THE BULLETIN OF SYMBOLIC LOGIC
Volume 26, Number 2, June 2020

A RECONSTRUCTION OF STEEL'S MULTIVERSE PROJECT

PENELOPE MADDY AND TOBY MEADOWS

Abstract. This paper reconstructs Steel's multiverse project in his 'Gödel's program' (Steel, 2014), first by comparing it to those of Hamkins (2012) and Woodin (2011), then by detailed analysis what's presented in Steel's brief text. In particular, we reconstruct his notion of a 'natural' theory, describe his multiverse axioms and his translation function, and assess the resulting status of the Continuum Hypothesis. In the end, we reconceptualize the defect that Steel thinks *CH* might suffer from and isolate what it would take to remove it while working within his framework. As our goal is to present as coherent and compelling a philosophical and mathematical story as we can, we allow ourselves to augment Steel's story in places (e.g., in the treatment of Amalgamation) and to depart from it in others (e.g., the removal of 'meaning' from the account). The relevant mathematics is laid out in the appendices.

The stubborn recalcitrance of some independent set-theoretic statements, most prominently the Continuum Hypothesis (*CH*), and the proliferation of powerful techniques for generating new models have led some observers to champion a stark revision in our understanding of the set-theoretic project: the goal isn't to develop a theory, as complete as possible, describing a single universe of sets; rather, the target is an array of universes, a multiverse. Several such theories have been proposed, and the general idea is now prevalent enough to have made its way into the prose of at least one textbook (Weaver, 2014). To take the example of *CH*, most such theories posit an array of universes with *CH* true in some and false in others, which is taken to show that it has no determinate truth value, that efforts to settle it definitively, one way or the other, are misguided. Against this backdrop, John Steel's approach is particularly intriguing: he offers his multiverse theory instead as a means toward assessing *CH*, of exploring whether or not it's defective, whether or not the old enterprise of attempting to settle it is in fact viable. It's this undertaking of Steel's that we intend to examine here.

Steel's presentation of the motivations, structure, and current status of his multiverse project appears in condensed form in his paper 'Gödel's Program' (Steel, 2014). Our goal is to tell as coherent and compelling a philosophical story as we can while capturing what we take to be the spirit of Steel's

Received December 18, 2018.
2020 *Mathematics Subject Classification*. 00A30, 00A35, 03A05, 03E50, 03E55, 03E65.
Key words and phrases. philosophy of mathematics, foundations of mathematics, set theory, continuum hypothesis, large cardinals, forcing, generic multiverse.

1079-8986/20/2602-0002
DOI:10.1017/bsl.2020.5

enterprise. Often this involves some filling in,[1] some elaboratation,[2] and some outright departures, especially around his appeals to meaning and synonymy,[3] but also occasionally in the mathematics. In the end, we can't claim to have made every turn of the argument entirely air-tight, so we conclude with a brief discussion of a few lingering questions.

We begin in Section 1 with a sketch of the historical background against which multiverse thinking first emerged. Section 2 introduces Steel's approach by contrasting it with those of Hamkins (2012) and Woodin (2011). The central notion of 'natural theory' is examined in Section 3, and Steel's multiverse theory itself is presented in Section 4. Section 5 explores the relationship between the multiverse language and that of ordinary set theory. *CH* is treated in Section 6, before the concluding Section 7.

§1. Historical background. In 1878, soon after proving that there are more reals than naturals, Cantor asked 'into how many and what [cardinality] classes do [infinite sets of reals] fall?'[4] He famously conjectured that the answer is 'two' – the Continuum Hypothesis – which he reformulated in 1883 to the claim that the reals have the cardinality of the set of countable ordinals, and in the 1890s to the now-standard $2^{\aleph_0} = \aleph_1$. Cantor may have hoped to prove *CH* by doing so for closed sets and generalizing from there, but this method was doomed.[5] At the famous international congress in Heidelberg in 1904, König claimed to have disproved *CH* by showing that the reals can't be well-ordered, but by the next day Zermelo had found an error in the proof.[6] Hilbert's well-known 1925 paper, 'On the infinite', included an attempted proof that *CH* is true.[7] Other less well-known efforts to resolve *CH* were similarly unsuccessful.[8]

---

[1]E.g., we take the discussion of natural theories in Section 3 and the emphasis on foundational theories to motivate *MV* at the beginning of Section 4 to fill in Steel's line of thought in Sections 2 and 3 and Section 5 of Steel (2014), and the description of a possible route back to a universe theory in Section 6 to fill in his line of thought in Sections 5 and 6.

[2]E.g., the explicit appeal to axiomatizability in defense of Amalgamation in Section 4. (The key Theorem 34 was provided by Woodin in response to our query.) See also footnote 77.

[3]E.g., the replacement of 'settled by the meaning currently assigned to $\mathcal{L}_\in$' with 'impartiality' and the replacement of 'synonymous with $t(\varphi)$ for some φ in $\mathcal{L}_{MV}$' with 'legitimate $_T$' in Section 5.

[4]Translated by Jourdain in the introduction to his translation of Cantor's articles of 1895 and 1897 (Cantor, 1952, p. 45).

[5]The idea was to generalize the Cantor-Bendixson theorem, which says that any uncountable closed set of reals has a perfect subset (and hence has the size of the continuum). Unfortunately, as Bernstein showed in 1908, this can't work, because the Axiom of Choice implies that the reals can be decomposed into two uncountable sets, neither of which contains a perfect subset.

[6]This inspired Zermelo to formulate the Axiom of Choice and to establish on that basis that the continuum can be well-ordered.

[7]Hilbert (1967). Van Heijenoort's introduction to the paper describes the relation between Hilbert's attempted proof and Gödel's later proof of the relative consistency of *CH*.

[8]See Moore (1989).

The forces behind this impasse were gradually revealed: in the 30s, Gödel used the inner model L to show that *ZFC* (if consistent) can't disprove *CH*, and in the 60s, Cohen used his new technique of forcing to show that *ZFC* (if consistent) can't prove it either. Foreseeing Cohen's result as early as the 40s, Gödel proposed a search for new axioms, leading with the suggestion of large cardinal axioms. At the time, inaccessible cardinals and Mahlo cardinals were the best on offer, and Gödel recognized that 'there is little hope of solving [*CH*] by means of ... axioms of infinity' like these, because, for example, his proof of 'the undisprovability of [*CH*] goes through for all of them without any change' (Godel, 1990, p. 182). By 1964, he held out some hope for large cardinals 'based on different principles' (Godel, 1990, p. 261) – a footnote discusses measurables – but a postscript added in 1966, in light of Cohen's work, notes that 'it seems to follow that the axioms of infinity mentioned in [the] footnote ... are not sufficient to answer the question of the truth or falsity of Cantor's continuum hypothesis' (ibid., p. 270). Lévy and Solovay confirmed this for all standard large cardinals in Lévy and Solovay (1967).

A new species of axiom candidate emerged in the late 60s, using the notion of determinacy.[9] Determinacy was quickly shown to imply other, more familiar regularity properties – if all sets of reals are determined (AD), then they're also Lebesgue measurable and have the Baire and perfect subset properties – while the Axiom of Choice guaranteed the existence of an undetermined set. To preserve Choice, interest focused on positing the determinacy of definable sets: the projective sets (*PD*) or the sets constructible from $\mathbb{R}$ ($\mathrm{AD}^{L(\mathbb{R})}$). These hypotheses settled questions of descriptive set theory that had been open since the 20s and that Godel's and Cohen's techniques showed could not be answered from *ZFC* alone. The perceived downside was their lack of intrinsic support: 'No one claims direct intuitions ... either for or against determinacy hypotheses' (Moschovakis, 2009, p. 472). This shortcoming was remedied in the late 80s, when Martin, Steel, and Woodin derived *PD* and $\mathrm{AD}^{L(\mathbb{R})}$ from large cardinal axioms. Unfortunately, it was known even before this that determinacy assumptions can't settle *CH*.[10]

This long history of failure to settle *CH* has led some observers to despair and some skeptics to press their advantage:

> The striking thing, despite all such progress, is that – contrary to Gödel's hopes – the Continuum Hypothesis is *still* completely undecided ... That may lead one to raise doubts not only about Gödel's program but its very presumptions. Is the Continuum Hypothesis a definite problem as Gödel and many current set theorists believe? (Feferman, 2000, pp. 404–405)

[9] A subset A of Baire space (${}^{\omega}\omega$) is determined iff one or the other player has a winning strategy in an infinite game in which they alternate choosing natural numbers and the first player wins iff the result is in A.

[10] See Steel (2016) for the history.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Even believers in the determinateness of *CH* admit that the skeptics have a point:

> Those who argue that the concept of set is not sufficiently clear to fix the truth value of *CH* have a position which is at present difficult to assail. As long as no new axiom is found which decides *CH*, their case will continue to grow stronger, and our assertion that the meaning of *CH* is clear will sound more and more empty. (Martin, 1976, pp. 90–91)[11].

In other words, perhaps it isn't that our methods have failed to crack *CH*, but that the problem itself is somehow ill-formed.

This is the state of affairs that inspires multiverse thinking. Maybe it's a mistake to pursue a unified theory of a single domain of sets; maybe we should allow for a range of theories describing a range of domains. Instead of doggedly demanding an answer to the Continuum Problem, maybe we're failing to recognize the solution that's right before our eyes:

> The answer to *CH* consists of the expansive, detailed knowledge set theorists have gained about the extent to which it holds and fails in the multiverse, about how to achieve it or its negation in combination with other diverse set-theoretic properties ...the most important and essential facts about *CH* are deeply understood, and these facts constitute the answer to the *CH* question. (Hamkins, 2012, p. 429)

Various of these themes appear throughout the multiverse literature. We now sketch three approaches, due to Hamkins, Woodin, and Steel.

§2. Motivation. A first step toward understanding Steel's multiverse language and theory is to recognize that his motivation is different from those of other multiverse theorists, most prominently Hamkins and Woodin. All three are concerned with *CH* in one way or another, all three engage in multiverse thinking, but they do so for quite different reasons, with quite different metaphysics and methods, and quite different outcomes. Our hope is that highlighting these differences will bring all three into sharper focus.

Hamkins's case for his multiverse is grounded in the phenomenology of set-theoretic practice; his multiverse is posited to account for that mathematical experience:

> Our most powerful set-theoretic tools, such as forcing, ultrapowers, and canonical inner models, are most naturally and directly understood as methods of constructing alternative set-theoretic universes. ... we have a robust experience in these worlds ... The multiverse view ... explains this experience by embracing them all as real. (Hamkins, 2012, p. 418)

[11] Martin is more optimistic in his paper for Koellner's EFI project (Martin, 2019).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

> We seem to have discovered the existence of other mathematical universes ... and the multiverse view asserts that yes, indeed, this is the case. (Ibid., p. 425)

This generates a rich platonistic metaphysics:

> The multiverse view is one of higher-order realism – Platonism about universes. (Ibid., p. 417)

It includes worlds for many different set theories, both weak and strong:[12]

> There seems to be no reason to restrict inclusion to only ZFC models, as we can include models of weaker theories ZF, ZF^{-}, KP, and so on, perhaps even down to second-order number theory. (Ibid., p. 436)

Hamkins rejects any call for an explicit axiomatization, but he does identify certain principles that 'we might expect to find in the multiverse' (ibid., p. 436). (It isn't clear what sort of epistemic access we have to these principles, what reason we have to think the platonic multiverse has these features.) Finally, as we've seen, Hamkins takes the status of CH to be resolved:

> On the multiverse view ... the continuum hypothesis is a settled question; it is incorrect to describe CH as an open problem. (Ibid., p. 429)

In sum, then, the line of thought goes like this: the phenomenology of set-theoretic practice is explained by, and therefore justifies, a generous abstract ontology; some facts about this abstract ontology provide a final answer to the CH.

Woodin's concerns are in some ways orthogonal to Hamkins's. He isn't out to explain set-theoretic experience, but to block what he sees as a way of denying that CH has a determinate truth value:

> Refinements of Cohen's method of forcing in the decades since his initial discovery of the method and the resulting plethora of problems shown to be unsolvable ... have ... almost compelled one to adopt the generic-multiverse perspective. (Woodin, 2011, pp. 16–17)
>
> Let the *multiverse* (of sets) refer to the collection of possible universes of sets. The truths of ...Set Theory are the sentences which hold in each universe of the multiverse. The multiverse is the *generic-multiverse* if it is generated from each universe of the collection by closing under generic extensions (enlargements) and under generic refinements (inner models of a universe which the given universe is a generic extension of). (Woodin, 2011, p. 14)

Clearly, the generic multiverse is less encompassing than Hamkins's wild menagerie, including as it does only generic extensions and refinements.

[12] Philosophers of mathematics may be reminded of the Plenitudinous Platonism of Balaguer (1998).

Since a set-theoretic claim is to be true in this multiverse just when it's true simpliciter in all its worlds, we see that CH is neither true nor false.

To block this conclusion, Woodin focuses on the notion of multiverse truth and shows (modulo a proper class of Woodin cardinals and the Ω-conjecture) that its Π_2 truths are Turing reducible to the truths of an initial segment that's uniformly definable in any one of its universes; this reduction, he claims, is inconsistent with 'the very nature of [the] conception' of set-theoretic truth (Woodin, 2011, p. 17).[13] (Here, once again, it's unclear what sort of epistemic access is involved.) In this way, Woodin claims to undercut the generic-multiverse conception of truth, and with it, the purported challenge to the determinateness of CH.

Different as they are, there's a sense in which Hamkins and Woodin stand together on one side of a divide that separates them both from Steel. For the two of them, the multiverse promise (Hamkins) or threat (Woodin) is that the pretheoretic subject matter of set theory isn't a single universe but an array of universes.[14] From this multiverse perspective, the language of set theory is still $\mathcal{L}_\in$, but because it's understood as describing a different pretheoretic metaphysics, some statements in that language – CH most conspicuously – have a new status. In stark contrast, Steel bypasses any pretheoretic metaphysics. For Steel, the question at issue is whether the language of set theory should be $\mathcal{L}_\in$, a language of sets, or a multiverse language of sets and universes. As we'll see (in Section 4), the pretheoretic subject matter guiding the formulation of Steel's multiverse language and multiverse theory is again just theories, first-order theories in the language $\mathcal{L}_\in$ of set theory.

It should be noted that this philosophically consequential contrast between Hamkins and Woodin on one side and Steel on the other is compromised when Steel formulates the central question about CH in this way: is 'the truth value of CH ... determined by the meaning we currently assign to' $\mathcal{L}_\in$? (Steel, 2014, p. 154).[15] Here 'the current meaning of $\mathcal{L}_\in$' looks to function as a new sort of pretheoretic subject matter.[16] In what follows, we present this as Steel's preferred way of identifying the defect CH might suffer from – not being settled by the current meaning assigned to $\mathcal{L}_\in$ – but eventually we argue not only that this is inessential, but that it in fact clashes with the central mechanisms of Steel's position. We offer a replacement that's better suited to the job and that removes any hint of back-sliding in the direction of Hamkins and Woodin.

---

[13]See Meadows (2020) for an assessment of this argument.

[14]Koellner employs a more general terminology of 'pluralism' and 'nonpluralism', which disagree on whether or not 'there is an objective [unique?] mathematical realm' (Koellner, 2014, paragraph 2).

[15]This is an improvement over the formulation of Steel's question in Section 3 of Maddy (2017): 'is CH meaningful?'. To suggest that CH might be meaningless in any ordinary sense is a nonstarter, but the possibility that a meaningful statement might lack a truth value is not.

[16]This begins to resemble contemporary versions of conceptualism. See, e.g., the quotation from Martin on p. 5.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Steel's focus, then, is linguistic, that is, on theories, and in particular on 'framework theories', that is, theories suited to the traditional foundational role:[17]

> Why not simply develop all natural theories ... ? Let 1000 flowers bloom! ... The problem with this ... is that we do not want everyone to have his own private mathematics. We want one framework theory, to be used by all, so that we can use each other's work. It is better for all our flowers to bloom in the same garden ...
>
> The goal of our framework theory is to *maximize interpretive power*, to provide a language and theory in which all mathematics, of today and of the future so far as we can anticipate it, can be developed. (Steel, 2014, pp. 164–165)

Given the overarching goal of 'maximizing interpretive power',[18] Steel takes large cardinals to be a good start on how to proceed beyond ZFC: they provide an effective measure of consistency strength; there's good evidence for their consistency, especially for those with canonical inner models (Steel, 2014, pp. 156, 164). The question, for him, is how we go on from there, and what bearing this has on the meaning of set-theoretic language, and hence on the determinacy of CH.

The key to Steel's answer is his contention that the natural theories aren't a chaotic collection, that

> In fact, the different natural theories ... are not independent of one another. (Steel, 2014, p. 164)

His goal, then, is to give all these natural theories fair and equal consideration:

> We seek a language in which all these theories can be unified, without bias toward any, in a way that exhibits their logical relationships ... We want a neat package they all fit into. (Steel, 2014, p. 165)

In this way, Steel hopes to address what is for him the central question – is CH settled by current set-theoretic meaning? – without prejudging the answer. To understand how this goes, we must first understand natural theories and the sense in which they aren't 'independent of one another'.

[17]Steel treats the terms 'foundational theory' (Steel, 2014, p. 154) and 'framework theory' (ibid., p. 164) interchangeably. For more on the foundational aspects of universe and multiverse theories, see Maddy (2017).

[18]Steel distinguishes 'interpretive power' from 'consistency strength': 'Maximizing interpretive power entails maximizing consistency strength, but it requires more, in that we want to be able to translate other theories/languages into our framework theory/language in such a way that we preserve their meaning. The way we interpret set theories today is to think of them as theories of inner models of generic extensions of models satisfying some large cardinal hypothesis, and this method has had amazing success' (Steel, 2014, p. 165). This preference for so-called 'meaning preserving interpretations' is implicit in the analysis of 'natural theories' in Section 3 below.

§3. Natural theories. To begin at the beginning, what are 'natural' theories? Understandably, Steel isn't precise about this, but he does give us a hint:

> By 'natural' we mean considered by set theorists, because they had some set-theoretic idea behind them. (Steel, 2014, p. 157)

We might say a natural set theory is one with a serious mathematical motivation. This is a broader class of theories than framework theories – foundational theories in which 'all mathematics ... can be developed' – but presumably all framework theories are natural. On the other extreme, unnatural theories would include those 'using self-referential sentences, for example' (ibid.) or what we might call 'Gödelian trickery' (e.g., $ZFC + \neg Con(ZFC)$). This obviously isn't enough to firmly delimit the class of natural theories, but the intention behind the notion should be clear enough. The claim, then, is that all such theories are interrelated. We can see this, Steel tells us, in 'logical relationships ... brought out in our relative consistency proofs' (ibid., p. 164).

The reference here is to the proofs involved in establishing that the hierarchy of large cardinal axioms provides an apt measure of consistency strength. What's emerged over the years is that many theories set theorists consider turn out to be equiconsistent with *ZFC* extended by one large cardinal axiom or another. Moreover, these large cardinal axioms are linearly ordered by their consistency strength. Of course it's possible to concoct a theory for which this fails, but as a straightforward matter of empirical fact, it has been true for 'natural' theories entertained to date. So, for example:

Theorem 1.

(1) $Con(ZFC + \text{ projective sets are Lebesgue measurable}) \leftrightarrow Con(ZFC + \exists \text{inaccessible cardinal})$.
(2) $Con(ZFC + \omega_1 \text{ has a precipitous ideal}) \leftrightarrow Con(ZFC + \exists \text{measurable cardinal})$.
(3) $Con(ZFC + \Delta^1_2\text{-determinacy}) \leftrightarrow Con(ZFC + \exists \text{Woodin cardinal})$.
(4) $Con(ZF + AD) \leftrightarrow Con(ZFC + \exists \text{infinitely many Woodin cardinals})$.[19]

The pervasiveness of this phenomenon has led to the widespread belief that the consistency strength of all natural theories can be measured by the large cardinal hierarchy:

> Phenomenon 1. Every natural theory extending *ZFC* is equiconsistent with a theory of the form $ZFC + LCA$, where *LCA* is some large cardinal axiom.

