (1980)). Thus, in a sense I'm defending a kind of Logicism about set theory. But I don't mean to claim that my foundational principles are analytic, cognitively trivial, or impossible for any rational being to doubt. I merely claim they're clearly true in the sense evoked by Putnam above.[7] I also don't mean to suggest that facts about conditional logical possibility discussed in this book constitute some kind of metaphysical free lunch.[8]

Fourth, we must distinguish the foundational project in this book from a less ambitious justificatory project. Actualist philosophers have sometimes aimed to find a unified conception of set theory from which all the various ZFC axioms clearly follow – without worrying whether this conception itself is clearly coherent. This project can be valuable in various ways, e.g., in showing the naturalness and appeal of certain mathematical hypotheses (like proposed large cardinal axioms) which also follow from the relevant conception. However, finding such a unified conception doesn't suffice for my foundational project. For if the unifying conception isn't clearly *consistent* then, surely, it isn't clearly *true* (even on a view which allows

should start to say in philosophy journals can differ radically from what physics journals would or should publish. Perhaps Sider's distinction between metaphysical semantics and linguistic semantics discussed in Section 11.4.2 suggests a similar line of response (Sider 2011).

However, the motivations I urge for Potentialist set theory are closer to those for foundational projects within mathematics than the explicitly philosophical motivations Hellman and Sider reference. Thus, I think the Potentialist paraphrases I advocate might be accepted by extreme naturalist readers, who would reject the above suggestion that philosophy or metaphysics could provide a legitimate further level of analysis beneath the sciences. Also note that the motivations for Potentialist set theory I press in this text aren't among the specific philosophical motivations for Nominalist formalizations of mathematics which Burgess and Rosen (1997) criticize.

[7] I take the axiom of choice to be prima facie clearly true, despite the fact that it can be doubted on grounds like the Banach–Tarski paradox. But readers who find Choice less immediately appealing can read this as a claim to justify Replacement from principles "as prima facie obvious as the other axioms of *ZF* set theory" instead.

[8] I take accepting a primitive modal notion of (conditional) logical possibility to be a significant, but warranted, addition to our fundamental ideology.

https://doi.org/10.1017/9781108992756.001 Published online by Cambridge University Press

mathematicians to introduce arbitrary coherent structures). So, we haven't succeeded in justifying all theorems of ZFC set theory from premises that seem clearly true.

Finally, the set-theoretic foundational project of this book also differs from a *more* ambitious project, along the following lines. Philosophers sometimes seek the most metaphysically joint-carving successor to the naive concept of sets which generates Russell's paradox (something which might, e.g., be hoped to connect intimately with the true answer to the liar paradox). So, for example, you could ask whether the iterative hierarchy conception of sets is remotely on the right track, or whether the "best" successor to naive set theory is something like Quine's New Foundations instead.

I think this more philosophically speculative project is legitimate, but not required for the foundational justificatory project attempted in this book. I will try to show that Potentialist translations which attractively explicate theorems of current mainstream set theory follow from principles that seem clearly true. But I won't take a position on what the most metaphysically illuminating successor to naive set theory is.