Steel is calling attention to the nature of the proofs involved in establishing these equiconsistencies because, as a matter of fact, they all follow a

[19] (1) is due to Solovay (1970) and Shelah (1984). (2) is due to Jech et al. (1980). (3) and (4) are due to Woodin (see Woodin and Koellner's chapter in Foreman and Kanamori, 2009).

certain important pattern. If T is a natural theory extending *ZFC* and Φ is the relevant large cardinal axiom, we proceed roughly as follows: in one direction, using Cohen's forcing technique, we start from a countable model $\mathcal{M}$ of $ZFC + \Phi$ and define a poset $\mathbb{P}$ in $\mathcal{M}$ such that whenever G is $\mathbb{P}$-generic over $\mathcal{M}$, $\mathcal{M}[G]$ thinks T, that is, such that $\mathbb{P}$ forces T; in the other direction, starting from a model $\mathcal{M}$ of T, we define an inner model $\mathcal{N}$ such that $\mathcal{N}$ thinks $ZFC + \Phi$.[20]

The prevalence of this form of proof suggests:

Phenomenon 2. For every natural theory extending *ZFC*, there's an *LCA* such that the $ZFC + LCA$ proves that theory holds in an inner model or a forcing extension.

The significance of Phenomenon 2 is precisely what Steel claims: it shows how natural theories are interrelated. To see this, suppose T and S are theories extending *ZFC* that are connected in the way these equiconsistency proofs require, that is, suppose that in any model of S, we can define a $\mathbb{P}$ that forces T and that any model of T contains an inner model of S. Then it's a straightforward consequence of the Lévy-Shoenfield absoluteness theorem that they have the same Σ^1_2 consequences:

Theorem 2. *Suppose S and T are theories extending ZFC such that*:

(1) *S proves that there is some $\mathbb{P}$ such that $\Vdash_{\mathbb{P}} T$.*
(2) *Every model $\mathcal{M}$ of T has an $\mathcal{M}$-definable inner model $\mathcal{N}$ with the same ordinals that satisfies S.*

Then $T =_{\Sigma^1_2} S$; i.e., T and S have the same Σ^1_2 sentences as consequences.

Proof. First recall that the Lévy-Schoenfield theorem[21] tells us that if ψ is Σ^1_2, $\mathcal{M}, \mathcal{N}$ are models of ZFC and $\mathcal{M}$ is an inner model of $\mathcal{N}$, then

$$\mathcal{M} \models \psi \Leftrightarrow \mathcal{N} \models \psi.$$

Let ψ be Σ^1_2. We show that $T \models \psi$ iff $S \models \psi$.

($\rightarrow$) Suppose $T \models \psi$ and let $\mathcal{M}$ be a countable model of S. We claim that $\mathcal{M} \models \psi$. To see this let G be $\mathbb{P}$-generic over $\mathcal{M}$. (1) tells us that $\mathcal{M}[G] \models T$ and so $\mathcal{M}[G] \models \psi$. Then since $\mathcal{M}$ is an inner model of $\mathcal{M}[G]$, we see that $\mathcal{M} \models \psi$.[22]

[20]For example, in (3) of Theorem 1, we take a countable model $\mathcal{M}$ of *ZFC* with a Woodin cardinal δ and then show that if we collapse δ using a G that's $\mathrm{Col}(\omega,\delta)$-generic over $\mathcal{M}$, we obtain a model $\mathcal{M}[G]$ in which Δ^1_2-determinacy is true. In the other direction, we take a model of *ZFC* in which Δ^1_2-determinacy is true and show that in a definable inner model $\mathcal{N}$, the HOD of $\mathcal{N}$ thinks $\mathcal{N}$'s ω_2 is Woodin. See Neeman's chapter in Foreman and Kanamori (2009) for a detailed account of the forward direction, and Koellner and Woodin's following chapter for the converse.

[21]For a detailed account of this theorem, see Theorem 25.20 of Jech (2003) or Theorem 13.15 of Kanamori (2003).

[22]Note that we are forcing over models of *ZFC* that might not be well-founded. This is well known to be harmless (see Corazza, 2007 for a comprehensive account of forcing in such situations).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

($\leftarrow$) Suppose $S \models \psi$ and let $\mathcal{M}$ be a countable model of T. We claim that $\mathcal{M} \models \psi$. By (2), we may fix an $\mathcal{M}$-definable inner model $\mathcal{N}$ of $\mathcal{M}$ such that $\mathcal{N} \models S$. By our assumption we see that $\mathcal{N} \models \psi$; and since $\mathcal{N}$ is an inner model of $\mathcal{M}$ we see that $\mathcal{M} \models \psi$. $\dashv$

Given Phenomenon 2, it follows that no natural theories extending *ZFC* can disagree about Σ^1_2 statements:

If T, S are natural theories extending *ZFC*, then either

$$T \subseteq_{\Sigma^1_2} S \text{ or } S \subseteq_{\Sigma^1_2} T.$$

The idea here is that even if T and S are not equiconsistent, they will each be equiconsistent with some large cardinal, so one of the directions from Theorem 2 is always available. More specifically, exploiting Phenomenon 1, suppose T and S have been shown to be equiconsistent via forcing or an inner model to large cardinal axioms LC_T and LC_S respectively. Then since large cardinal axioms are linearly ordered by consistency strength, we may suppose without loss of generality that LC_T interprets LC_S via an inner model. By repeated use of Phenomenon 2 and Theorem 2 we then see that: T and LC_T have the same Σ^1_2 consequences; LC_T has possibly more Σ^1_2 consequences than LC_S;[23] and LC_S and S have the same Σ^1_2 consequences. This means that natural theories extending *ZFC* cannot disagree about Σ^1_2 sentences.[24]

And this style of consequence continues:[25]

For any two natural theories extending	$ZFC + \forall x\, x^{\#}$ exists	we have	$T \subseteq_{\Sigma^1_3} S$	or	$S \subseteq_{\Sigma^1_3} T$
	ZFC+ infinite Woodins		$T \subseteq_{\Sigma^1_\omega} S$		$S \subseteq_{\Sigma^1_\omega} T$
	ZFC+ infinite Woodins & Measurable above		$T \subseteq_{Th(L(\mathbb{R}))} S$		$S \subseteq_{Th(L(\mathbb{R}))} T$

[23]For this we rely on the fact that the known large cardinals are linearly ordered by consistency strength. Moreover the proof that the consistency of LC_T implies the consistency of LC_S can – at worst – be established by defining a model M of LC_S from a model N of LC_T and both of these models can be understood has having the same ω_1. This then suffices for a further use of the Lévy-Schoenfield theorem.

[24]Note that it is crucial that T and S are natural theories. For a pathological example, observe that $ZFC + \neg Con(ZFC)$ is equiconsistent with ZFC but it clearly doesn't agree with ZFC on all Σ^1_2 sentences.

[25]By "infinite Woodins," we mean infinitely many Woodin cardinals.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

The proofs of specific cases of this phenomenon also follow the template above.[26] The upshot is that as we add more large cardinal axioms, we remove the possibility of disagreement between natural theories extending them:

> (*) As natural theories proceed up the large cardinal hierarchy in consistency strength, they agree on an ever-increasing class of mathematical statements.

It's worth recalling that these ever-increasing classes have their origin in the work of the French analysts Baire, Borel, and Lebesgue in the early years of the twentieth century. Alarmed by the role of pathological functions in the foundations of analysis in the late nineteenth century, they set out to bring order to the study of functions from reals to reals by classifying them according to their complexity. The process began with the Borel hierarchy ($\{\mathbf{\Sigma}^0_\alpha\}_{\alpha<\omega_1}$), where the complexity of functions is reduced to that of sets (for example, a function is Borel iff the inverse image of every Borel set is Borel). Complexity for sets of reals is then defined in familiar topological terms, and as hoped, the Borel sets turned out to be fairly well-behaved, enjoying regularity properties like Lebesgue measurability and the perfect set property. The $\mathbf{\Sigma}^1_n$s involved in (*) arose as the effort to domesticate parts of analysis continued in the Russian school of Lusin and Souslin. The regularity properties were extended to $\mathbf{\Sigma}^1_1$, but stalled at the perfect subset property for $\mathbf{\Pi}^1_1$ and Lebesgue measurability for $\mathbf{\Sigma}^1_2$. Unbeknownst to Lusin and Souslin, their failures weren't from lack of imagination: *ZFC* isn't enough to settle these matters. With the introduction of determinacy hypotheses, eventually derived from large cardinals, regularity was extended to the entire projective hierarchy. Thus the original goal of delimiting the more civilized, more well-behaved portion of analysis was extended.

So these classification hierarchies of Borel and projective sets of reals originated in an effort to isolate the more straightforward, down-to-earth portion of analysis; Steel refers to statements involving these sets as 'concrete'. In these terms, (*) becomes:

> PHENOMENON 3. As natural theories proceed up the large cardinal hierarchy in consistency strength, they agree on an ever-increasing class of concrete mathematical statements.

---

[26] For example, in the cases where every set has a sharp, we show that there is a kind of generalized proof theory for the Σ^1_3 sentences in the sense that their truth can be witnessed by the ill-foundedness of a certain tree. In this case, the Martin-Solovay tree suffices (see Chapter 15 of Kanamori, 2003). For the other two, we rely on a generalisation of this known as a homogeneous tree (see Chapter 32 of Kanamori, 2003 and Neeman's chapter in Foreman and Kanamori, 2009). One might also think of the Lévy-Shoenfield tree used in Theorem 2 as providing a kind of proof theory for Σ^1_2 facts. One might think of the ill-foundedness of such a tree as being analogous the existence of an open branch in a proof tree or tableau in first order logic. It turns out that – in the presence of sufficient sharps – this proof theory remains intact through forcing and inner model constructions. Thus if T and U are natural theories – so linked by either forcing or an inner model construction – then they agree about how this proof theory works. In this way, the Σ^1_3 sentences are preserved and disagreement is removed.

The trouble, of course, is that *CH* is immune to this kind of disagreement-removal: it isn't concrete; large cardinal axioms aren't enough. This is what raises the specter, for Steel, that 'the truth value of *CH* is not determined by the meaning we currently assign' to the language of set theory (Steel, 2014, p. 154). The language of ZFC all by itself might be luring us into asking questions with no answers. This possibility can't be ignored:

> Certainly we do not want to employ a syntax which encourages us to ask pseudo-questions, and the problem then becomes how to flesh out the current meaning, or trim back the current syntax, so that we can stop asking pseudo-questions. (Steel, 2014, p. 154)

As we'll see, Steel's multiverse language is his tool for this project. We focus first on the 'trimming' option and return to the 'fleshing out' toward the end of Section 6.

§4. Multiverse language and theory. With this understanding of natural theories in hand, we return to the motivation for Steel's multiverse. We have good reason to adopt $ZFC + LCs$,[27] but don't know how to go on from there. We've seen (in Section 3) that all natural theories will agree on concrete mathematics. Maybe this is all we should ask of our foundational or framework theory:

> Why not simply develop all natural theories? ... Let 1000 flowers bloom! (Steel, 2014, p. 164)

But we've also seen (in Section 2) that Steel rejects this solution on the grounds that it wouldn't provide a unified framework:

> We do not want everyone to have his own private mathematics. We want one framework theory, to be used by all, so that we can use each other's work. It is better for all our flowers to bloom in the same garden. If truly distinct frameworks emerged, the first order of business would be to unify them. (Ibid.)

The trouble is that our guiding principle – maximize interpretive power – has given out. So Steel suggests:

> Before we try to decide whether some such theory is preferable to the others, let us try to find a neutral common ground on which to compare them. We seek a language in which all these theories can be unified, without bias toward any. (Ibid., p. 165)

The trick then is to find such a neutral language.

Phenomenon 2 of the previous section suggests that all natural theories are realized in inner models or forcing extensions of models of $ZFC + LCA$ for some large cardinal axiom LCA. Rather than trying to codify the natural

[27] We use 'LCA', e.g., in Phenomena 1 and 2, as a stand-in for some particular cardinal axiom or other and 'LCs' as a rough term for traditional large cardinal axioms in general.

theories directly, we could concoct a collection of 'worlds' that manage to realize each of them, a collection of worlds, then, that's closed under inner models and forcing extensions.

But notice that for present purposes, we don't actually want to represent all natural theories by worlds in our neutral ground. The point of the exercise, after all, is the hope it might help us determine whether or not independent statements like CH are settled by the current meaning we assign to the language of set theory. In the past, questions of this sort have been answered by finding ways to extend our current list of axioms, as we did by adding large cardinal axioms to ZFC; this would be to determine that 'some such theory is preferable to the others' (quoted above). Given our current commitment to $ZFC + LCs$, then, the theories we're interested in, the candidates for the foundational role, are extensions of $ZFC + LCs$.

Unfortunately, there's no precise characterization of LCs, of what are often called 'traditional large cardinal axioms', so we can't just narrow the range of natural theories in our neutral ground by stipulating that all our 'worlds' must satisfy $ZFC + LCs$. But we could, at the very least, try to avoid including theories with 'antilarge-cardinal axioms', like $V = L$, which foreclose the addition of some large cardinals. Theories like this tend to be realized in definable inner models, so one expedient would be to resist closing our collection of worlds under those.[28] Inner models generated by forcing refinements don't have this drawback, so they can be included.

Still, though the natural theories realized by definable inner models aren't represented by worlds in this collection, Steel emphasizes that 'they are already there, we can talk about them in the multiverse language already' (Steel, 2014, p. 167). He means, of course, that they can be defined in a world. This is crucial, as Steel has explained elsewhere, given that the theories realized by those worlds are to be regarded as candidates for a foundation:

> It is a familiar but remarkable fact that all mathematical language can be translated into the language of set theory, and all theorems of 'ordinary' mathematics can be proved in ZFC. In extending ZFC, we are attempting to strengthen this foundation.
>
> ...In this light we can see why most set theorists reject $V = L$ as restrictive: adopting it restricts the interpretative power of the language of set theory. *The language of set theory as used by the believer in $V = L$ can certainly be translated into the language of set theory as used by the believer in measurable cardinals, via the translation*

[28]Steel doesn't explicitly make the argument of this paragraph and the previous, but it's clear from the fact that he requires all worlds in his multiverse to satisfy ZFC that he's not expecting all natural theories to be represented by worlds (e.g., $ZF + AD$ is left out). Late in the paper, he remarks that 'Our current understanding of the possibilities for maximizing interpretive power [i.e., what a foundational theory is supposed to do] has led us to one theory of the concrete, and *a family of theoretical superstructures for it*, *each containing all the large cardinal hypotheses*' (Steel, 2014, p. 178, emphasis added). This family is the multiverse, and speaking informally now, he indicates that each world ought include all large cardinals.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

> $\varphi \to \varphi^L$. There is no translation in the other direction. (Steel, 2000, p. 423, emphasis added)

In other words, part of what makes large cardinals preferable to $V = L$ is that $ZFC + V = L$ can be realized in a definable inner model,[29] so it's important that those inner models are 'already there'.

To this point, then, our collection of worlds is closed under generic extension and refinement. Unfortunately, this doesn't tell us as much as it might seem. Sets appear here only in some world or other, so given a world V and a poset $\mathbb{P}$ in V, the existence of sets $\mathbb{P}$-generic over V depends on which worlds exist. For that matter, the single world V by itself is 'closed under generic extension' in the sense that for every poset $\mathbb{P}$ in V and every G that's $\mathbb{P}$-generic over V, $V[G]$ is in the collection – but only because there aren't any such Gs! We can do a little better by stipulating that

> (Extension) Given a world V and $\mathbb{P} \in V$, there is a world U and a G in U such that G is $\mathbb{P}$-generic over V and $U = V[G]$,

but so far, this guarantees one $\mathbb{P}$-generic set G and one generic extension $V[G]$, for each V and $\mathbb{P}$. So the question arises, how many should there be in our multiverse?

One immediately appealing answer is: all of them! This is the answer Woodin intends, but just to say this in the intuitive setting where we're currently operating doesn't help: 'for every world V, every poset $\mathbb{P}$ in V, and every G that's $\mathbb{P}$-generic over V, there is a world U such that $U = V[G]$' doesn't determine how many such Gs there are. To properly convey what he has in mind – 'to illustrate the concept of the generic-multiverse' (Woodin, 2011, p. 14) – Woodin gives us a set-theoretic toy model: we start with a countable transitive model (ctm) M of ZFC (or $ZFC + LCA$) and form a collection of ctms by closing under generic extension and refinement.[30] Because these models reside in our background universe, all $\mathbb{P}$-generic Gs are available when we close under generic extension. Here Woodin relies on our understanding of that background universe to deliver this full array of generics.

This approach would seem to align with Steel's goals – such a generous array of generic extensions promises to help realize all candidate foundational theories. But it turns out that Steel has reason to resist Woodin's answer.[31] To see this, we need to look more closely at Woodin's toy model.

Definition 3. If M is a ctm of ZFC, $\mathbb{V}_M$ is the smallest collection of models containing M and such that:[32]

---

[29] And of course the natural theory $ZF + AD$, mentioned in the previous footnote, will be satisfied in an inner model in any world with sufficient large cardinals.

[30] We adopt the convention of denoting transitive models by M and arbitrary models by $\mathcal{M}$.

[31] We take up the question of just how good his reason is in Section 7, but for now our goal is to give the best formulation we can muster.

[32] For ease of exposition, we state this definition of Woodin's $\mathbb{V}_M$ and Definition 6 of Steel's M^G in terms of countable transitive models, but we actually intend a slight

(1) If $N \in \mathbb{V}_M$ and G is $\mathbb{P}$-generic over N for some $\mathbb{P} \in N$, then $N[G] \in \mathbb{V}_M$; and
(2) If $N \in \mathbb{V}_M$ and $N = N'[G]$ where G is $\mathbb{P}$-generic over N' for some $\mathbb{P} \in N'$, then $N' \in \mathbb{V}_M$.

$\mathbb{V}_M$ has the following striking feature:

THEOREM 4 (Woodin). *For M a ctm of ZFC, there exist Cohen reals c and d over M such that there is no ctm N of ZFC such that*:

$$M[c] \subseteq N \supseteq M[d],$$

where N has the same ordinals as M.[33]

We might say that the two extensions $M[c]$ and $M[d]$ can't be amalgamated. It would be difficult to specify exactly how often this happens – the phenomenon is not restricted to Cohen reals – but at least we can say that in Woodin's intuitive multiverse, the following claim is false:

(Amalgamation) If V and V' are worlds, then there exist posets $\mathbb{P}$ and $\mathbb{P}'$ in V and V', respectively, and a world U, a $G \in U$ that's $\mathbb{P}$-generic over V, and a $G' \in U$ that's $\mathbb{P}'$-generic over V', such that

$$V[G] = U = V'[G'].$$

For a precise account of what hangs on this, we need more machinery – we get to this in a moment – but first a rough and informal sketch. In addition to Extension, we've also been assuming that the multiverse is closed under generic refinement:

(Refinement) If V is a world and $V = U[G]$ where G is $\mathbb{P}$-generic over U for some $\mathbb{P}$ in U, then U is a world.

Obviously, Woodin's toy model satisfies both Extension and Refinement but not Amalgamation, so speaking loosely for now, it follows that

(i) Extension + Refinement doesn't imply Amalgamation.

Using a different toy model, we show below (p. 19) that

(ii) Extension + Refinement is consistent with Amalgamation.[34]

---

generalization of both to all countable models (see Appendix B for $\mathbb{V}_{\mathcal{M}}$ and Definition 25 of Appendix A for $\mathcal{M}^G$). The no-go theorem for Woodin's multiverse (p. 17 and Theorem 34 of Appendix B) holds on either version of the definition. The point of the generalization is that it enables theorems like 8 and 12 on Steel's multiverse, which capture its purely linguistic character (as opposed to metaphysical approaches involving an abstract ontology of universes). The generalized versions are used in the two completeness claims on p. 17 to allow a direct comparison.

[33]See Fuchs et al. (2015), Section 2, Observation 35 for a proof of this. Also note that this result holds for the more general case of arbitrary countable models.

[34]This follows directly from Theorem 26 in Appendix A.

So we see that Extension and Refinement alone are too weak for a viable multiverse theory: they aren't enough to settle even the elementary matter of Amalgamation.

How is the question of Amalgamation to be settled? Steel's thinking here comes out in his complaint that

> Neither Hamkins nor Woodin presented a language and a first-order theory in that language, both of which would seem necessary for a true foundational theory. (Steel, 2014, p. 170)

So far, the candidate foundational theories Steel has been out to bring together for comparison in the multiverse are universe theories in the first-order language of set theory, like $ZFC + LCs$. Presumably in this quotation, when Steel speaks of a foundational theory in a multiverse language, he's thinking of what would happen if our exploration of all candidate universe theories in the 'neutral common ground' of the multiverse were to conclude that no such theory is 'preferable' to the others as a foundation (Steel, 2014, p. 165).[35] In that eventuality, our current foundational theorizing about a single universe would be prompting us to pose questions with no answers, that is, to ask which of a range of candidate foundational theories is 'correct' where there is no correct or incorrect. Some change would be in order, perhaps to a multiverse theory as foundation, in which case, obviously, we'd need an explicit multiverse theory to replace $ZFC + LCs$. What Steel actually ends up proposing in that eventuality (see Sections 5 and 6) is that we 'trim back current syntax [i.e., the syntax of $\mathcal{L}_\in$], so that we can stop asking pseudo-questions' (ibid., p. 154), which for him roughly comes to the same as adopting a particular multiverse theory as our foundation. Either way, an explicit multiverse theory is essential to Steel's project. Speaking of Woodin's multiverse, he writes 'it is not at all clear what its theory would be' (ibid., p. 170). The question, then, is whether Woodin's multiverse can be suitably axiomatized.[36]

To address this question, we need to adopt a meta-mathematical perspective on Woodin's intuitive multiverse, and the only way we know to do this is to drop back into the theory we can agree on – $ZFC + LCs$ – and reason there about a set-theoretic surrogate. At this point, we again call on Woodin's $\mathbb{V}_M$, this time not as an intuitive guide to his intentions (as on p. 14), not as a simple tool for proving (i) (as on p. 15), but as a set-theoretic surrogate for his intuitive multiverse in our meta-mathematical inquiry. Any such meta-mathematical surrogate will be imperfect in some ways – even the original identifications of validity with truth in all set-theoretic models and of ordinary proof with a formal proof predicate have their infelicities – but without such surrogates, meta-mathematics is impossible.

[35]This leaves open the likelihood that some theories will be of more mathematical interest than others.

[36]See footnote 77.

The question, then, is whether there's a recursive set of axioms, T, in a suitable multiverse language such that for all φ in this language the following are equivalent:

(1) $T \vdash \varphi$.
(2) If $\mathcal{M}$ is a countable model of ZFC, then $\mathbb{V}_{\mathcal{M}} \models \varphi$.

Alas, the answer is no.[37]

This means that if Steel wants an axiomatizable multiverse, he can't follow Woodin and include 'all' generic extensions. In fact, a look at the proof of this no-go theorem reveals that it's Theorem 4, the violation of Amalgamation, that blocks the possibility of axiomatization.[38] So the natural move for Steel is to add Amalgamation to Extension and Refinement. Syntactically, Steel's multiverse theory, MV, builds on formal versions of these three assumptions. Semantically, he presents a toy model, $\mathcal{M}^G$, for some countable model $\mathcal{M}$ of ZFC,[39] much as Woodin gives us $\mathbb{V}_{\mathcal{M}}$, and that model, $\mathcal{M}^G$, appears as a meta-mathematical device in the proof of (ii), just as $\mathbb{V}_{\mathcal{M}}$ does for (i). It can then be proved that for all φ in the multiverse language, the following are equivalent

(1) $MV \vdash \varphi$.
(2) If $\mathcal{M}$ is a countable model of ZFC, then $\mathcal{M}^G \models \varphi$.

So Steel has reason to resist Woodin's approach and to answer the question 'how many generics' by adding Amalgamation to the characterization of his intuitive multiverse.

To make this rough story precise, we need a first-order language and theory. Steel proposes $\mathcal{L}_{MV}$, a two-sorted language with a sort for worlds and a sort for sets. (We reserve upper case letters like $V, U, V_0, \dots U_0, \dots$ for worlds and lower case letters like $x, y, z, x_0, \dots$ for sets.[40]) The language has a single relation symbol $\in$ and the atomic well-formed formulae include $x \in y$ and $x \in V$ but *not* $V \in x$. Steel formulates his multiverse theory, MV, in this language.

The first two axioms of MV codify our basic understanding of the multiverse's worlds and sets. At a minimum, worlds must be extensional and worlds must think all axioms of ZFC hold:

MV-0 $\forall V \forall U (\forall x (x \in V \leftrightarrow x \in U) \rightarrow V = U)$;
MV-1$_\varphi$ $\forall V \varphi^V$

[37] As remarked in footnote 32, we use the generalized definitions of both $\mathbb{V}_{\mathcal{M}}$ and $\mathcal{M}^G$ in these equivalences to maintain a strict parallelism. The no-go theorem for Woodin's multiverse is Theorem 34 of Appendix B.

[38] In his comparison of Woodin's multiverse with his own, Steel explicitly notes that the former 'does not satisfy amalgamation' (Steel, 2014, p. 170, footnote 22).

[39] Recall footnote 32.

[40] In a slight departure from Steel, we avoid W as world variable and reserve it for the class term in Theorem 5.

for any axiom φ of ZFC.[41] Steel points out that 'one can add large cardinal hypotheses that are preserved by small forcings ... as follows: given an large cardinal hypothesis φ, we add " φ^W for all worlds W"' (Steel, 2014, p. 166). Also, all worlds are transitive and have the same ordinals:

MV-2 $\forall V \forall x \in V\, x \subseteq V$;

MV-3 $\forall V \forall U \forall x (x \in On^V \leftrightarrow x \in On^U)$.

The fourth axiom guarantees that the only way sets appear in the multiverse is in worlds:

MV-4 $\forall x \exists V\, x \in V$.

The next group of axioms for MV specifies the structure of the multiverse by codifying the principles we've been treating informally. First Extension:

MV-5 (Extension) $\forall V \forall p \in V \exists U \exists g \in U (g \text{ is } p\text{-generic}/V \wedge U = V[g])$.[42]

Informally, this says just what it should: given a poset in some world V, there will be a V-generic g for that poset such that $V[g]$ is a world.

Stating Refinement is trickier, requiring the following theorem:[43]

Theorem 5 (NBG) (Laver, Woodin). *There is a class term $W_{(\cdot)} \in \mathcal{L}_\in$ such that the following are equivalent*:[44]

(1) *N is a generic refinement of the universe*; *and*
(2) *$N = W_r$ for some r.*

In other words, every generic refinement of a universe can be defined in that universe using a parameter. Refinement then becomes:

MV-6 (Refinement) $\forall V \forall r \in V \exists U (U = (W_r)^V)$.

Informally, this says that for any world V and parameter $r \in V$, there is a world corresponding to the generic refinement as calculated by the formula W_r in V.[45]

[41] The first of these was not included in the original axioms provided by Steel, however it is required. (Thanks to Gabriel Goldberg for this observation.) The second of these is – strictly speaking – an axiom schema.

[42] Contrary to ordinary practice we are using lower case letters for posets and generics to indicate that these are sets. We shall only do this in the *official* version of the axioms and when it will help avoid confusion. We write g is p-generic $/V$ to mean that p is a poset and g is a filter over p which has a nonempty intersection with every dense subset of p in V; and we write $U = V[G]$ to mean that for all $x \in U \exists \sigma \in V^p\ Val(\sigma, g) = x$ and for all $\sigma \in V^p \exists x \in U\ x = Val(\sigma, g)$. Val is a valuation function as defined in Kunen (2006) and V^p is the class of p-names according to V.

[43] For a detailed account of this theorem, see Reitz (2007).

[44] By a class term, we mean a term which defines a class and in this case takes a parameter as an argument. For another example, consider $L[x]$.

[45] Since Laver and Woodin's theorem is provable in ZFC, the class term $W_{(\cdot)}$ is available in any universe, V, of the multiverse.

The final axiom of MV is Amalgamation:

MV-7 (Amalgamation)

$$\forall U_0 \forall U_1 \exists p_0 \in U_0 \exists p_1 \in U_1 \exists V \exists g_0, g_1 \in V$$
$$(g_0 \text{ is } p_0\text{-generic}/U_0 \wedge g_1 \text{ is } p_1 \text{ -generic}/U_1 \wedge V = U_0[g_0] = U_1[g_1]).$$

Informally, this says that given any two worlds U_0, U_1, there is a world V which is a generic extension of both those worlds. The theory MV then consists of the schema $MV\text{-}1_\varphi$ and the axioms MV-0 and MV-2 through MV-7.

Like Woodin, Steel provides a toy model to illustrate the structure of his multiverse. Informally, the construction goes like this. Begin (as Woodin does) with a countable transitive model, M, of ZFC; take some poset $\mathbb{P} \in M$ and form a generic extension $M[G]$ of M; take some poset $\mathbb{Q} \in M[G]$ and form a generic extension $M[G][H]$; repeat this process transfinitely to form a sequence of models where each new model is a generic extension of all of its predecessors. Along the way, make sure that every poset from any model in the sequence is used cofinally often to generate new extensions.[46] Finally, add all generic refinements of models in this sequence. The result is M^G. From the construction, it's obvious that M^G satisfies Extension and Refinement. As for Amalgamation, notice that every world is a generic refinement of some witnessing world in the sequence, so it's also a generic refinement of every world further along the sequence. This means that any two worlds are generic refinements of witnesses somewhere in the sequence, so both will be refinements of whichever of these witnesses appears furthest along. (See Theorem 26 in the Appendix A for the proof.)

To formalize this idea, the universality of collapse forcing means that we needn't close under arbitrary posets, that those of the form $\mathrm{Col}(\omega, \alpha)$ for $\alpha \in M$ are enough. Since the collapse forcings absorb smaller forcings, we can find those models by looking back into the generic refinements. Thus we end up with the following definition:[47]

Definition 6. For M a ctm of ZFC and G, $\mathrm{Col}(\omega, {<}\mathrm{Ord}^M)$-generic over M, let M^G be the set of countable models N such that for some $\alpha \in \mathrm{Ord}^M$, $\mathbb{P} \in N$ and H, $\mathbb{P}$-generic over N

$$N[H] = M[G \restriction \alpha].$$

[46]Strictly speaking, we need to use iteration at the limit stages and we need to demand that the poset used at such a limit is a set in M. Then Theorem 34 of Fuchs et al. (2015) tells us that there is a generic extension of M which has all of the predecessors as generic refinements.

[47]For ease of exposition, we've described and expressed this in terms of ctms, but what follows actually employs a generalization, $\mathcal{M}^G$, defined in terms of all countable models. As explained in footnote 32, this adjustment, including the parallel change to $\mathbb{V}_{\mathcal{M}}$, makes no significant difference in Woodin's case and preserves the fundamentally linguistic character of Steel's project (as opposed to metaphysical approaches like those of Woodin and Hamkins). The particulars are spelled out in Appendices A and B.

Since $\mathcal{M}^G$ is a model of MV, it follows that Amalgamation is consistent with Extension and Refinement, as advertised in (ii) on p. 15, above. In fact, we're now in a position to show that MV is equiconsistent with ZFC:

PROPOSITION 7. $Con(ZFC) \leftrightarrow Con(MV)$.

PROOF. ($\rightarrow$) This follows from the proof of Theorem 26 in Appendix A.1. ($\leftarrow$) Suppose $ZFC \vdash \psi \wedge \neg\psi$ for some $\psi \in \mathcal{L}_\in$. Fix finite $\Delta \subseteq ZFC$ such that $\Delta \vdash \psi \wedge \neg\psi$. Then $MV \vdash \forall V\ (\bigwedge \Delta)^V$; and so $MV \vdash \forall V\ (\psi \wedge \neg\psi)^V$.[48] ⊣

So MV is a reasonable theory with natural models of the form $\mathcal{M}^G$.

At this point, recall the key drawback of Woodin's multiverse: the theory of his meta-mathematical surrogate, $\mathbb{V}_\mathcal{M}$, can't be axiomatized. We saw that for Steel this is disqualifying, because his goal is essentially linguistic: he wants to determine whether we should 'trim' the syntax of $\mathcal{L}_\in$ to avoid asking pseudo-questions; to accomplish this, he seeks to axiomatize the theory of his meta-mathematical surrogate, $\mathcal{M}^G$. We know that the axiom system MV is sound for models of the form $\mathcal{M}^G$, but we need for it to be complete, as well. And it is:

THEOREM 8. *For all φ in the multiverse language, the following are equivalent*:

(1) $MV \vdash \varphi$.
(2) *If $\mathcal{M}$ is a countable model of ZFC, then $\mathcal{M}^G \models \varphi$, where G is $Col(\omega, {<}Ord)^\mathcal{M}$-generic over $\mathcal{M}$.*

(See Appendix A.) So Steel has given a multiverse language and theory intended to characterize the range of candidate foundational theories and to realize his goal of a 'neutral ground on which to compare them ...without bias toward any' (Steel, 2014, p. 165).[49] We can now try to determine whether *CH* is indeed a pseudo-question.

§5. The translation function.

Recall from Section 2 that the central question for Steel is whether *CH* is settled by the meaning we currently assign to $\mathcal{L}_\in$. More generally, the worry is that a range of sentences of $\mathcal{L}_\in$ may be defective in this way, that attempting to answer them is chasing pseudo-questions. It's fairly easy to explicate this type of concern in a metaphysical theory like Hamkins's or Woodin's – there's an abstract ontology of worlds in some of which *CH* is true and others of which it's false – but we've seen that Steel's thinking is strictly linguistic. In that context, it's not obvious

[48] This proof is overly detailed for such a simple observation. However, we want to draw attention to the supervaluation-like aspect of the translation employed as it becomes significant later on.

[49] We raise a question about this in Section 7.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

how to characterize the potential problem without providing a substantive theory of meaning (no hint of which appears in Steel, 2014).[50]

Perhaps we can make more progress on characterizing the defect at issue by coming at the question from the other end, that is, by looking at the role it ends up playing in Steel's thought. Recall that in search of a 'neutral common ground' on which to compare all candidate foundational theories, he devises a multiverse language, $\mathcal{L}_{MV}$, and multiverse theory, MV – having previously argued that all such candidates will be represented by its worlds (the extended discussion of Section 3 and the opening pages of Section 4). Now suppose that none of these theories 'is preferable to the others' as a foundation (Steel, 2014, p. 165).[51] In that case, Steel contemplates two options: we could 'flesh out the current meaning' or we could 'trim back the current syntax, so that we can stop asking pseudo-questions' (Steel, 2014, p. 154). The option we're exploring here (until late in Section 6) is the latter. Given that all candidates are on equal footing, Steel's suggestion is that only the sentences of $\mathcal{L}_{MV}$ (and their synonyms) express propositions.[52] Assuming Steel's reference to 'expressing propositions' is a way of indicating that the sentences in question are capable of being true or false ('truth-apt' in the philosophical jargon), this appears to be a way of saying that the sentences of $\mathcal{L}_{MV}$ aren't subject to the defect Steel has in mind; to put it the other way around, they enjoy the virtue of being settled by the meaning assigned to $\mathcal{L}_{MV}$.

The parenthetical proviso – 'and their synonyms' – is included so that many of our familiar $\mathcal{L}_{\in}$ sentences will also enjoy this virtue, namely those synonymous with $\mathcal{L}_{MV}$ sentences and thus without defect. To isolate these favored sentences of $\mathcal{L}_{\in}$, Steel offers a translation from $\mathcal{L}_{MV}$ to $\mathcal{L}_{\in}$; sentences of $\mathcal{L}_{\in}$ in the range of that translation function mean the same as the corresponding $\mathcal{L}_{MV}$ sentences.[53] So, for example, Steel suggests that the $\mathcal{L}_{MV}$ claim that every world contains a measurable cardinal is synonymous

[50]It's not even clear what kind of theory would be called for. Obviously the everyday meaning of 'set' isn't what's at issue. If anything, a more limited community of trained set theorists would be the relevant target, but even if the boundaries of that group could be drawn in some principled way, would we turn to linguists or sociologists or anthropologists for answers? Most likely, some sort of a priori philosophical theory of meaning would be required. We consider it an attraction of the reconstruction described below that it avoids this prospect.

[51]This leaves open the likelihood that some theories will have more mathematical interest than others. (Recall footnote 35.)

[52]This appears to be the Weak Relativist Thesis of Steel (2014), p. 167.

[53]To be clear, Steel isn't suggesting that the $\mathcal{L}_{MV}$ sentence serves to confer meaning on the corresponding $\mathcal{L}_{\in}$ sentence via the translation function; rather the translation allows us to isolate from among the sentences of $\mathcal{L}_{\in}$, all of which are antecedently meaningful, (some of) those that are without defect. In fact, Steel goes further; he regards $\mathcal{L}_{MV}$ as meaningless syntax until the translation function is introduced. This seems to us problematic: it's hard to see how picking some meaningless syntax and mapping it somehow to $\mathcal{L}_{\in}$ could tell us anything significant about the sentences in its range. We take the discussion of Section 3 and the opening pages of Section 4 to provide a robust understanding what the formalization in $\mathcal{L}_{MV}$ and MV is intended to capture.

with the $\mathcal{L}_\in$ claim that there is a proper class of measurable cardinals, so this $\mathcal{L}_\in$ claim is not defective.[54] Still, he continues,

> Clearly we cannot state the *CH* in this way. The same goes for the many other statements about the uncountable which are sensitive to set forcing, no matter what large cardinals there may be. (Steel, 2014, p. 167)

But this isn't the end of the story. *CH* may still not be defective, because there may be 'traces of *CH* and these other sentences in the multiverse language' (ibid.) – that is, there may be other sentences of $\mathcal{L}_{MV}$ with which they're synonymous.

Steel goes on to explore this possibility – we follow him on this in the next section – but for now our focus is on the characterization of the potential defect and its corresponding virtue. It seems to us that the gloss 'unsettled/settled by the current meaning' is a problematic fit for the role of defective/virtuous as just described. First, assuming $\mathcal{L}_{MV}$ sentences enjoy this virtue, it isn't obvious that a translation would preserve it. To take a familiar example, whatever being 'settled by the current meaning' comes to, it seems there might well be a sentence of $\mathcal{L}_\in$ that translates to a synonymous sentence of $\mathcal{L}_\mathbb{N}$ (the language of arithmetic), where the former is settled by the current meaning of $\mathcal{L}_\in$ but the latter is unsettled by the current meaning of $\mathcal{L}_\mathbb{N}$ (e.g., a strong consistency statement). Second, more importantly, what reason is there to think that all $\mathcal{L}_{MV}$ sentences enjoy this virtue in the first place? Why should the meaning currently assigned to $\mathcal{L}_{MV}$ do any better at settling all sentences of $\mathcal{L}_{MV}$ than the meaning currently assigned to $\mathcal{L}_\in$ does at settling all sentences of $\mathcal{L}_\in$?[55] There may be answers to these challenges, but clearly much more would need to be said.

Fortunately, there's no need to get into these thickets; we offer an alternative characterization of defective/virtuous that has the added advantage of tying more directly into the line of thought we've been tracing.[56] To see this, recall we're assuming that our examination of the various candidates shows them all to be on equal footing and that our best response is to trim the syntax of $\mathcal{L}_\in$. On those assumptions, consider the state of two imaginary set theorists, a universe theorist and a multiverse theorist. The universe theorist speaks $\mathcal{L}_\in$, embraces $ZFC + LCs$, and persists in trying to figure out the 'correct' way to extend it;[57] under our current assumptions, this

[54]See Steel (2014), p. 167, and Proposition 41 in Appendix C.

[55]There are, of course, the familiar Gödel sentences unsettled by $\mathcal{L}_{MV}$, but – to get ahead of our story, well into Section 6 – is there any more reason to think that the meaning currently assigned to $\mathcal{L}_{MV}$ settles CH^{C} than to think that the meaning currently assigned to $\mathcal{L}_\in$ settles *CH*?

[56]Here we appear to depart from Steel, but see footnote 59.

[57]The universe theorist bears some resemblance to Steel's strong absolutist (Steel, 2014, p. 168): perhaps it's unobjectionable to say that she understands $\dot{V}$ – it's just her V – but initially she isn't privy to $\mathcal{L}_{MV}$, so she takes no stand on whether $\dot{V}$ is expressible there. She might be characterized as a thin realist (Maddy, 2011) who hasn't yet considered the possibility of switching from a theory of sets to a theory of sets and universes.

universe theorist is just wrong, making a mistake. In contrast, our multiverse theorist is aware that no candidate is preferable, speaks $\mathcal{L}_{MV}$, and embraces MV. This multiverse theorist thinks, with considerable justification on our assumptions, that the universe theorist is missing the fact that all the candidate foundational theories represented by worlds in the multiverse have equal standing.[58]

To put this another way, we might say that from the multiverse theorist's perspective, the universe theorist's $\mathcal{L}_{\in}$ sentences may reflect an improper bias, restricting attention to one world, while all $\mathcal{L}_{MV}$ sentences are suitably impartial. Steel expresses this idea with an analogy:[59]

> 'the laws of physics are the same in all inertial frames' has a parallel to 'the laws of set theory are the same in all universes of the generic multiverse'. Just as there is something called coordinate-free geometry, MV and its extensions might be called 'coordinate-free set theory'. (Private communication, 12/10/18, quoted with permission.)

Just as statements of coordinate geometry can be partial to one coordinate system, the sentences of $\mathcal{L}_{\in}$ can be partial to one world of the multiverse. Just as statements of coordinate-free geometry aren't partial to any particular coordinate system, the sentences of $\mathcal{L}_{MV}$ aren't partial to any particular world of the multiverse. This *impartiality*, we submit, is the virtue that $\mathcal{L}_{MV}$ sentences enjoy, that the translation function must preserve, and that CH might lack.

How this works obviously depends on the specifics. The mathematical key to this line of thought is a recursive function, $t : \mathcal{L}_{MV} \to \mathcal{L}_{\in}$, defined in ZFC. The hope is that the universe claim $t(\varphi)$ in some sense captures the spirit of the multiverse claim φ, in particular, that it preserves its multiverse virtue of impartiality. The definition of t (see Appendix C) rests on three basic facts: (1) since Steel's multiverse includes Amaglamation, for any worlds V_0 and V_1, there exists another world U that's a generic extension of them both, so any world in the multiverse is just a generic extension followed by a generic refinement away from any other; (2) the forcing relation is definable;[60] and (3) generic refinements are definable (Theorem 5). With these building blocks, we can show that

THEOREM 9. *There is a recursive function $t : \mathcal{L}_{MV} \to \mathcal{L}_{\in}$ such that for all sentences $\varphi \in \mathcal{L}_{MV}$, MV proves that the following are equivalent*:

(1) φ;
(2) $\forall U t(\varphi)^U$;
(3) $\exists U t(\varphi)^U$.

---

[58] Again, equal standing as candidates for our fundamental foundational theory. See footnote 51.

[59] Cf. footnote 56. Perhaps we aren't departing so much, after all.

[60] Here we just mean definable in the sense of the standard fact that the forcing relation for Σ_n formulae is definable for any $n \in \omega$.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

PROOF. See the proof in Appendix C. (As noted there, in Proposition 40, Amalgamation is essential for this result.) ⊣

So the multiverse theorist, speaker of $\mathcal{L}_{MV}$, supporter of the theory MV, sees that each world of his multiverse contains an encoding of each multiverse truth. From his multiverse perspective, a universe theorist, speaker of $\mathcal{L}_{\in}$, is confined to one of the multiverse's worlds; she's a supporter of the theory of that particular world. But, the multiverse theorist continues, despite this universist's parochial stance, she still has access to multiverse truth via the t function. (The multiverse theorist also sees that the universe theorist is making a serious mistake, the nature of which will come clear in just a moment.)

But this encoding of multiverse truth isn't quite enough. If the impartiality of $\varphi \in \mathcal{L}_{MV}$ is to be found in the universe language $\mathcal{L}_{\in}$, that impartiality has to be something the speaker of $\mathcal{L}_{\in}$, the universe theorist, can appreciate. But an $\mathcal{L}_{\in}$ speaker doesn't understand $\mathcal{L}_{MV}$ and has no reason to be moved by Theorem 9, a proof-theoretic fact about an entirely foreign theory, MV. $t(\varphi)$ is there in her language, but it might be said that she understands it in a sense analogous to that in which I 'see' a well-camouflaged pheasant hiding in the brush: photons from its feathers reach my eyes, I'm aware of its colors, but the bird itself I don't discern.

What brings the significance of $t(\varphi)$ home to the universe theorist is the following theorem, which is loosely speaking equivalent to Theorem 9:

THEOREM 10. *There is a recursive function* $t : \mathcal{L}_{MV} \to \mathcal{L}_{\in}$ *such that if* $\mathcal{M}$ *is a countable model of ZFC, G is* $Col(\omega, < Ord)^{\mathcal{M}}$*-generic over* $\mathcal{M}$, *and* $\varphi \in \mathcal{L}_{MV}$, *then*:

$$\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{M}^G \models \varphi.$$

PROOF. See Appendix C. ⊣

For any countable model $\mathcal{M}$, our universe theorist can clearly understand the structure $\mathcal{M}^G$. She can then understand $\mathcal{L}_{MV}$ as interpreted in $\mathcal{M}^G$ and MV as a theory that's true there. And, finally, Theorem 10 shows her the significance of $t(\varphi)$: in Steel's words, the $\mathcal{L}_{\in}$ sentence $t(\varphi)$, evaluated in $\mathcal{M}$, says in effect that the $\mathcal{L}_{MV}$ sentence 'φ is true in some (equivalently all) multiverse(s) obtained from me' (Steel, 2014, p. 166). Thus the pheasant emerges from its surroundings.

In fact, there's a perfectly ordinary sense in which t is 'meaning-preserving': φ makes a certain claim about the multiverse; $t(\varphi)$, evaluated in $\mathcal{M}$, makes the very same claim about $\mathcal{M}^G$. As an $\mathcal{L}_{MV}$ sentence, φ is impartial because it takes into account the full range of theories represented in the multiverse; the $\mathcal{L}_{\in}$ sentence $t(\varphi)$ inherits the virtue of impartiality via t, because it takes into account the full range of $\mathcal{M}^G$, the universe theorist's best understanding of the multiverse. So, here is (our version of) Steel's proposal: on the assumed outcome of our investigation – that is, on the assumption that there are no good reasons to prefer any of our candidate foundational theories to the others – the language $\mathcal{L}_{\in}$ is prompting us to pose pseudo-questions, so we should 'trim back our syntax', the syntax of $\mathcal{L}_\in$, to favor sentences in the range of t. (This where the multiverse theorist takes the universe theorist to be making a serious mistake: she thinks all sentences of $\mathcal{L}_\in$ are just as good as those that fall in the range of t.)

How is this to be done? We're out to specify which sentences of $\mathcal{L}_\in$ are legitimate, which don't tempt us to pseudo-questions. We've seen that what makes sentences of $\mathcal{L}_{MV}$ legitimate is their impartiality, and that t is meaning-preserving in the rough-and-ready sense sketched above, so certainly $\mathcal{L}_\in$ sentences in the range of t are legitimate. But it won't do to simply stop there. As is clear from the sketch above and the official definition in Appendix C, the outputs of t are quite complex constructions, not the sort of thing that turns up in ordinary set-theoretic practice; for example, even the sentence affirming the existence of the empty set – $\exists x \forall y\, (y \notin x))$ – isn't in the range of t. To procure its legitimacy, Steel would require that $\exists x \forall y\, (y \notin x)$ be *synonymous* with $t(\varphi)$ for some $\varphi \in \mathcal{L}_{MV}$, but $t(\varphi)$ involves a great deal more mathematical machinery than $\exists x \forall y\, (y \notin x)$ – and Steel acknowledges that this makes the synonymy claim problematic (Maddy, 2017, p. 314). Short of requiring synonymy in some recognizable sense, it seems to us enough that it's provably equivalent to a formula of the form $t(\varphi)$. So we say

DEFINITION 11. $\varphi \in \mathcal{L}_\in$ is *legitimate*$_T$, where T extends ZFC, if there is some $\psi \in \mathcal{L}_{MV}$ such that

$$T \vdash \varphi \leftrightarrow t(\psi).$$

Here T can extend ZFC, as in stronger versions of MV. Without the subscript, we use 'legitimate' more loosely, for statements of $\mathcal{L}_\in$ that preserve the impartiality of $\mathcal{L}_{MV}$. So being legitimate$_T$ for some reasonable T is a good indicator of being legitimate.

On this definition, not only $\forall y\, (y \notin x)$, but any $\varphi \in \mathcal{L}_{MV}$ that provably holds (or fails) in every world of the multiverse is also legitimate:

THEOREM 12. *Let $\varphi \in \mathcal{L}_\in$. Then following are equivalent*:

(1) $MV \vdash (\forall V \varphi^V \vee \forall V \neg\varphi^V)$.
(2) *φ is legitimate$_{ZFC}$*.

It follows immediately that all theorems of ZFC are legitimate. In fact, even some undecidable sentences like $\neg Con(ZFC)$ are legitimate; because such simple claims are unaffected by forcing, they have the same truth value in all worlds of any particular $\mathcal{M}^G$. Steel's example noted above – the $\mathcal{L}_\in$ sentence 'there is a proper class of measurable cardinals' – also comes out legitimate because it's equivalent in ZFC to $t(\forall V(\text{there is a measurable cardinal})^V)$ (see Appendix C4). Indeed, at this point, with ZFC as our background theory, the legitimate statements of $L_\in$ are exactly those that are determinate in Woodin's multiverse, that is, true in all worlds or false in all worlds

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

of $\mathbb{V}_{\mathcal{M}}$.[61] Steel's focus in the next section is on the possibility of undermining this extensional equivalence, of extending legitimacy to some statements that are indeterminate in Woodin's multiverse.

In sum, then, though this reconstruction doesn't match the letter of Steel's presentation – substituting 'impartiality' for 'settled by the current meaning' as the relevant virtue and 'provable equivalence' for 'synonymy' as the impartiality-preserving relation within $\mathcal{L}_\in$ – it seems to us roughly in the spirit of his project. Notably, it doesn't require a notion of 'meaning' any more sophisticated than the explicit and transparent relation described above between φ and $t(\varphi)$.[62] In contrast to the metaphysical theorizing of Hamkins and Woodin, its treatments of the multiverse language, the multiverse theory, and the relations between these and our current universe language and theory are all purely linguistic and proof-theoretic. Those of us who prize the independence of mathematical practice from external questions of meaning, truth, and existence will welcome these outcomes.

§6. The status of *CH*. We turn at last to the continuum problem. Concerned that our only guide to formulating foundational theories is 'maximize interpretive power', recognizing that it gives out at $ZFC + LCs$ and that this theory is too weak to settle CH, we embraced a multiverse of candidate foundational theories. Assuming, as we have been, that no new guide has emerged to give us reason to prefer some of these candidates over others, we propose to trim the syntax of $\mathcal{L}_\in$ to legitimate claims, as described in the previous section. The problem of CH can now be formulated with some precision: is CH legitimate, that is, is there a φ in $\mathcal{L}_{MV}$ such that the $\mathcal{L}_\in$ sentence $t(\varphi)$ is provably equivalent to CH in some reasonable theory?

Now obviously the legitimacy of CH can't be established by the route suggested in Theorem 12 because CH is true in some worlds and false in others. But (as noted in passing above) Steel doesn't see this as the end of the story:

> The multiverse language is ... sufficiently expressive to state versions of the axioms of ZFC, and of the large cardinal hypotheses preserved by set forcing: we replace φ by 'for all worlds W, φ^W'. Clearly we cannot state CH this way. The same goes for the many other statements about the uncountable which are sensitive to set forcing, no matter what large cardinals there may be. Whether there are traces of CH and these other sentences in the multiverse language is the issue we consider next. (Steel, 2014, p. 167)

[61] A little care is required here. Woodin's original definition of determinateness focused on countable transitive models. So a sentence is determinate if it is true in $\mathbb{V}_M$ for some ctm M of ZFC. The sentences which are determinate in this sense are not the legitimate$_{ZFC}$ sentences. However, this distinction is merely an artifact of the move between ctms and the arbitrary countable models used in the definition of legitimate. So if we relax the transitivity restriction and consider $\mathbb{V}_{\mathcal{M}}$ for arbitrary countable $\mathcal{M}$ then they are the same.

[62] We regard this as an improvement. Recall footnote 50.

How might this happen? Well, perhaps there's 'a distinguished reference world ...an individual world that is definable in the multiverse language' (Steel, 2014, p. 168). Steel credits Woodin with the observation that such a world would be unique and contained in all other worlds. This special world, if it exists, is called the 'core' of the multiverse:

Definition 13. A world C is a *core* of the multiverse iff $\forall x(x \in \mathsf{C} \rightarrow \forall U(x \in U))$.

So defined, C is included in every world, thus in their intersection, and because it's one of the worlds being intersected, it's equal to the intersection. Of course the intersection itself will be definable in any case, but it's only the core if it's a world:

There is a core iff $\exists U \forall x(x \in U \rightarrow \forall V(x \in V))$.

For $\varphi \in \mathcal{L}_{MV}$, let φ^{C} be φ relativized to $\forall U(x \in U)$.[63] Supposing, then, there is such a core, Steel considers CH^{C} in $\mathcal{L}_{MV}$[64] and suggests that in $\mathcal{L}_{\in}$, $t(CH^{\mathsf{C}})$ might be 'synonymous' with CH itself. In that way, CH^{C} would be the 'trace of CH' in the multiverse language.

Before examining this move, we should note that since the appearance of Steel (2014), Usuba (2017) has proved that the multiverse *does* have a core, assuming the existence of an extendible cardinal (EXT):[65]

Theorem 14. $MV + \forall U(EXT^{U}) \vdash \exists U(U \textit{ is the core})$.[66]

Returning to the question of CH, Steel grants that any ordinary kind of synonymy is a stretch here: $t(CH^{\mathsf{C}})$ involves mathematical machinery unknown to Cantor, but surely we don't want to suggest that Cantor didn't understand the meaning of CH![67] Though the switch to 'legitimate $_T$' (Definition 11) replaced the problematic 'synonymous' in this context with provable equivalence, this is of no immediate help here, because

$$ZFC \nvdash (CH \leftrightarrow t(CH^{\mathsf{C}})).^{68}$$

That is, CH isn't legitimate $_{ZFC}$. But this isn't really the question we should be asking, since the existence of the core depends on an extendible cardinal,

[63] Observe that relativizing to the core in this way does not imply that there is a core.

[64] It might seem that CH^{C} violates impartiality even in the multiverse language because it only involves what happens in the single world C, but recall that C itself is defined in terms of all worlds.

[65] A cardinal κ is η-extendible iff there is an elementary embedding $j : V_{\kappa+\eta} \rightarrow V_{\theta}$ with critical point κ for some ordinal θ. κ is extendible if it is η-extendible for every ordinal η. In Steel's terminology (Steel, 2014, pp. 167–168), Usuba's result suggests that a weak relativist is also a weak absolutist.

[66] In fact, we only need to add that $\exists U(EXT^{U})$ for this to work. We've used the stronger assumption because it's Steel's standard way of adding large cardinals to MV (see the parenthetical comment after MV-1 $_{\varphi}$ in Section 4).

[67] See footnote 58 of Maddy (2017), p. 314.

[68] See Proposition 44(1) in Appendix C.

well beyond the reach of ZFC on its own. For that matter, how does the universe theorist, speaker of $\mathcal{L}_\in$, understand talk of the core in the first place?

To answer this question, recall Theorem 10 and the surrounding discussion. The universe theorist understands an $\mathcal{L}_{MV}$ claim φ by interpreting it in $\mathcal{M}^G$, and the theorem tells her that the $\mathcal{L}_\in$ claim $t(\varphi)$, interpreted in $\mathcal{M}$, encodes the claim that ' φ holds in the multiverse generated from me'. So consider the case of the $\mathcal{L}_{MV}$ claim 'there is a core' as above. t('there is a core'), interpreted in $\mathcal{M}$, says that M^G thinks there is a core. So the universe theorist understands t('there is a core') itself to say that the multiverse generated from her single universe – if there were such a thing – would have a core.

This line of thought can be expressed directly in $\mathcal{L}_\in$ using $W_{(\cdot)}$ from Theorem 5.

There is a core iff $\exists r \forall x(x \in W_r \leftrightarrow \forall s(x \in W_s))$.[69]

This says that the core is a generic refinement of every generic refinement. Usuba's theorem can be reformulated to

Theorem 15. *$ZFC + EXT \vdash$ 'there is a core'.*

In this sense, then, if we allow our universe theorist to augment ZFC to $ZFC + EXT$, just as our multiverse theorist has augmented MV to $MV + \forall U(EXT^U)$, she can see that the multiverse generated from her single universe – if there were such a thing – would have a core. With this machinery in hand, our question can be improved to

$$ZFC + EXT \vdash (CH \leftrightarrow t(CH^{\mathsf{C}}))?$$

But the answer is still no; CH isn't legitimate$_{ZFC+EXT}$, either.[70] Still more is needed if CH is to be legitimized.

Addressing this challenge, Steel remarks that

> One can think of [$\mathcal{L}_\in$] as the multiverse language, together with a constant symbol $\dot{\mathsf{V}}$ for a reference universe. (Steel, 2014, p. 167)

The idea is that the multiverse theorist might understand the universe theorist simply as speaking of that reference universe; in other words, what the universe theorist takes for the single universe, V,[71] is really the

[69]Observe that this sentence merely says that W_r has no generic refinements, not that it is the intersection of all the worlds. So W_r is a world that satisfies what Rietz calls the Ground Axiom (i.e., W_r is not a generic extension of any world), but we may wonder if there are other worlds also satisfying it (Reitz, 2007). This is ruled out by Usuba's Downward Directed Ground theorem (Usuba, 2017), which tells us the generic refinement relation is downward directed; i.e., given W_r and W_s, generic refinements of the universe, there is some W_t which is a generic refinement of both W_r and W_s.

[70]See Proposition 44(2) in Appendix C.

[71]We use V for the universe theorist's unique world, in contrast with the variable V over the multiverse theorist's many worlds.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

interpretation of V in the extended multiverse language. But this only works if H is given a definition in $\mathcal{L}_{MV}$. Steel continues:

> If the multiverse has a core, then surely it is important, whether it is the denotation of $\dot{\text{V}}$ or not! Indeed, if there is an inclusion-least world in the multiverse, why not use $\dot{\text{H}}$ to denote it, and agree to retire $\dot{\text{V}}$ until we need it? (Steel, 2014, p. 169)

This is more than a little cryptic, but the suggestion seems to be that the multiverse theorist understand the universe theorist as talking about the core or even that the universe theorist take the core to be her single universe (at least until the need arises for something else). The universe theorist is being asked to agree that there are no proper generic refinements of her single universe, that is, to agree that $\forall x \forall r (x \in W_r)$ or more succinctly, $\mathsf{V} = \mathsf{C}$.[72] This does the trick:

$$ZFC + EXT + \mathsf{V} = \mathsf{C} \vdash (CH \leftrightarrow t(CH^{\mathsf{C}})).$$

(This follows from Theorem 17.) In fact, every statement of $\mathcal{L}_{\in}$ is legitimate$_{ZFC+EXT+\mathsf{V}=\mathsf{C}}$, potentially removing any worry about pseudo-questions! But why should the universe theorist regard $ZFC + EXT + \mathsf{V} = \mathsf{C}$ as a reasonable theory? In particular, why should she identify her universe with C?

Assuming, as we have been, that the available guides offer us no reason to prefer some candidate foundational theories with large cardinals over the others, there can be no reason to opt for $\mathsf{V} = \mathsf{C}$ over all the rest, but there's another way to look at the role of the core in Steel's thinking here. We've seen how, at the beginning of the paper, he imagines two possible reactions 'for those who …believe that the truth value of CH is not determined by the meaning we currently assign to the syntax of' $\mathcal{L}_{\in}$, namely, 'trim back the current syntax' or 'flesh out the current meaning' (Steel, 2014, p. 154). We've replaced 'believing the truth value of CH isn't determined by the current meaning' with the less-loaded 'taking "maximizing interpretive power" to be the only guide we have to formulating foundational theories and $ZFC + LCs$ to be all we can glean from this guide'; and we've been exploring how to 'trim' under this assumption – which allows no reason to opt for $\mathsf{V} = \mathsf{C}$. The 'fleshing out' alternative comes up in a moment, but for now, it's important to note that there's actually an intermediate position in play here.

In Steel's meaning-theoretic idiom, that intermediate position is the possibility that there's more to the current meaning than we've so far appreciated, that $ZFC + LCs$ isn't all that's implicit there. In our less-loaded idiom, there might be good reasons to favor some candidate foundational theories represented in the multiverse over others. Though the complexities of 'trimming' make it easy to lose sight of this fact, the original motivation of the multiverse was 'to find a neutral common ground on which to compare'

[72]This is known as the Ground Axiom in Reitz (2007), however we are using it to assert that the universe is a solid bedrock in Reitz's terminology. These can be seen to be equivalent by the Downward Directed Grounds theorem (Usuba, 2017).

the various candidate theories 'without bias toward any' (Steel, 2014, p. 165). Steel's thought was that attending to the multiverse, placing all the candidate theories side-by-side, might reveal hidden aspects of meaning, or, in our terms, uncover previously unnoticed reasons for preferring one to another candidate. He now points out that attending to the multiverse has, in fact, revealed a new 'fundamental question' (ibid., p. 169): does the multiverse have a core? The suggestion above, that the universe theorist is talking about the core, is really the suggestion that thinking about the core in the context of the multiverse gives us reason to prefer its theory to the other candidates for our fundamental, foundational theory.

So far so good, but why should this be so, what makes the theory of C preferable? Think of it this way. We've just recalled that Steel introduced the multiverse as a way of drawing all the candidate theories together, to be viewed side-by-side, on equal footing ('without bias'). What's emerged from this exercise, from the pure mathematics of the situation, is that the candidate theories aren't actually on a par – one is singled out as more fundamental. Furthermore, from the perspective of that special theory, each of the other candidates is only a forcing extension away. This means that a universe theorist working in the core can understand everything about alternative theories simply by studying what's forced by particular posets;[73] in this sense, the entire multiverse of theories is accessible to her. The question then arises: why not regard the core as the interpretation of $\dot{\mathrm{V}}$? What would be lost by preferring the theory $ZFC + EXT + \mathsf{V} = \mathsf{C}$?

This is undeniably a provocative line of thought. Perhaps a compelling case could be made for answering these questions affirmatively, though we won't attempt to fill one in here. But, if this could be done, it might be argued that multiverse thinking would, in the end, reveal something important about $\mathcal{L}_{\in}$, namely that all its sentences are legitimate, after all. Steel describes the situation this way:

> Perhaps ... some future mathematics [will be] built around an understanding of the symbol $\dot{\mathrm{V}}$ that does not involve defining $\dot{\mathrm{V}}$ in the multiverse language [i.e., as the core]. But at the moment, it is hard to see what that is. (Steel, 2014, p. 168)

Short of such a future, this would be a remarkable outcome: by openly facing the possibility that some sentences of $\mathcal{L}_{\in}$ are defective, by giving multiverse thinking its due, Steel would end up dispelling the very concern that he started with.

As it happens, though, even this dramatic conclusion wouldn't have the mathematical power it might appear to promise: it turns out that $ZFC + EXT + \mathsf{V} = \mathsf{C}$ is a much less informative theory than one might have hoped. In the wake of Usuba's theorem, Steel[74] alludes to work of the set-theoretic

[73] We might think in terms of Kunen's 'forcing over V' in Kunen (2006), pp. 234–235, or 'forcing over the Universe' in Kunen (2011), pp. 281–282.

[74] Private correspondence, 2/19/17, quoted with permission.

geologists, who have shown, for example, that $\mathsf{V} = \mathsf{C}$ can't settle CH. But this failure is just the beginning:

THEOREM 16 (Reitz, 2007). *If V satisfies ZFC, then there is a class forcing extension $V[G]$ such that $V[G] \models \mathsf{V} = \mathsf{C}$.*

In other words, a model of $\mathsf{V} = \mathsf{C}$ can be generated from any model of ZFC. Moreover, the generated model can be constructed in such a way as to preserve any rank initial segment of the original universe. Or to put it the other way 'round, a model of $ZFC + \mathsf{V} = \mathsf{C}$ might have as an initial segment, an initial segment of any universe whatsoever. This suggests that assuming $\mathsf{V} = \mathsf{C}$ pins down very little about what V is actually like.

At this point, it seems unlikely that any more can be gleaned from the 'current meaning' of $\mathcal{L}_\in$, so Steel turns at last to the possibility that we might 'flesh out' that meaning in some principled way. Notice, by the way, that in our less-loaded idiom, the contrast between 'revealing hidden aspects of the meaning' and 'fleshing out the meaning' is a distinction without a difference: both come down to finding sound mathematical reasons to prefer one theory in the multiverse to others. In any case, Steel sees the line between 'revealing hidden aspects' and 'fleshing out' as having been crossed when we try to say more about the core.

We've just seen that the key role of the core here is to legitimize all sentences of $\mathcal{L}_\in$:

THEOREM 17. *For each $\varphi \in \mathcal{L}_\in$, there is a $\psi \in \mathcal{L}_{MV}$ such that*

$$ZFC + \mathsf{V} = \mathsf{C} \vdash \varphi \leftrightarrow t(\psi).$$

PROOF. See Appendix C.1. ⊣

As it happens, a number of familiar inner models could substitute for C in this theorem – L, $L[0^\#]$, the core model K – because each of these satisfies $\mathsf{V} = \mathsf{C}$. But, as we've seen (in Section 4), these don't appear as worlds in Steel's multiverse due to their antilarge-cardinal effects. There is, however, a developing program around an axiom candidate designed in the hope of overcoming this problem, namely, V=Ultimate-L.[75] This axiom implies $\mathsf{V} = \mathsf{C}$,[76] and if all goes as advertised, it would also be consistent with all traditional large cardinals; in a sense to be made precise, V=Ultimate-L would imply that the universe is as L-like as possible without antilarge-cardinal effects.

To this point, then, Steel has argued that $\mathsf{V} = \mathsf{C}$ is implicit in the meaning of $\mathcal{L}_\in$. But we've also seen that $ZFC + LCs + \mathsf{V} = \mathsf{C}$ is too weak a theory to settle CH, or much else. In stark contrast, V=Ultimate-L would be a very powerful assumption indeed – beginning, though by no means ending, with

[75] See Woodin (2017).

[76] Woodin (2017), Theorem 7.8.

a solution to the Continuum Problem:

$$ZFC + \mathsf{V} = \text{Ultimate-L} \vdash CH.$$

So here, at last, is Steel's second option: assuming the Ultimate-L project succeeds, rather than 'trimming back the syntax' of $\mathcal{L}_{\in}$ to avoid asking pseudo-questions, he proposes that we opt instead to 'flesh out', to extend its meaning to include V=Ultimate-L, that we move from $ZFC + LCs + \mathsf{V} = \mathsf{C}$ to $ZFC + LCs + \mathsf{V}$=Ultimate L. Indeed something like this may have been what Steel had in mind in 'Gödel's program', before Usuba's theorem: 'the multiverse may indeed have a core, and this core may admit a detailed fine-structural analysis that resembles Gödel's L' (Steel, 2014, p. 178).

Steel presents no detailed case for this idea; presumably his reasons would dovetail with those offered by other advocates of V=Ultimate-L. But for now, this remains speculative, as Steel clearly acknowledges:

> Perhaps the mathematics will turn out some other way. Perhaps the multiverse has no core [written pre-Usuba], but some other, more subtle structure. There are many basic open questions at the foundations of set theory: the extent of generic absoluteness, the existence of iterable structures, the Ω-Conjecture, the form of canonical inner models with supercompacts, and the properties of *HOD* in models of determinacy, to give my own partial list. Our path toward a stronger foundation will be lit by the answers to such questions. (Steel, 2014, p. 179)

§7. Caveats and conclusions. This completes our reconstruction of Steel's multiverse project. We've done our level best to tell the story as fully and persuasively as we can, but certainly don't pretend that the line of thought sketched here is air-tight at every turn. We conclude with a brief look at some lingering concerns.

The most troubling questions center on the assumption of Amalgamation. Recall that this axiom is essential to the turn of argument in Section 4: the failure of Amalgamation is what makes (the meta-mathematical surrogate for) Woodin's multiverse unaxiomatizable; with Amalgamation, Steel's axioms successfully axiomatize his natural toy models. One awkward question then is why we should insist on axiomatization in this strong sense.[77] Steel is after a theory of sets and worlds that represents the full range of candidate foundational theories. It's obvious that he needs axioms, but why wouldn't it be sufficient to isolate a set of axioms that captures this central idea well enough to generate a mathematically successful theory, even if it wasn't complete for some natural collection of toy models? Without a satisfactory answer to this question, we have no reason to adopt the

[77] This requirement isn't explicit in Steel (2014), but it seems implicit in his remark about Woodin's multiverse: 'it is not at all clear what its theory would be' (Steel, 2014, p. 170). (If 'axiomatization' in a looser sense were intended, why not MV-Amalgamation?) In any case, appeal to full axiomatizabilty was the only way we could find to mount a principled case for Amalgamation.

axiomatizabilty requirement, and we're left without a principled argument for Amalgamation.

Unfortunately, Amalgamation can't simply be jettisoned. Its involvement in the mathematics of Steel's theorizing goes far deeper than the question of axiomatizability: it's essential to the translation function of Theorems 9 and 10 in Section 5, which could hardly be more fundamental. The history of set theory has accustomed us to the notion that an axiom can be defended by appeal to its mathematical benefits – these are so-called extrinsic justifications, going back to Zermelo, endorsed by Gödel, now playing a central role in contemporary set-theoretic practice – so one thought would be to defend the addition of Amalgamation to MV by pointing to its welcome mathematical consequences.

The trouble with this approach is that MV isn't a pure mathematical theory; it's a meta-mathematical theory – a piece of applied mathematics intended to provide representations for all candidate foundational theories – and for applied mathematical theories, mathematical benefits aren't enough, we need representational accuracy.[78] Not perfect accuracy, of course, but whenever simplifications or idealizations are employed, we have to have good reason to believe that no relevant distortions are being introduced.[79] That's what we don't have in this case. Presumably the candidate foundational theories are what they are independently of our theorizing about them. When Steel narrows his multiverse theory by adding Amalgamation, what reason do we have for thinking that he hasn't ruled out some perfectly good candidates?

Finally, at least in passing, we should flag the assumption that 'maximizing interpretive power' is the only guide we have to extending our theory of sets. Though Steel has good reasons for this position, it bears noting that there are differing opinions in the field, particularly among supporters of forcing axioms.[80] But 'noting' is enough for our purposes – this obviously isn't the place to engage that debate!

So, in sum, what have we learned from Steel's discussion of his multiverse? The concluding argument for V=Ultimate-L more-or-less coincides with what its current defenders offer as a straightforward case for adding it directly as a new axiom to some extension of ZFC. What distinguishes Steel's project is the path he travels to get there: after taking seriously the possibility that there are no grounds on which to extend past $ZFC + LCs$, he explores his multiverse language and theory; the mathematics itself leads him to ask whether the multiverse has a core, and post-Usuba, we find that it does; Steel then suggests that $\mathsf{V} = \mathsf{C}$, which returns us to universe thinking about the theory of that structure and the debate over V=Ultimate-

[78]For comparison, there would be mathematical benefits to describing fluid flow with something more tractable than the Navier-Stokes equations, but as it happens, the world just isn't that cooperative.

[79]Sticking to fluid dynamics for comparison, applied mathematicians who assume that fluids are continuous present detailed analyses of when and why this idealization is benign.

[80]See, e.g., Magidor (2019), Todorcevic (2019). For a broader discussion, see Schatz (2019).

L. Metaphysically, we've seen (in Section 2) that much of the current multiverse discussion involves a robust ontology of sets and universes. Here Steel presents a stark alternative: a linguistically defined multiverse intended to capture the full range of candidates for extending $ZFC + LCs$. His goal isn't metaphysical truths about an abstract realm, but a fair adjudication of the whether CH is a viable set-theoretic question. Though the discussion is couched in terms of 'the current meaning assigned to $\mathcal{L}_{\in}$', we've argued that this is inessential, that the substance of Steel's thought can be formulated more effectively in philosophically innocent mathematical terms. By these means, we steer away from the vagaries of mathematical meaning, truth, and existence and toward the methodologically central questions: how exactly do we select our theories and by what right? At that point, Steel's approach offers something rare: a novel and distinctive approach to answering them.[81]

Appendix. The goal of this appendix is to fill in the mathematical aspects of the story we've told above. In parts A and B, we prove two claims from Section 4: that MV is sound and complete for the natural class of models and that no such theory is available for Woodin's multiverse. In part C, we define the translation function t from $\mathcal{L}_{MV}$ to $\mathcal{L}_{\in}$, and prove the key results (Theorem 9, Theorem 10, and Theorem 12 from Section 5, and Theorem 17 from Section 6). In some cases, we provide proofs for explicit claims of Steel (2014); in others, both claim and proof are part of our reconstruction. We don't believe any of this material has previously appeared in print.

Because our target audience includes mathematically-informed philosophers as well as set theorists, we have included more details and examples than usual. Anyone acquainted with forcing, large cardinals, and ultrapowers at the level of Kunen (2011) and the first seventeen chapters of Jech (2003) should be well-served. Citations to ancillary resources have also been included where they might be helpful. Our hope is that this explicit treatment will serve as a foundation for future philosophical work on the generic multiverse and multiverse theories more generally.

In overview, our proof in part A of the soundness and completeness of MV begins with a sketch of a general framework for a particular class forcing over arbitrary countable models of ZFC. (We're grateful to Gabriel Goldberg for suggesting this framework, which allowed us to shorten the appendices significantly.) In part B, we prove that Woodin's generic multiverse is not amenable to a similar soundness and completeness theorem: using a

[81] Many thanks to John Steel for his rich and challenging paper, for his patience with our many questions and misunderstandings, and for his invaluable comments on many earlier drafts. Though we haven't reflected his thought faithfully at every turn – our departures have been noted along the way – perhaps we've at least provided a starting point for further investigations. Thanks also to Peter Koellner, Jeffrey Schatz, an anonymous referee, and especially to Gabriel Goldberg and Hugh Woodin (grateful acknowledgments of their specific contributions appear elsewhere in the text and footnotes) for helpful comments on earlier drafts. We'd also like to thank our *BSL* editor, Patricia Blanchette, for her patient and judicious handling of the manuscript.

result Woodin kindly allowed us to include, we show that Woodin's generic multiverse – when generalised to arbitrary countable models – has a theory from which the full theory of analysis can be computed. Thus, it has no recursive axiomatisation. Finally, in part C, we give a full definition of Steel's translation function t and prove that it works as claimed in Steel (2014). The appendix ends with the proof that adding $\mathsf{V} = \mathsf{C}$ to ZFC is enough to remove all threat of illegitimacy from sentences of $\mathcal{L}_\in$.

A. Soundness and completeness for MV.

A.1. Preliminaries. In this section we establish that MV is sound and complete with respect to a natural class of models (see Theorem 8). Before we launch into the main proofs, we first provide some background on a couple of slightly exotic elements that we need: class forcing to collapse all cardinals; and forcing over ill-founded models. First we discuss some class forcing basics.[82] Our goal is to show how to collapse all of the cardinals in a model while retaining some ability to refer to the ground model. We make use of this in the proof of Soundness and in defining the translation function.

We work for the moment with transitive models. Let $\mathbb{P} = Col(\omega, <Ord)$. Loosely following Kunen (2011), we work in a forcing language $\mathcal{FL}_\mathbb{P}(\check{V}) = \{\check{V}, \in, \sigma\}_{\sigma \in V^\mathbb{P}}$ which expands $\mathcal{L}_\in$ with constant symbols from $V^\mathbb{P}$ and an extra 1-place relation symbol $\check{V} = \{\langle \check{x}, 1\rangle \mid x \in V\}$. It should be clear that $\check{V}$ denotes V in any generic extension; i.e., $Val(\check{V}, G) = V$ whenever G is $\mathbb{P}$-generic over V. To deal with $\check{V}$ in the definition of the forcing relation, we add an extra clause. For a $\mathbb{P}$-name σ and $p \in \mathbb{P}$, we let

$$p \Vdash \sigma \in \check{V} \text{ iff } \{q \in \mathbb{P} \mid \exists x \in V\ q \Vdash \sigma = \check{x}\} \text{ is dense below } p.$$

$\check{V}$ gives us sufficient ability to refer to the ground model in the context of the forcing relation.

If we force using $\mathbb{P}$, then we cannot preserve all of ZFC. In particular, it is easy to see that powerset must fail. However, the rest of the axioms survive. Indeed, we can even preserve uses of replacement which make use of $\check{V}$. Let ZFC^- be ZFC without the powerset axiom. For $\dot{N}$ a one-place relation symbol, let $\mathcal{L}_\in(\dot{N})$ be the expansion of $\mathcal{L}_\in$ with $\dot{N}$. Let $ZFC^-(\dot{N})$ be ZFC^- with $\dot{N}$ allowed in the Replacement and Separation schemata.

Fact 18. *Let M be a transitive model of ZFC. Let $\mathbb{P} = Col(\omega, <Ord)$. Then*

(1) *If G is $\mathbb{P}$-generic over M, then*

$$\langle M[G], M, \in\rangle \models \varphi((\tau_0)_G, \dots, (\tau_n)_G) \Leftrightarrow \exists p \in G\ M \models \text{“}p \Vdash \varphi(\tau_0, \dots, \tau_n)\text{”},$$

where $\varphi(v_0, \dots, v_n)$ is a formula of $\mathcal{FL}_\mathbb{P}(\check{V})$.

(2) $\Vdash_\mathbb{P} ZFC^-(\check{V})$.

[82] For a detailed discussion see Friedman (2000).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

This can be established by showing that $\mathbb{P}$ is pretame. The proof is identical to the proof of Lemma 2.26 in Friedman (2000). Alternatively, it can be shown quite directly by using a couple of facts about $\mathbb{P}$. First, the homogeneity of $\mathbb{P}$ and its factors can be used to show that whenever $p \Vdash_{\mathbb{P}} \varphi(\sigma_0, \dots, \sigma_n)$, then $p \upharpoonright \alpha \Vdash_{\mathbb{P}\upharpoonright\alpha} \varphi(\sigma_0, \dots, \sigma_n)$ where α is the supremum of those β such that $q(\beta) \neq \top$ for some $q \in \mathbb{P}$ in the transitive closure of some σ_i for some $i \leq n$. This is enough to ensure that definability holds. Second, since $\mathbb{P}$ is essentially an *Ord*-length finite support product of *Ord*-*cc* forcings, every antichain in $\mathbb{P}$ is set-sized. This – in conjunction with the previous fact – allows us to obtain a version of the mixing lemma which then allows us to prove a simple modification of the standard proof that Replacement is forced.[83]

Now we consider forcing over possibly ill-founded models. We'll just describe some standard results.[84] Let $\mathcal{M}$ be a countable model of *ZFC*. Let $\mathbb{P}$ be a poset which is such that either: $\mathbb{P} = (Col(\omega, <))^{\mathcal{M}}$; or $\mathbb{P} \in \mathcal{M}$. Let $\mathcal{M}^{\mathbb{P}}$ be the $\mathbb{P}$-names as defined within $\mathcal{M}$. We say $G \subseteq \mathbb{P}$ is $\mathcal{M}$-generic if G is a filter and for all dense $(D \subseteq \mathbb{P})^{\mathcal{M}}$ with $D \in \mathcal{M}$, we have $G \cap D \neq \emptyset$.[85] We define $\mathcal{M}^{\mathbb{P}}/G$ as follows. Let the domain of $\mathcal{M}^{\mathbb{P}}/G$ be the set of $[\sigma]^{\mathcal{M}}_G$ where $\sigma \in \mathcal{M}^{\mathbb{P}}$ and $[\sigma]^{\mathcal{M}}_G = \{\tau \mid (\tau \in V^{\mathbb{P}})^{\mathcal{M}} \wedge \tau \sim_G \sigma\}$ where $\tau \sim_G \sigma$ iff $\exists p \in G (p \Vdash \tau = \sigma)^{\mathcal{M}}$. Let the membership relation of $\mathcal{M}^{\mathbb{P}}/G$, denoted $\in^{\mathcal{M}}_G$, be such that for $\sigma, \tau \in \mathcal{M}^{\mathbb{P}}$ we have

$$[\sigma]^{\mathcal{M}}_G \in^{\mathcal{M}}_G [\tau]^{\mathcal{M}}_G \Leftrightarrow \exists p \in G\, (p \Vdash \sigma \in \tau)^{\mathcal{M}}.$$

And for $\check{V}$ we let

$$[\sigma]^{\mathcal{M}}_G \in^{\mathcal{M}}_G [\check{V}]^{\mathcal{M}}_G \;\Leftrightarrow\; \exists p \in G\, (p \Vdash \sigma \in \check{V})^{\mathcal{M}}.$$

Note that versions of the truth and definability lemmas still hold.

FACT 19. *For* $\varphi(v_{0,n}) \in \mathcal{FL}_{\mathbb{P}}(\check{V})$ *and* $\tau_0, \dots, \tau_1 \in \mathcal{M}^{\mathbb{P}}$ *we have*

$$\langle \mathcal{M}^{\mathbb{P}}/G, \mathcal{M}, \in^{\mathcal{M}}_G \rangle \models \varphi([\tau_0]^{\mathcal{M}}_G, \dots, [\tau_n]^{\mathcal{M}}_G) \;\Leftrightarrow\; \exists p \in G\, (p \Vdash \varphi(\tau_0, \dots, \tau_n))^{\mathcal{M}}.$$

Moreover, a representation $[\dot{G}]^{\mathcal{M}}_G$ of G exists in $\mathcal{M}^{\mathbb{P}}/G$.

FACT 20. $G = \{p \mid (p \in \mathbb{P})^{\mathcal{M}} \wedge ([\check{p}]^{\mathcal{M}}_G \in [\dot{G}]^{\mathcal{M}}_G)^{\mathcal{M}^{\mathbb{P}}/G}\}$ *where* $\dot{G} = \{\langle \check{p}, p \rangle \mid p \in \mathbb{P}\}^{\mathcal{M}}$.

If $\mathcal{M}^{\mathbb{P}}/G$ is well-founded, we'll follow standard conventions and assume that it has been collapsed into a transitive model. But in general this will not be the case. If $\mathcal{M}^{\mathbb{P}}/G$ is ill-founded it is essentially an uncollapsed ultrapower

[83] For example, see page 254 of Kunen (2011).

[84] For a detailed account see Corazza (2007).

[85] We follow Kunen's notation of relativisation here (see page 141 in Kunen, 2006). Thus for $\mathcal{M} = \langle M, E \rangle$ and $\varphi \in \mathcal{L}_{\in}$ we write $\varphi^{\mathcal{M}}$ to indicate the result of replacing $\in$ by E and relativising all quantifiers to M. Note also that $G \cap D$ doesn't capture our strict intention, since G is a set in V (the ambient universe) while $D \in \mathcal{M}$ which might not have $\in$ as its membership relation. Strictly, we should say for all $D \in \mathcal{M}$ with (D is dense in $\mathbb{P})^{\mathcal{M}}$ there is some $g \in G$ with $g\,E\,D$ (i.e., $(g \in D)^{\mathcal{M}}$. For convenience, we'll adopt the sloppy notation since it should cause no confusion.

and so it is not strictly the case that $\mathcal{M}$ is a generic refinement of $\mathcal{M}^{\mathbb{P}}/G$. Indeed $\mathcal{M}$ is not literally a submodel of $\mathcal{M}^{\mathbb{P}}/G$. However, we do have the following:

FACT 21. *The embedding* $i_G^{\mathcal{M}} : \mathcal{M} \to \mathcal{M}^{\mathbb{P}}/G$ *where for all* $x \in \mathcal{M}$

$$i_G^{\mathcal{M}}(x) = [\check{x}]_G^{\mathcal{M}}$$

is such that $\mathcal{M}^{\mathbb{P}}/G$ *thinks that it is the generic extension of* $i_G^{\mathcal{M}}``\mathcal{M}$ *by* $i_G^{\mathcal{M}}``G$ *over* $i_G^{\mathcal{M}}(\mathbb{P})$.

With this in mind, it is convenient to modify $\mathcal{M}^{\mathbb{P}}/G$ in such a way that it does contain $\mathcal{M}$ as a submodel.

DEFINITION 22. Let $\mathcal{M}_{ult}[G]$ be the result of replacing $i_G^{\mathcal{M}}``\mathcal{M}$ by $\mathcal{M}$; i.e., let the domain of $\mathcal{M}_{ult}[G]$ be

$$(\mathcal{M}^{\mathbb{P}}/G \setminus i_G^{\mathcal{M}}``\mathcal{M}) \cup \mathcal{M}$$

and for $x, y \in \mathcal{M}_{ult}[G]$, let the membership relation, $\in_G$, be such that

$$\begin{aligned} x \in_G y \Leftrightarrow & (x, y \in \mathcal{M} \wedge (x \in y)^{\mathcal{M}}) \vee \\ & (x, y \in \mathcal{M}^{\mathbb{P}}/G \wedge x \in_G^{\mathcal{M}} y) \\ & (x \in \mathcal{M} \wedge y \in \mathcal{M}^{\mathbb{P}}/G \wedge i_G^{\mathcal{M}}(x) \in_G^{\mathcal{M}} y). \end{aligned}$$

Then we are able to ensure that (from the perspective of $\mathcal{M}_{ult}[G]$) $\mathcal{M}_{ult}[G]$ is a generic extension of $\mathcal{M}$ in the conventional sense of say Kunen (2006).

PROPOSITION 23.

(1) $Ord^{\mathcal{M}} = Ord^{\mathcal{M}_{ult}[G]}$ *and* $\mathcal{M}_{ult}[G]$ *is an end extension of* $\mathcal{M}$.[86]
(2) $\mathcal{M}_{ult}[G]$ *thinks it is a generic extension of* $\mathcal{M}$ *by* G *over* $\mathbb{P}$; *i.e.*, $x \in \mathcal{M}_{ult}[G]$ *iff there is some* τ *in* $\mathcal{M}^{\mathbb{P}}$ *such that* $(Val(\tau, [\dot{G}]_G^{\mathcal{M}}) = x)^{\mathcal{M}_{ult}[G]}$.
(3) *There exists* $r \in \mathcal{M}_{ult}[G]$ *such that* $i_G^{\mathcal{M}}``\mathcal{M} = (W_r)^{\mathcal{M}_{ult}[G]}$ *if* $\mathbb{P} \in \mathcal{M}$.

If $\mathcal{M}$ is transitive, let us use the standard notation $\mathcal{M}[G]$ to denote the $\{Val(\sigma, G) \mid \sigma \in \mathcal{M}^{\mathbb{P}}\}$.[87] Turning our attention back to the case where $\mathbb{P} = (Col(\omega, <))^{\mathcal{M}}$, we see that $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$ is able to define all of the generic refinements of $\mathcal{M}_{ult}[G]$.[88]

PROPOSITION 24. *Let* $\mathcal{M} \models ZFC$ *and* G *be* $Col(\omega, <Ord)^{\mathcal{M}}$*-generic over* $\mathcal{M}$. *Then for all* $\alpha \in Ord^{\mathcal{M}}$

(1) $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle \models ZFC^{-}(\mathcal{M})$;
(2) $G \restriction \alpha \in \mathcal{M}$;
(3) $(\mathcal{M}[G \restriction \alpha])^{\mathcal{M}_{ult}[G]}$ *is definable in the model* $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$.
(4) *For all* $\alpha \in Ord^{\mathcal{M}}$ *and all* $r \in \mathcal{M}[G \restriction \alpha]$, $(W_r)^{\mathcal{M}[G \restriction \alpha]}$ *is definable in*

$$\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle.$$

[86] $\mathcal{N} = \langle N, F \rangle$ is an end extension of $\mathcal{M} = \langle M, E \rangle$ iff whenever $xFy \in N \cap M$, $x \in M$.
[87] This only makes sense if $\mathcal{M}$ is well-founded.
[88] Note that since this is a class forcing, we cannot make use of Laver and Woodin's theorem.

A.2. Soundness. With this in hand, we can now provide a general definition of Steel's generic multiverse.

Definition 25. Suppose $\mathcal{M}$ is a model of ZFC and G is $Col(\omega,<Ord)^{\mathcal{M}}$-generic over $\mathcal{M}$. We let $\mathcal{M}^G$ be the model with sets from $\mathcal{M}_{ult}[G]$ and whose worlds are those $\mathcal{N}$ such that there exists $\alpha \in Ord^{\mathcal{M}}$ and $r \in \mathcal{M}[G \restriction \alpha]$ such that[89]

$$x \in \mathcal{N} \Leftrightarrow \langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle \models x \in (W_r)^{\mathcal{M}[G \restriction \alpha]}.$$

Recall that – by our definitions – if $\mathcal{M}$ is transitive, then $\mathcal{M}_{ult}[G] = \mathcal{M}[G]$ and observe that if $\mathcal{N}$ is a world in $\mathcal{M}^G$, then $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$ thinks that $\mathcal{N}[H] = \mathcal{M}[G \restriction \alpha]$ for some $\mathcal{N}$-generic H.[90] We now show that $\mathcal{M}^G$ provides a model of our MV. This is the required fact for establishing soundness.

Theorem 26 (Steel). *Suppose $\mathcal{M}$ is a model of ZFC and G is $(Col(\omega,<Ord))^{\mathcal{M}}$-generic over $\mathcal{M}$. Then $\mathcal{M}^G \models MV$.*

Proof. By Fact 18 we see that $\langle \mathcal{M}[G], \mathcal{M}, \in_G \rangle \models ZFC^-(\mathcal{M})$. Then using Proposition 24, it can be seen that $\mathcal{M}[G \restriction \alpha]$ and all of its generic refinements are definable in $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$. Moreover, from the perspective of $\langle \mathcal{M}[G], \mathcal{M}, \in_G \rangle$ each of these worlds is transitive. This is sufficient for us to carry out the rest of the proof within $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$.

(MV-0) This holds by definition. (MV- 1_φ) Let φ be an axiom of ZFC and let $\mathcal{N}$ be some world in $\mathcal{M}^G$. Then $\mathcal{N} \models ZFC$ by definition. We note that (MV-2) and (MV-3) are entailed by (MV-7), so we'll cover the former two with a proof of the latter at the end. (MV-4) This holds by the definition of $\mathcal{M}^G$.

(MV-5: Extension) Suppose $\mathcal{N}$ is a world and $\mathbb{P} \in \mathcal{N}$ is a poset. Fix $\mathbb{Q} \in \mathcal{N}$ and H, which is $\mathbb{Q}$-generic over V, such that $\mathcal{N}[H] = \mathcal{M}[G \restriction \alpha]$ for some $\alpha \in Ord^{\mathcal{M}}$. Fix $\beta \in Ord^{\mathcal{M}}$ such that $\beta > |\mathbb{P}|^{\mathcal{N}}$. Then we may fix a complete embedding $\sigma : \mathbb{P} \to Col(\omega, [\alpha, \beta))$ with $\sigma \in \mathcal{N}$.[91] Then we see that $J = \sigma^{-1}(G \restriction [\alpha, \beta))$ is $\mathbb{P}$-generic over $\mathcal{M}[G \restriction \alpha]$; and since $\mathcal{N} \subseteq \mathcal{M}[G \restriction \alpha]$, J is also $\mathbb{P}$-generic over $\mathcal{N}$. Finally since $\mathcal{N} \subseteq \mathcal{N}[J] \subseteq \mathcal{M}[G \restriction \beta]$, we see by the quotient lemma that there is some $\mathbb{R} \in \mathcal{N}[J]$ and K, which is $\mathbb{R}$-generic over $\mathcal{N}[J]$, such that $\mathcal{M}[G \restriction \beta] = \mathcal{N}[J][K]$.[92] Thus $\mathcal{N}[J]$ is a generic refinement of $\mathcal{M}[G \restriction \beta]$ and is thus, a world.

(MV-6: Refinement) Suppose $\mathcal{N} \in \mathcal{M}^G$ and $\mathcal{N} = \mathcal{V}[J]$ for J, $\mathbb{P}$-generic over $\mathcal{V}$. Since $\mathcal{N} \in M^G$, we may fix $\mathbb{Q} \in \mathcal{N}$ and H, $\mathbb{Q}$-generic over $\mathcal{N}$, such that

$$\mathcal{N}[H] = \mathcal{M}[G \restriction \alpha]$$

[89] Thanks to Goldberg for suggesting the framework which allows for this simple definition.

[90] We don't need to expand the signature to accommodate $\mathcal{N}$ since Laver and Woodin's theorem guarantees that $\mathcal{N}$ is definable from $\mathcal{M}[G \restriction \alpha]$ in $\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle$.

[91] This can be obtained using the embedding of Proposition 10.20 of Kanamori (2003).

[92] For the quotient lemma see, for example, exercises VII(D4) and VII(D5) in Kunen (2006).

for some α. But then $\mathcal{M}[G \restriction \alpha] = \mathcal{V}[J][H]$ is a generic extension of V by the iteration lemma and so $\mathcal{V} \in \mathcal{M}^G$.[93]

(MV-7: Amalgamation) Let $\mathcal{N}$ and $\mathcal{V}$ be worlds such that $\mathcal{N}[H] = \mathcal{M}[G \restriction \alpha]$ and $\mathcal{V}[I] = \mathcal{M}[G \restriction \beta]$ where H and I are $\mathcal{N}$-generic and $\mathcal{V}$-generic respectively and $\alpha, \beta \in Ord^{\mathcal{M}}$. Without loss of generality, suppose that $\alpha \geq \beta$. Then

$$\begin{aligned} \mathcal{N}[H] &= \mathcal{M}[G \restriction \alpha] \\ &= \mathcal{M}[G \restriction \beta][G \restriction [\beta, \alpha)] \\ &= \mathcal{V}[I \times G \restriction [\beta, \alpha)]. \end{aligned}$$ ⊣

A.3. Completeness. Now we show that any countable model of MV is of the desired natural form. This is the fact required for the completeness proof.

THEOREM 27 (Steel). *Let $\mathcal{W}$ be a model of MV satisfying extensionality for worlds with $\mathcal{M} \in \mathcal{W}$ and $|\mathcal{W}| = \omega$. Then for some $Col(\omega, {<}Ord^{\mathcal{M}})$-generic H*

$$\mathcal{W} = \mathcal{M}^H.$$

PROOF. We shall define a sequence of generics $\langle G_i \mid i \in \omega \rangle$ and indices $\langle n_i \mid i \in \omega \rangle$ which will allow us to capture every world in $\mathcal{W}$ as a generic refinement. Then using Lemma 28, we will use $\langle G_i \mid i \in \omega \rangle$ to define our desired H. Let $\langle \kappa_n \mid n \in \omega \rangle$ be a sequence of $\mathcal{M}$-cardinals which is cofinal in $\mathcal{M}$. Let $\langle \mathcal{N}_n \mid n \in \omega \rangle$ enumerate the worlds in $\mathcal{W}$. Let

- Let n_0 be least such that there exists $\mathrm{Col}(\omega, \kappa_{n_0})$-generic G over $\mathcal{M}$ where $\mathcal{N}_0$ is a ground of $\mathcal{M}[G] \in \mathcal{W}$. Let G_0 be such a G.
- Let n_{i+1} be least such that there exists $\mathrm{Col}(\omega, [\kappa_{n_i}, \kappa_{n_{i+1}}))$-generic G over $\mathcal{M}[\prod_{j \leq i} G_j]$ where $\mathcal{N}_{i+1}$ is a ground of $\mathcal{M}[\prod_{j \leq i} G_j \times G] \in \mathcal{W}$. Let G_{i+1} be such a G.

It can be seen using the universality properties of collapse forcing and the axioms of MV that this sequence is well-defined in the sense that such an n_i and G_i exist for all $i \in \omega$. $\langle G_i \mid i \in \omega \rangle$ can then be used to define a sequence $\langle G^*_\alpha \mid \alpha {<} \mathrm{Ord}^{\mathcal{M}} \rangle$ where each G^*_α is $\mathrm{Col}(\omega, \{\alpha\})$-generic over $\mathcal{M}$ and every product of a finite subsequence is generic over $\mathcal{M}$. We then use Lemma 28 to obtain H which is $\mathrm{Col}(\omega, {<}\mathrm{Ord}^{\mathcal{M}})$-generic over $\mathcal{M}$ and such that for all $\alpha \in \mathrm{Ord}^M$

$$\mathcal{M}[H \restriction \{\alpha\}] = \mathcal{M}[G^*_\alpha].$$

CLAIM. $\mathcal{M}^H = \mathcal{W}$.

PROOF. ($\subseteq$) Suppose $N \in \mathcal{M}^H$. Fix $\mathcal{N}$-generic J such that $\mathcal{N}[J] = \mathcal{M}[H \restriction \alpha]$ for some $\alpha \in Ord^{\mathcal{M}}$. Then we see that since $\mathcal{M}[H \restriction \alpha] \in \mathcal{W}$ by construction, we must have $N \in \mathcal{W}$ by Refinement. ($\supseteq$) Suppose $\mathcal{N} \in \mathcal{W}$.

[93]See Proposition 10.9 of Kanamori (2003).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Then $\mathcal{N} = \mathcal{N}_i$ for some $i \in \omega$. Then we see from our construction that $\mathcal{N}_i$ is generically extended by $\mathcal{M}[H \upharpoonright \kappa_{n_i}]$. Thus, $\mathcal{N}_i \in \mathcal{M}^H$. $\dashv$

The following lemma allows us to take a countable set of finitely mutually generic sets and make a finite support product of equivalent generics over homogeneous posets. It's a generalization of a theorem of Hamkins. [94]

Lemma 28. *Let $\mathcal{M}$ be a countable model of ZFC. Let $\langle \mathbb{P}_\alpha \mid \alpha \in Ord^{\mathcal{M}} \rangle$ be definable in $\mathcal{M}$ where each $\mathbb{P}_\alpha \in \mathcal{M}$ is weakly homogeneous according to $\mathcal{M}$. Let $\langle G^*_\alpha \mid \alpha \in Ord^{\mathcal{M}} \rangle$ be finitely mutually generic; i.e., such that for all finite partial functions $f : \omega \rightharpoonup Ord^{\mathcal{M}}$,*

$$\prod_{i \in dom(f)} G^*_{f(i)} \text{ is } \prod_{i \in dom(f)} \mathbb{P}_{f(i)}\text{-generic over } \mathcal{M}.$$

Then there exists H such that:[95]

(1) *H is $\prod^{fin}_{\alpha \in Ord^{\mathcal{M}}} \mathbb{P}_\alpha$-generic over $\mathcal{M}$; and*
(2) *$\mathcal{M}[H \upharpoonright \{\alpha\}] = \mathcal{M}[G^*_\alpha]$ for all $\alpha \in Ord^{\mathcal{M}}$.*

Proof. We take an enumeration of the dense sets of $\mathbb{P} = \prod^{fin}_{\alpha \in Ord} \mathbb{P}_\alpha$ which are definable in $\mathcal{M}$ and a countable enumeration of the ordinals of $\mathcal{M}$ and use this to construct our generic in the fashion of the proof of the Baire category theorem. Along the way, we exploit the homogeneity assumption to make some adjustments to the elements of the $\langle G_\alpha \mid \alpha \in \mathrm{Ord}^{\mathcal{M}} \rangle$ sequence. Let $\langle D_n \mid n \in \omega \rangle$ enumerate the dense subsets of $\mathbb{P}$ which are definable in $\mathcal{M}$.

Let $f : \omega \cong Ord^{\mathcal{M}}$ enumerate the ordinals of $\mathcal{M}$. We use this to rearrange the sequence $\langle G_\alpha \mid \alpha \in Ord^{\mathcal{M}} \rangle$ into a sequence of length ω which will allow us to define the generic by constructing an ω-sequence of points which intersects every dense set. To implement this and make the following more readable, we let $\mathbb{P}^\dagger_n = \mathbb{P}_{f(n)}$; $\mathbb{P}^\dagger = \prod^{fin}_{n \in \omega} \mathbb{P}^\dagger_n$ and $G^\dagger_n = G^*_{f(n)}$ for all $n \in \omega$. Then for $p \in \mathbb{P}$, let $p^\dagger(n) = p(f(n))$; and for $n \in \omega$ let $D^\dagger_n = \{p^\dagger \mid p \in D_n\}$.[96] We'll define sequences $\langle H^\dagger_n \mid n \in \omega \rangle$ and $\langle \bar{p}_n \mid n \in \omega \rangle$ by recursion such that for all $n \in \omega$:

- $\mathcal{M}[H^\dagger_n] = \mathcal{M}[G^\dagger_n]$; and
- $\bar{p}_n \in D^\dagger_n \cap \prod^{fin}_{n \in \omega} H^\dagger_n$ and $\bar{p}_n \leq \bar{p}_m$ for all $m < n$.

We then let $H_\alpha = H^\dagger_{f^{-1}(\alpha)}$ for all $\alpha \in Ord^M$ and $H = \prod^{fin}_{\alpha \in Ord^M} H_\alpha$. This will suffice for the lemma. To obtain this we just need to define these sequences such that for all $n \in \omega$, $\bar{p}_n \in D^\dagger_n$, $\bar{p}_n \upharpoonright (n+1) \in H^\dagger_0 \times \cdots \times H^\dagger_n$ where $\mathcal{M}[G^\dagger_n] =$

[94] See Theorem 13 of (Hamkins, 2015). The setting of that theorem is the concrete context of adding Cohen reals and, as such, it may be helpful for the reader to consult that proof first.

[95] I'm using Kunen's notation for infinite products with finite support; i.e., $\prod^{fin}_{n \in \omega} \mathbb{P}_n = \{\bar{p} \in \prod_{n \in \omega} \mathbb{P}_n \mid |\{n \in \omega \mid \bar{p}(n) \neq 1_{\mathbb{P}_n}\}| < \omega\}$.

[96] It's probably worth observing that $\mathbb{P}^\dagger$ cannot be defined in $\mathcal{M}$, while $\prod_{i<n} \mathbb{P}^\dagger_i \in \mathcal{M}$ for all $n \in \omega$.

$\mathcal{M}[H_n^\dagger]$ and $\bar{p}_n \leq \bar{p}_m$ for all $m < n$. Suppose we've done this up to n. We define $\bar{p}_{n+1}$ and $H_{n+1}^\dagger$ as follows. First we note that there is some $\bar{p} \leq \bar{p}_n$ such that $\bar{p} \in D_{n+1}^\dagger$ and $\bar{p} \restriction (n+1) \in H_0^\dagger \times \cdots \times H_n^\dagger$. To see this observe that

$$D^* = \{q \in \mathbb{P}_0^\dagger \times \cdots \times \mathbb{P}_n^\dagger \mid \exists \bar{q} \in D_{n+1}^\dagger\ (\bar{q} \restriction (n+1) = q \wedge \bar{q} \leq \bar{p}_n)\}$$

is dense below $\bar{p}_n \restriction (n+1)$. Thus, we may fix some $p \in D^* \cap (H_0^\dagger \times \cdots \times H_n^\dagger)$ and thus some $\bar{p} \in \mathbb{P}^\dagger$ witnessing that $\bar{p} \in D^*$. Let $\bar{p}_n$ be such a $\bar{p} \in \mathbb{P}^\dagger$. Now it might not be the case that $\bar{p}_n \in H_0^\dagger \times \cdots \times H_n^\dagger \times G_{n+1}^\dagger$. So using the weak homogeneity of $\mathbb{P}_{n+1}$, fix $\sigma : \mathbb{P}_{n+1}^\dagger \cong \mathbb{P}_{n+1}^\dagger$ such that $\sigma(\bar{p}_n(n)) = g$ for some $g \in G_0^\dagger$. Then let $H_{n+1}^\dagger = (\sigma^{-1})``G_{n+1}^\dagger$. This ensures that the sequences have the desired properties and the result follows. ⊣

A.4. Main theorem. The soundness and completeness theorems now follow easily.

THEOREM 8. *For all φ in the multiverse language, the following are equivalent*:

(1) $MV \vdash \varphi$.
(2) *If $\mathcal{M}$ is a countable model of ZFC, then $\mathcal{M}^G \models \varphi$, where G is $Col(\omega, <Ord)^{\mathcal{M}}$-generic over $\mathcal{M}$.*

PROOF. $(1 \to 2)$ Suppose $MV \vdash \varphi$ and $\mathcal{M}$ is a countable model of ZFC. Then by Theorem 26, we see that $\mathcal{M}^G \models ZFC$ and so by (1) and the soundness theorem, $\mathcal{M} \models \varphi$.

$(2 \to 1)$ Suppose $\mathcal{W}$ is an arbitrary model of MV. Taking a Skolem hull if necessary, we may assume that $\mathcal{W}$ is countable. Then by Theorem 27 $\mathcal{W} = \mathcal{M}^G$ for some $Col(\omega, <\text{Ord})^{\mathcal{M}}$-generic over $\mathcal{M}$. Thus, by (2) we see that $\mathcal{W} \models \varphi$ and by the completeness theorem, we have $MV \vdash \varphi$. ⊣

B. Unaxiomatisability of $\mathbb{V}_{\mathcal{M}}$. In this section, we show that Woodin's generic multiverse – unlike Steel's – is not amenable to axiomatisation. First recall the following theorem of Usuba.

THEOREM 29. (*Usuba*) *For all r_0, r_1 there exists s such that*

$$W_s \subseteq W_{r_0} \cap W_{r_1}.$$

Less formally, this tells us that when V is a generic extension of U_0 and U_1, then V is a generic extension of some $U_2 \subseteq U_0 \cap U_1$. Now note the following Corollary of Theorem 29:

COROLLARY 30. *Suppose N is a countable transitive model of ZFC. Then $N \in \mathbb{V}_M$ iff there exists $r \in M$, $\mathbb{P} \in (W_r)^M$ and $\mathbb{P}$-generic G over $(W_r)^M$, such that*

$$N = (W_r)^M[G].$$

In other words, $\mathbb{V}_M$ is the set of ctms N accessible by a generic refinement followed by a generic extension from M. We use this fact to define a provisional version of Woodin's generic multiverse adapted to the non-well-founded setting. The obvious way to attempt to form $\mathbb{V}_{\mathcal{M}}$ over a possibly ill-founded $\mathcal{M}$ would be just use the corollary above with $\mathcal{M}$ instead of transitive M. However, this leads to problems. For example, suppose M is transitive; $\mathbb{P}, \mathbb{Q} \in M$ are posets; G is $\mathbb{P}$-generic over M and H is $\mathbb{Q}$-generic over $M[G]$. Then $G \times H$ is $\mathbb{P} \times \mathbb{Q}$-generic over M and importantly, $M[G][H] = M[G \times H]$. However, the situation is different if we use the ill-founded method above. If $\mathcal{M}$ is ill-founded, $\mathbb{P}, \mathbb{Q} \in \mathcal{M}$ are posets; G is $\mathbb{P}$-generic over $\mathcal{M}$ and H is $\mathbb{Q}$-generic over $\mathcal{M}[G]$, then we do get that $G \times H$ is $\mathbb{P} \times \mathbb{Q}$-generic over $\mathcal{M}$; but in general $\mathcal{M}[G][H]$ is merely isomorphic and not identical to $\mathcal{M}[G \times H]$.

We address this problem by identifying – for any model $\mathcal{M}$ of *ZFC* – a canonical structure that is simply definable over $\mathcal{M}$ and from which $\mathcal{M}$ may be recovered up to isomorphism. These structures will be used to represent the worlds of our generalised generic multiverse. (This definition is due to Woodin; we're grateful for his permission to include it here.)

For $\mathcal{M} = \langle M, \in_{\mathcal{M}} \rangle$ a countable model of *ZFC*, consider the structure

$$\langle Ord^{\mathcal{M}}, \in_{\mathcal{M}}, OrdRel(\mathcal{M}) \rangle,$$

where

$$OrdRel(\mathcal{M}) = \{p \in \mathcal{M} \mid \mathcal{M} \models p \subseteq Ord \times Ord\}$$

is the set of ordinal-domain relations according to $\mathcal{M}$. Call this the *ord-structure* of $\mathcal{M}$. We claim that there is only one model of *ZFC* up to isomorphism with this ord-structure. To see this suppose $\mathcal{N} = \langle N, \in_{\mathcal{N}} \rangle$ is a countable model of *ZFC* that has the same ord-structure as $\mathcal{M}$. Let $\pi : M \to N$ be defined as follows. For each $x \in M$, fix $A_x \in OrdRel(\mathcal{M})$ be such that $\mathcal{M}$ thinks that the transitive collapse of A_x is the transitive closure of $\{x\}$. Let $\pi(x)$ be the $\in_{\mathcal{N}}$-greatest element of what $\mathcal{N}$ thinks is the transitive collapse of A_x. This makes sense since $A_x \in OrdRel(\mathcal{M}) = OrdRel(\mathcal{N})$. We can then see that π is an isomorphism but we'll just show that π is an injection as the rest of the proof is similar. Suppose that $\pi(x) = \pi(y)$. Then $\mathcal{N}$ thinks that the transitive collapse of A_x is the transitive collapse of A_y. This means that there is some $f \in \mathcal{N}$ such that $\mathcal{N}$ thinks that $f : A_x \cong A_y$. And since it is clear that $f \in OrdRel(\mathcal{N}) = OrdRel(\mathcal{M})$, we see that $\mathcal{M}$ also thinks that the transitive collapse of A_x is the transitive collapse of A_y. Thus, $x = y$.

Now we can generalise the generic multiverse to the case of ill-founded models. Recall that we have set up a forcing definition for ill-founded models such that generic extensions literally contain their grounds as submodels. Thus, it is easy to see that every world in the generic multiverse uses the same set of objects for ordinals and has the same ordering upon them. This means that we can represent each world $\mathcal{N}$ in the multiverse by $OrdRel(\mathcal{N})$ as that is the only part of the ord-structure which varies.

Definition 31. Let $\mathbb{V}_{\mathcal{M}}$ be the set of $OrdRel(\mathcal{N})$ for which there exists $r \in \mathcal{M}$, $\mathbb{P} \in (W_r)^{\mathcal{M}}$, $\mathbb{P}$-generic G over $(W_r)^{\mathcal{M}}$ such that

$$\mathcal{N} = ((W_r)^{\mathcal{M}})_{ult}[G].$$

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Now we might worry that since we are representing worlds by sets of ordinal-domain relations, we are restricted to only being able to compare such relations between worlds and not arbitrary sets in those worlds. This is – in general – correct and appears to be forced upon us by our use of ill-founded models. However, much can still be recovered. For example, if we start with a countable transitive model M, then we can recover the Woodin's original generic multiverse by taking a model $\mathcal{N}$ for each world in the multiverse and then collapsing each of those models into transitive models.[97]

In the more general case of ill-founded models, we cannot perform this collapse. However, we can still unambiguously *identify* many sets across different worlds. We illustrate this with a couple of examples. Suppose $\mathcal{N}_0$ and $\mathcal{N}_1$ are in $\mathbb{V}_{\mathcal{M}}$ where $\mathcal{M}$ is ill-founded. First, suppose $x \in L^{\mathcal{N}_0}$. Then let $\alpha \in Ord^{\mathcal{N}_0}$ be such that $\mathcal{N}_1 \models Enum(x, \alpha)$ where $Enum(v_0, v_1)$ a formula of $\mathcal{L}_\in$ which says that x is the α^{th} set in the canonical enumeration of L and this formula absolute for L. Then it is clear that x should be identified with that $y \in \mathcal{N}_1$ where $\mathcal{N}_1 \models Enum(y, \alpha)$. Second we note that if $X_0 \in \mathcal{N}_0$ is what $\mathcal{N}$ thinks is a family of ordinal-domain relations, then we can check whether X_0 is *represented* in $\mathcal{N}_1$ by asking whether there is some $X_1 \in \mathcal{N}_1$ such that for all $p \in OrdSet(\mathcal{N}_0) = OrdSet(\mathcal{N}_1)$ such that

$$\mathcal{N}_0 \models p \in X_0 \;\Leftrightarrow\; \mathcal{N}_1 \models p \in X_1.$$

This allows us to – so to speak – find the reals of one world in any other, if they are present. We can then generalise this to compare families of families of ordinal-domain relations; and we can continue iterating this idea along the well-founded part of $Ord^{\mathcal{M}}$, but not beyond. Even in the case where our generating model $\mathcal{M}$ is not an ω-model this gives us more than enough agreement to be able to carry out the following argument.

B.1. $\mathbb{V}_{\mathcal{M}}$ is unaxiomatisable. We now show that there cannot be a recursive axiomatization of $\mathbb{V}_{\mathcal{M}}$. (These results are due to Woodin who has kindly allowed us to include them here.) Our strategy is to demonstrate that – in fact – the true theory of arithmetic can be obtained from the theory of any $\mathbb{V}_{\mathcal{M}}$! To do this we show that the *real* ω can be defined in any such $\mathbb{V}_{\mathcal{M}}$ - even though its version of ω may be nonstandard. We then define a translation which gives us the required reduction. Given a countable model $\mathcal{M}$ of ZFC, we shall suppose below – without loss of generality – that the real V_ω is a submodel of the well-founded part of $\mathcal{M}$.

Lemma 32 (Woodin). *For $\mathcal{M}$ an arbitrary (and possibly ill-founded) countable model of ZFC, there is a formula $\varphi_\omega \in \mathcal{L}_\in$ such that for all $x \in \omega^{\mathcal{M}}$*

$$x \in \omega^V \;\Leftrightarrow\; \mathbb{V}_{\mathcal{M}} \models \varphi_\omega(x).$$

Proof. Our strategy is as follows. We first show that in $\mathbb{V}_{\mathcal{M}}$, there exist reals $c, d \in (2^\omega)^{\mathcal{M}}$ which instantiate an alternating pattern which has ω^V-many alternations. We call such a pair an *ω-pair*. We then observe that there

[97]The extra collapse step is required since we are using our ultrapower approach to forcing.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

is a formula φ_{alt} which when given reals $c, d \in (2^\omega)^{\mathcal{M}}$ can identify how many alternations have occurred according to $\mathcal{M}$. From the perspective of V, this could be a nonstandard number of alternations. We then use φ_{alt} to define a formula which identifies the *real* natural numbers.

To see that an ω-pair exists in $\mathbb{V}_{\mathcal{M}}$, let $\langle D_n \rangle_{n\in\omega}$ enumerate (in V) the subsets of $\mathbb{P} = (2^{<\omega})^{\mathcal{M}}$ which are dense according to $\mathcal{M}$. We define c and d by recursion as follows. Let $c_0 \in \mathcal{M}$ be such that $\mathcal{M} \models c_0 \in D_0$ and let p_0 be the empty sequence.

Let $d_n \in \mathcal{M}$ be such that the following are satisfied in $\mathcal{M}$:

- $d_n \in D_n$; and
- $d_n = d_{n-1}^\frown s_n^\frown \langle 1 \rangle^\frown q_{n+1}$ where s_n is a sequence of length $lh(c_n)$ which constantly outputs 0 and where we let d_{n-1} be the empty sequence if $n = 0$.

Let $c_{n+1} \in \mathcal{M}$ be such that the following are satisfied in $\mathcal{M}$:

- $c_{n+1} \in D_{n+1}$; and
- $c_{n+1} = c_n^\frown t_n^\frown \langle 1 \rangle^\frown p_{n+1}$ where t_n is a sequence of length $lh(q_n)$ which constantly outputs 0.

It should then be clear that the sequences $\langle c_n \rangle_{n\in\omega}$ and $\langle d_n \rangle_{n\in\omega}$ can be used to define Cohen reals for $\mathcal{M}$. The following diagram might be helpful.

<table>
<tr><td>c</td><td>c_0</td><td colspan="2">0's</td><td>1</td><td>p_1</td><td colspan="2">0's</td><td>1</td><td>p_2</td><td>...</td></tr>
<tr><td>d</td><td>0's</td><td>1</td><td>q_0</td><td colspan="2">0s</td><td>1</td><td>q_1</td><td colspan="2">0's</td><td>...</td></tr>
</table>

Since these sequences clearly yield $\mathcal{M}$-generics, it can be seen that $c = \bigcup_n c_n \in 2^\omega$ and $d = \bigcup_n d_n \in 2^\omega$ are both represented in $\mathbb{V}_{\mathcal{M}}$. We now define φ_ω. But first we let $\varphi^*_{alt}(x, c, d)$ be a formula in the language of arithmetic with two function parameters which says that there is some β-function with domain x tracking some of the alternating pattern of 0's which could occur between c and d as outlined above. So if we were working in the standard model of arithmetic $\mathbb{N}$, then $\langle \mathbb{N}, c, d \rangle \models \varphi^*_{alt}(n, c, d)$ holds when at least n blocks of 0's occur in the pattern described above.

Then let $\varphi_{alt}(x, c, d)$ be the standard translation of $\varphi^*_{alt}(x, c, d)$ from the language of arithmetic into the language of set theory. This means that if $c, d \in (2^\omega)^{\mathcal{M}}$ then we'll have $\mathcal{M} \models \varphi_{alt}(x, c, d)$ iff $\mathcal{M}$ thinks there are at least $x \in \omega^{\mathcal{M}}$ many blocks of 0's occurring as in the diagram above, where x could – in general – be nonstandard. Note also that since $\omega^{\mathcal{M}} = \omega^{\mathcal{N}}$ for all worlds $\mathcal{N}$, we'll have $\mathcal{M} \models \varphi_{alt}(x, c, d)$ iff $\mathbb{V}_{\mathcal{M}} \models \varphi_{alt}(x, c, d)$. In this situation, let us say that x is *captured by* c *and* d in the sense that there at least x many alternations in the pattern instantiated by c and d. We then let $\varphi_\omega(x)$ say that for all $c, d \in 2^\omega$

- if $\forall y(\varphi_{alt}(y, c, d) \rightarrow \exists z(z > y \wedge \varphi_{alt}(z, c, d)))$,
- then $\varphi_{alt}(x, c, d)$.

Informally, this says that x is captured by every pair c, d that tracks an alternation which has a limit length.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Claim. *For all $x \in \omega^{\mathcal{M}}$*

$$x \in \omega^V \;\Leftrightarrow\; \mathbb{V}_{\mathcal{M}} \models \varphi_\omega(x).$$

Proof. ($\Rightarrow$) Suppose $x \in \omega^V$. Suppose that $c, d \in \mathbb{V}_{\mathcal{M}}$ support an alternation of limit length, then the length of the alternation cannot be finite, so x must be captured by c and d. ($\Leftarrow$) Suppose $x \notin \omega^V$. Then $x \in \omega^{\mathcal{M}}$ and must be nonstandard. Work in $\mathbb{V}_{\mathcal{M}}$. Fix an ω-pair c, d in $\mathbb{V}_{\mathcal{M}}$. Then it can be seen that x is not captured by c and d since they only capture V-finite naturals. ⊣

Lemma 33 (Woodin). *There is a recursive function $f : \omega \to \omega$ such that for all countable models $\mathcal{M}$ of ZFC and all sentences ψ in the language of arithmetic, we have*

$$\mathbb{N} \models \psi \;\Leftrightarrow\; f(\psi) \in Th(\mathbb{V}_{\mathcal{M}}),$$

where $\mathbb{N}$ is the standard model of arithmetic.

Proof. Using Lemma 32, we can define ω with a formula $\varphi_\omega(v_0) \in \mathcal{L}_{MV}$ in the $\mathbb{V}_{\mathcal{M}}$ for any countable $\mathcal{M}$. For a formula ψ from the language of arithmetic, we let $f(\psi)$ be the result of: first using the standard translation of arithmetic into set theory; and then relativising all the quantifiers using the formula $\varphi_\omega(v_0)$. ⊣

With this in hand, we are able to establish the main result of this section.

Theorem 34 (Woodin). *There is no recursive $T \subseteq \mathcal{L}_{MV}$ such that for all $\varphi \in \mathcal{L}_{MV}$*

$$T \vdash \varphi \Leftrightarrow \forall \mathcal{M}(|\mathcal{M}| = \omega \wedge \mathcal{M} \models ZFC \to \mathbb{V}_{\mathcal{M}} \models \varphi).$$

Proof. Suppose not and fix such a recursive T. Let f be the recursive function given by Theorem 33. Then we see that

$$\begin{aligned} \varphi \in Th(\mathbb{N}) &\Leftrightarrow \forall \mathcal{M}\ (|\mathcal{M}| = \omega \wedge \mathcal{M} \models \mathrm{ZFC} \to \mathbb{V}_{\mathcal{M}} \models f(\varphi)) \\ &\Leftrightarrow T \vdash f(\varphi). \end{aligned}$$

But this means that $Th(\mathbb{N})$ can be computed from T which means that T is not recursive. ⊣

C. The translation. In this section, we provide a definition of the Steel's translation function and show that it works.[98]

Definition 35. Assume M is an inner model of ZFC.

[98]Thanks to Goldberg for providing this simplified approach to the proof. Our original strategy provided a direct proof of Theorem 9 from the theory MV. However, the strategy was extremely syntactic and difficult to read.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

- A triple $\langle r, \mathbb{P}, G\rangle$ *defines a world U relative to M* if $(W_r)^M$ is defined, $\mathbb{P}$ is a poset in $(W_r)^M$, and $G \subseteq \mathbb{B}$ is a $(W_r)^M$-generic ultrafilter such that

$$U = (W_r)^M[G].$$

- Let X_M denote the class of triples that define some world relative to M.
- Let $\sim$ be the equivalence relation defined on X_M by $x \sim y$ if x and y define the same world.
- Let S_M be the class of equivalence classes of $\sim$ using Scott's trick.

The following definition is carried out in the multiverse language.

DEFINITION 36. Let f be the function sending a pair of worlds $\langle M, U\rangle$ to the $\sim$-equivalence class in S_M of triples that represent U relative to M.

PROPOSITION 37. *MV proves that f is a total function.*

PROOF. Suppose U and M are worlds. Using Amalgamation fix a world U^* such that U and M are both grounds of U^*. Then by the Downward Directed Grounds theorem in U^* fix $U^\dagger$ such that $U^\dagger$ is a ground of both U and M. Thus there exists $r \in M$, $\mathbb{P} \in (W_r)^M$ and $G \in M$ which is $\mathbb{P}$-generic over $(W_r)^M$ such that $U = (W_r)^M[G]$. ⊣

LEMMA 38. *There is a total recursive function $e : \mathcal{L}_{MV} \to \mathcal{L}_\in(\dot{M})$ with the following property. Suppose $\mathcal{W} \models MV, M$ is a world of $\mathcal{W}$, and N is the collection of sets of $\mathcal{W}$ and $\in_\mathcal{W}$ is the membership relation in $\mathcal{W}$. Then*

$$\mathcal{W} \models \varphi(\vec{x}, U_0, \dots, U_{n-1}) \Leftrightarrow \langle N, M, \in_\mathcal{W}\rangle \models e(\varphi)(\vec{x}, u_0, \dots, u_{n-1})$$

where $u_i = (f(M, U_i))^\mathcal{W}$ for $i < n$.

PROOF. The function e is defined by recursion on formula complexity:

- If x is a set variable and u is a world variable,

$$e(x \in U) = \exists (r, \mathbb{B}, G) \in u\ (x \in W_r^{\dot{M}}[G]).$$

- For all atomic formulae φ not covered by the previous bullet, $e(\varphi) = \varphi$.
- $e(\neg\varphi) = \neg e(\varphi)$ and $e(\varphi \wedge \psi) = e(\varphi) \wedge e(\psi)$.
- If x is a set variable $e(\exists x \varphi) = \exists x\ e(\varphi)$.
- If U is a world variable, $e(\exists U \varphi) = \exists u \in S_{\dot{M}}\ e(\varphi)$.

A simple induction on the complexity of formulae of $\mathcal{L}_\in$ then suffices. ⊣

DEFINITION 39. If φ is a sentence in the multiverse language, then $t(\varphi)$ is the sentence in the language of set theory asserting that

$$\Vdash_{Col(\omega, <Ord)} e(\varphi)^*,$$

where we use the forcing language $\mathcal{FL}_\mathbb{P}(\check{V})$ and where $e(\varphi)^*$ is the result of replacing every instance of $\dot{M}$ in $e(\varphi)$ with $\check{V}$.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

THEOREM 10 (Steel). *If $\mathcal{M}$ is a countable model of ZFC, G is $Col(\omega, <Ord)^{\mathcal{M}}$-generic over $\mathcal{M}$ and $\varphi \in \mathcal{L}_{MV}$, then*:

$$\mathcal{M} \models t(\varphi) \Leftrightarrow \mathcal{M}^G \models \varphi.$$

PROOF. ($\Rightarrow$) Suppose $\mathcal{M}^G \models \varphi$. Then by Theorem 26 we see that $\mathcal{M}^G \models MV$ and so by Lemma 38

$$\langle \mathcal{M}_{ult}[G], \mathcal{M}, \in_G \rangle \models e(\varphi).$$

And by Lemma 18, we see that

$$\exists p \in G\, \mathcal{M} \models \text{“}p \Vdash_{Col(\omega,<Ord)} e(\varphi)^*\text{”}.$$

Then since $e(\varphi)^*$ only uses $\check{}$-names (i.e., $\check{V}$) we see by the homogeneity of collapse forcing[99] that:

$$\mathcal{M} \models \text{“} \Vdash_{Col(\omega,<Ord)} e(\varphi)^*\text{”};$$

i.e., $\mathcal{M} \models t(\varphi)$. ($\Leftarrow$) Similar. ⊣

THEOREM 9. *There is a recursive function $t : \mathcal{L}_{MV} \to \mathcal{L}_{\in}$ such that for all sentences $\varphi \in \mathcal{L}_{MV}$, MV proves that the following are equivalent*:

(1) φ;
(2) $\forall U t(\varphi)^U$;
(3) $\exists U t(\varphi)^U$.

PROOF. Let $\mathcal{W} \models MV$ and taking a Skolem hull if necessary suppose that $\mathcal{W}$ is countable. It suffices to show that (1), (2), and (3) have the same truth value in $\mathcal{W}$. First we note that by Theorem 27, $\mathcal{W} = \mathcal{M}^G$ for some world $\mathcal{M}$ in $\mathcal{W}$ and $Col(\omega,<)^{\mathcal{M}}$-generic G over $\mathcal{M}$. Suppose φ is true and let U be an arbitrary world. By Theorem 10 we see that $t(\varphi)^U$ holds. Thus $\forall U\, t(\varphi)^U$ and since $\mathcal{W}$ must contain at least one world we have $\exists U\ t(\varphi)^U$. Now suppose $\exists U\ t(\varphi)^U$ and fix such a U. We then see by Theorem 10 that φ is true. ⊣

C.1. Applications. Now we establish some applications of the translation function. First we observe that without the Amalgamation axiom, we cannot have a translation function.

PROPOSITION 40. *There is no recursive function $s : \mathcal{L}_{MV} \to \mathcal{L}_{\in}$ such that for all $\varphi \in \mathcal{L}_{MV}$*

$$MV \textit{ minus Amagamation} \vdash \varphi \leftrightarrow \forall V\ s(\varphi)^V.$$

PROOF. Suppose not and fix such a t. Let $\mathcal{M}$ be a countable model of ZFC and let $\mathcal{M}^G$ be the multiverse defined from $\mathcal{M}$ using some GCol$(\omega, <$ Ord$^{\mathcal{M}})$-generic over $\mathcal{M}$. Let $\mathbb{V}_{\mathcal{M}}$ be the Woodin generic multiverse generated from $\mathcal{M}$. Then we see that:

(1) $\mathcal{M}^G \models$ MV minus Amalgamation + Amalgamation;
(2) $\mathbb{V}_{\mathcal{M}} \models$ MV minus Amalgamation + ¬Amalgamation.

[99] See 10.19(a) in (Kanamori, 2003).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

Then since both $\mathcal{M}^G$ and $\mathbb{V}_{\mathcal{M}}$ are models of MV minus Amalgamation, we see that:

$$\begin{aligned}\mathcal{M}^G \models \text{Amalgamation} &\Leftrightarrow \mathcal{M} \models s(\text{Amalgamation})\\ &\Leftrightarrow \mathbb{V}_{\mathcal{M}} \models \text{Amalgamation}\end{aligned}$$

which is impossible. ⊣

Now we show how large cardinals are naturally represented in the language of the multiverse.

Proposition 41 (ZFC). *$t(\forall V \exists \kappa\ (\kappa$ is measurable$)^V)$ iff there is a proper class of measurable cardinals.*

Recall the following fact from Lévy and Solovay:

Fact 42. *Let κ be a cardinal and let $\mathbb{P}$ be a poset with $|\mathbb{P}| < \kappa$. Then*

(1) *if κ is measurable, then $\Vdash_{\mathbb{P}} \kappa$ is measurable; and*
(2) *if κ is not measurable, then $\Vdash_{\mathbb{P}} \kappa$ is not measurable.*

Proof. (of Proposition 41) First observe that $\forall V \exists \kappa\ (\kappa \text{ is measurable})^V$ translates by t as

$$\forall \mathbb{P} \Vdash \forall r \exists \kappa\ (\kappa \text{ is measurable})^{W_r}$$

($\leftarrow$) Let $\mathbb{P}$ be arbitrary. Let G be $\mathbb{P}$-generic over V. Let $r \in V[G]$.[100] Then we want to show that $(W_r)^{V[G]}$ still has a proper class of measurable cardinals. First we observe that we have a definable inner model $U = (W_r)^{V[G]}$ with some $\mathbb{Q} \in U$ and $H \in V[G]$ which is $\mathbb{Q}$-generic over U and such that

$$U[H] = V[G].$$

Now by Fact 42 (1), we see that for any $\kappa > |\mathbb{P}|^V$ which is measurable according to V, κ remains measurable in $V[G]$. Thus $V[G]$ still has a proper class of measurable cardinals. Suppose now – for a contradiction – that there are no measurable cardinals in U. Let $\kappa > |\mathbb{Q}|^U$ where κ is measurable according to $U[H]$. We can obtain such a cardinal since $U[H] = V[G]$. But since U says that κ is not measurable, Fact 42 (2), tells us that κ is not measurable according to $U[H]$ either: contradiction.

($\rightarrow$) Suppose α is such that for all $\kappa > \alpha$, κ is not measurable. Let $\mathbb{P} = \text{Col}(\omega, \{\alpha\})$ and let G be $\mathbb{P}$-generic over V. Let $\kappa > \alpha$; then by Fact 42 (1), we see that κ is not measurable according to $V[G]$. Let $r \in V[G]$ be a vacuous refinement parameter; i.e., let r be such that $V[G] = (W_r)^{V[G]}$. Then we see that:

$$V[G] \models \exists r \forall \kappa (\kappa \text{ is not measurable})^{W_r}$$

[100]Now that we're familiar with the generic multiverse, we'll allow ourselves to fall back into the conventional luxury of talking about V-generics, although – of course – such talk is easily removed.

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

and so exploiting the homogeneity of $\mathbb{P}$[101] we see that

$$\Vdash_{\mathbb{P}} \exists r \forall \kappa (\kappa \text{ is not measurable})^{W_r}$$
$$\Rightarrow \nVdash_{\mathbb{P}} \forall r \exists \kappa (\kappa \text{ is not measurable})^{W_r}$$

which suffices for our claim. ⊣

We now establish that the range of the t function is the set of sentences which are provably generically invariant.

THEOREM 12. *Let $\varphi \in \mathcal{L}_{\in}$. Then following are equivalent:*

(1) $MV \vdash \forall V \varphi^V \vee \forall V \neg \varphi^V$; *and*
(2) φ *is legitimate*$_{ZFC}$.

PROOF. $(1 \to 2)$ Suppose $\varphi \in \mathcal{L}_{\in}$. Let $\mathcal{M}$ be a countable model of ZFC. Recall that

$$\mathcal{M} \models t(\forall V \varphi^V) \;\Leftrightarrow\; \mathcal{M}^G \models \forall V \varphi^V.$$

And by assumption we have

$$\mathcal{M}^G \models \neg \forall V \varphi^V \leftrightarrow \forall V \neg \varphi^V.$$

It will suffice to show that

$$\mathcal{M} \models t(\forall V \varphi^V) \leftrightarrow \varphi.$$

Suppose $\mathcal{M} \models t(\forall V \varphi^V)$. Then $\mathcal{M}^G \models \forall V \varphi^V$ and so $\mathcal{M} \models \varphi$. Suppose $\mathcal{M} \models \neg t(\forall V \varphi^V)$. Then $\mathcal{M}^G \models \neg \forall V \varphi^V$. Thus $\mathcal{M}^G \models \forall V \neg \varphi^V$ and so $\mathcal{M} \models \neg \varphi$.

$(2 \to 1)$ Suppose (1) is false and fix a multiverse $\mathcal{W}$ such that $\mathcal{W} \models MV$ and there exist worlds $\mathcal{M}_0, \mathcal{M}_1 \in \mathcal{W}$ such that

$$\mathcal{M}_0 \models \varphi \text{ and } \mathcal{M}_1 \models \neg \varphi.$$

Now suppose for a contradiction that there is some $\psi \in \mathcal{L}_{MV}$ such that

$$ZFC \vdash t(\psi) \leftrightarrow \varphi.$$

Recall that if $\mathcal{N}$ is a countable model of ZFC, then for all multiverses $\mathcal{U}$ with $\mathcal{U} \models MV$ and $\mathcal{N} \in \mathcal{U}$, we have

$$\mathcal{U} \models \chi \;\Leftrightarrow\; \mathcal{N} \models t(\chi).$$

Thus, we see that our chosen ψ is such that:

(1) $\mathcal{W} \models \psi \;\Leftrightarrow\; \mathcal{M}_0 \models \varphi$; and
(2) $\mathcal{W} \models \psi \;\Leftrightarrow\; \mathcal{M}_1 \models \varphi$

which is clearly impossible. ⊣

It is then easy to see that this generalises if we extend our theories with large cardinals in the manner outlined above. For example, letting $ZFC +$

[101] See Theorem 10.19(a) of (Kanamori, 2003).

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

$pc(EXT)$ be the theory extending MV with a proper class of extendible cardinals, in every world.

THEOREM 43. *For $\varphi \in \mathcal{L}_\in$, the following are equivalent*:

(1) $MV + \forall V pc(EXT)^V \vdash \forall V \varphi^V \vee \forall V \neg\varphi^V$; *and*
(2) φ *is legitimate*$_{ZFC+pc(EXT)}$.

We then note that against the backdrop of ZFC, CH is not legitimate. Moreover, the mere existence of a core does not alter this.

PROPOSITION 44. *We have*:

(1) $ZFC \nvdash CH \leftrightarrow t(\psi)$ *for any* $\psi \in \mathcal{L}_{MV}$; *and*
(2) $ZFC + pc(EXT) \nvdash CH \leftrightarrow t(\psi)$ *for any* $\psi \in \mathcal{L}_\in$.

PROOF. (1) Suppose not. Then by Proposition 12 we see that $MV \vdash \forall V\, CH^V \vee \forall V \neg CH^V$. This is clearly impossible. (2) Similar except use Theorem 43. ⊣

Finally, we show that if we are at the core, then the obvious translation function allows us to establish that every sentence in the language of set theory is legitimate.

THEOREM 17. *For each $\varphi \in \mathcal{L}_\in$, there is a $\psi \in \mathcal{L}_{MV}$ such that*

$$ZFC + \mathsf{V} = \mathsf{C} \vdash \varphi \leftrightarrow t(\psi).$$

We'll first establish the key lemma.
Let $s : \mathcal{L}_\in \to \mathcal{L}_{MV}$ be such that for $\psi \in \mathcal{L}_\in$ we have

$$s(\psi) = \psi^C;$$

i.e., ψ relativized to the formula defining the core in $\mathcal{L}_\in$.

LEMMA 45. *For all sentences $\varphi \in \mathcal{L}_\in$*,

$$ZFC + \mathsf{V} = \mathsf{C} \vdash t \circ s(\varphi) \leftrightarrow \varphi.$$

PROOF. Let $\mathcal{W}$ be such that $\mathcal{W} \models MV$ and $\mathcal{M} \in \mathcal{W}$. Then we note that

$$\begin{aligned} \mathcal{M} \models t \circ s(\varphi) &\Leftrightarrow \mathcal{W} \models s(\varphi) \\ &\Leftrightarrow \mathcal{W} \models \varphi^C \\ &\Leftrightarrow \mathcal{M} \models \varphi. \end{aligned}$$

For the first $\Leftrightarrow$ we rely on Theorem 26 and Theorem 9. ⊣

PROOF. (of Theorem 17) Let $\varphi \in \mathcal{L}_\in$. Let ψ be $s(\psi)$. Then by Lemma 45, we see that

$$ZFC + \mathsf{V} = \mathsf{C} \vdash t(\psi) \leftrightarrow \varphi$$

as required. ⊣

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press

REFERENCES

M. Balaguer, ***Platonism and Anti-Platonism in Mathematics***, Oxford University Press, New York, 1998.

G. Cantor, *Contributions to the Founding of the Theory of Transfinite Numbers*. LaSalle, IL: Open Court 1952. Translated by Philip Jourdain.

P. Corazza, *Forcing with non-wellfounded models*. ***The Australasian Journal of Logic***, vol. 5 (2007), pp. 20–58.

S. Feferman, *Why the programs for new axioms need to be questioned*, this Journal, vol. 4 (2000), pp. 401–413.

M. Foreman and A. Kanamori, ***Handbook of Set Theory***, Springer, Netherlands, 2009.

S. D. Friedman, ***Fine Structure and Class Forcing***. De Gruyter Series in Logic and Its Applications, Walter de Gruyter, Berlin, 2000.

G. Fuchs, J. D. Hamkins, and J. Reitz, *Set-theoretic geology*. ***Annals of Pure and Applied Logic***, vol. 166 (2015), no. 4, pp. 464–501.

K. Gödel, *What is cantor's continuum problem*?, ***Collected Papers, vol. II*** (S. Feferman, et al., editors), Oxford University Press, Oxford, 1990, pp. 176–187, 254–270.

J. D. Hamkins, *The set-theoretic multiverse*. ***Review of Symbolic Logic***, vol. 5 (2012), no. 3, pp. 416–449.

———, *Upward closure and amalgamation in the generic multiverse of a countable model of set theory*, 2015, arXiv e-prints.

D. Hilbert, *On the infinite*, ***From Frege to Gödel*** (J. van Heijenoort, editor), Harvard University Press, Cambridge, 1967, pp. 369–392.

T. Jech, ***Set Theory***, Springer, Heidelberg, 2003.

T. Jech, M. Magidor, W. Mitchell, and K. Prikry, *Precipitous ideals*. ***The Journal of Symbolic Logic***, vol. 45 (1980), no. 1, pp. 1–8.

A. Kanamori, ***The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings***, Springer, New York, 2003.

P. Koellner, *Large cardinals and determinacy*, ***The Stanford Encyclopedia of Philosophy*** (E. N. Zalta, editor), Springer, New York, 2014. Available at https://plato.stanford.edu/archives/spr2014/entries/large-cardinals-determinacy/.

K. Kunen, ***Set Theory: An Introduction to Independence Proofs***, Elsevier, Amsterdam, 2006.

———, ***Set Theory***, second ed., College Publications, London, 2011.

A. Levy and R. Solovay, *Measurable cardinals and the continuum hypothesis*. ***Israel Journal of Mathematics***, vol. 5 (1967), pp. 234–248.

P. Maddy, ***Defending the Axioms***, Oxford University Press, Oxford, 2011.

———, *Set-theoretic foundations*, ***Foundations of Mathematics*** (A. Caicedo, et al., editors), American Mathematical Society, Providence, 2017, pp. 289–322.

M. Magidor, *Some set theories are more equal*, 2019. Available at http://logic.harvard.edu/efi.php#multimedia.

D. A. Martin, *Completeness or incompleteness of basic mathematical concepts*, 2019. Available at http://www.math.ucla.edu/dam/booketc/efi.pdf.

———, *Hilbert's first problem*: *The continuum hypothesis*, ***Mathematical Developments from Hilbert's Problems*** (F. Browder, editor), Proceedings of Symposia in Pure Mathematics, vol. 28, American Mathematical Society, Providence, RI, 1976, pp. 81–92.

T. Meadows, *Two arguments against the generic multiverse*. ***Review of Symbolic Logic*** (2020), forthcoming.

G. Moore *Towards a history of cantor's continuum problem*, ***The History of Modern Mathematics, vol. I*** (D. Rowe and J. McCleary, editors), Academic Press, Cambridge, 1989, pp. 79–121.

Y. Moschovakis, ***Descriptive Set Theory***, second ed., American Mathematical Society, Providence, RI, 2009.

J. Reitz, *The ground axiom*. ***Journal of Symbolic Logic***, vol. 72 (2007), no. 4, pp. 1299–1317.

J. Schatz, ***Axiom Selection and Maximize: Forcing Axioms vs. V=Ultimate-L***, UCI Ph.D. dissertation, 2019.

S. Shelah, *Can you take solovay's inaccessible away*? ***Israel Journal of Mathematics***, vol. 48 (1984), no. 1, pp. 1–47.

R. M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*. ***Annals of Mathematics***, vol. 92 (1970), no. 1, pp. 1–56.

J. Steel, *Mathematics needs new axioms*, this Journal, vol. 6 (2000), pp. 422–433.

———, *Gödel's program*, ***Interpreting Gödel: Critical Essays*** (J. Kennedy, editor), Cambridge University Press, Cambridge, 2014, pp. 153–179.

———, *Ordinal definability in models of determinacy*, ***Ordinal Definability and Recursion Theory*** (B. L. A. Kechris and J. Steel, editors), Cambridge University Press, Cambridge, 2016, pp. 3–48.

S. Todorcevic, *The power set of omega*-1 *and the continuum problem*, 2019. Available at http://logic.harvard.edu/efi.php#multimedia.

T. Usuba, *The downward directed grounds hypothesis and very large cardinals*. ***Journal of Mathematical Logic***, vol. 17 (2017), p. 2.

N. Weaver, ***Forcing for Mathematicians***, World Scientific, Singapore, 2014.

H. Woodin, *In search of ultimate-l*, this Journal, vol. 23 (2017), pp. 1–109.

———, *The continuum hypothesis, the generic-multiverse of sets, and the omega conjecture*, ***Set Theory, Arithmetic, and the Foundations of Mathematics*** (J. Kennedy and R. Kossak, editors), Cambridge University Press, Cambridge, 2011, pp. 13–42.

DEPARTMENT OF LOGIC AND PHILOSOPHY OF SCIENCE
UNIVERSITY OF CALIFORNIA, IRVINE
IRVINE, CA, USA
E-mail: pjmaddy@uci.edu
E-mail: meadowst@uci.edu

https://doi.org/10.1017/bsl.2020.5 Published online by Cambridge University Press