Title

The Epistemology of the Infinite

Author

Ryan, Patrick James

Publication Date

2024

Peer reviewed|Thesis/dissertation

The Epistemology of the Infinite

by

Patrick James Ryan

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Philosophy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Paolo Mancosu, Co-chair
Associate Professor Shamik Dasgupta, Co-chair
Professor Theodore Slaman

Summer 2024

The Epistemology of the Infinite

Abstract

The Epistemology of the Infinite

by

Patrick James Ryan

Doctor of Philosophy in Philosophy

University of California, Berkeley

Professor Paolo Mancosu, Co-chair

Associate Professor Shamik Dasgupta, Co-chair

The great mathematician, physicist, and philosopher, Hermann Weyl, once called mathematics the "science of the infinite." This is a fitting title: contemporary mathematics—especially Cantorian set theory—provides us with marvelous ways of taming and clarifying the infinite. Nonetheless, I believe that the epistemic significance of mathematical infinity remains poorly understood. This dissertation investigates the role of the infinite in three diverse areas of study: number theory, cosmology, and probability theory. A discovery that emerges from my work is that the epistemic role of the infinite varies, often in surprising ways, across different domains of knowledge.

My first chapter examines the role of mathematical infinity in number theory. It is reasonable to think that theorems concerning finite patterns and structures in the natural numbers are particularly "simple" or "elementary." Indeed, such statements are comprehensible to a wide range of investigators, regardless of their mathematical training. One might then expect *proofs* of these theorems to be similarly comprehensible. However, many proofs, especially those that utilize only finitary methods, are exceedingly difficult to understand. Consequently, one finds that finitary theorems are often re-proved using infinitary techniques. My claim is that this is because infinitary proofs are often explanatory, while finitary proofs are not. This chapter analyzes why this is the case. Along the way, I investigate other questions of long-standing interest in the philosophy of mathematics, e.g., the role of purity/impurity ascriptions and nature of the content of a theorem. In particular, I diagnose the explanatory potential of the infinite by articulating a new construal of *content*. This new construal both saves intuitive epistemic ascriptions made in mathematical practice and explains the unexpected role of the infinite in providing explanatory proofs of finitary statements. Thus, in number theory, my claim is that the *infinite* often plays an explanatory role.

My second chapter turns to the role of the infinite in cosmology. It investigates a question much discussed by philosophers and physicists alike: is the spatial extent of the universe finite or infinite? Contemporary cosmological research has indicated that one of the essential determinants of the extent of the universe is the *topology* we ascribe to space. Topology is a global property, which may suggest that it is not testable through local observation. Nonetheless, some cosmologists have indicated that it may be empirically detected, thereby providing an answer to the question of spatial extent. I argue that, in fact, the epistemic status of the topology of space is extremely subtle and not well captured by any of the categories commonly employed by philosophers of science. In particular, I argue that topological properties are neither empirical nor *a priori* (even in suitably weakened senses). Furthermore, I claim that we should prefer topological properties that generate *finite* universe models (consistent with our best data) in order to avoid extremely thorny issues concerning the physics of an infinite universe. I argue for such a preference on the grounds of the simplicity and explanatory power of finite universe models. Thus, in cosmology, my claim is that the *finite* often plays an explanatory and simplifying role.

My third chapter investigates several paradoxes that arise in the foundations of infinitary probability theory: the Label Invariance Paradox, God's Lottery, and Bertrand's Paradox. I argue that these have been poorly understood because they do not expressly concern probability theory, but rather our intuitions about—and formal techniques for dealing with—infinite sets. The paradoxes in question are, in fact, symptoms of our complete reliance upon Cantorian cardinality and its associated criterion of sameness of "size." That is, two sets have the same cardinality if and only if the elements of the sets can be placed in 1-1 correspondence. When applied to infinite sets, this criterion produces counterintuitive verdicts. For instance, given a fair lottery on the natural numbers, we expect that the probability of drawing an even number is 1/2, and likewise for drawing an odd number. However, one can construct "relabellings" of the naturals such that the probability of drawing an even number remains 1/2, while the probability for drawing an odd number becomes 1/4. I argue that, ultimately, it is the coarseness of Cantorian cardinality that generates the probabilistic paradoxes. I then propose that finer-grained measures of infinite sets from mathematical logic and number theory can help to dissolve the paradoxes in question. Thus, in probability theory, we find that *particular kinds* of infinitary techniques effectively systematize our theory, while others lead to paradox.

But as in landlessness alone resides the highest truth, shoreless, indefinite as God—so, better it is to perish in that howling infinite, than to be ingloriously dashed upon the lee, even if that were safety!

–*Moby-Dick*, Chapter 23, The Lee Shore.

"Infinity!" Törless had often heard the word in mathematics lessons. It had never meant anything in particular to him. The term kept on recurring; somebody had once invented it, and since then it had become possible to calculate with it as surely as with anything real and solid. It was whatever it stood for in the calculation; and beyond that Törless had never sought to understand it. But now it flashed through him, with startling clarity, that there was something terribly disturbing about this word. It seemed to him like a concept that had been tamed and with which he himself had been daily going through his little circus tricks; and now all of a sudden it had broken loose. Something surpassing all comprehension, something wild and annihilating, that once had been put to sleep by some ingenious operation, had suddenly leapt awake and was there again in all its terrifying strength. There, in the sky, it was standing over him, alive and threatening and sneering. At last he shut his eyes, the sight of it was such anguish to him.

–*Die Verwirrungen des Zöglings Törleß* (*The Confusions of Young Törless*).

Contents

Acknowledgments

I would like to thank, first and foremost, my advisor, Paolo Mancosu. Paolo's intellectual influence should be apparent from the first words of the dissertation. From the start of my time at Berkeley, Paolo showed me new ways to think philosophically about mathematics, logic, and their histories. He also made clear to me the incredible amount of thought, scholarship, and effort that must go into worthwhile research. I have tried, though perhaps have not always succeeded, to do justice to his insights in what follows. Besides this, he has been a constant source of patience, support, and encouragement over the course of some truly difficult years. He has my deepest gratitude.

I must also thank Shamik Dasgupta for graciously joining this project *in media res*. Shamik often helped me to consider my work in a new light, which is an invaluable part of the philosophical process. I have also benefitted from his requests for further clarity in presentation and high-level argumentation. (This should be evident in the selections of the dissertation that have been published.) Finally, Shamik got me excited about Reichenbach and Poincaré, and this greatly enhanced the philosophical interest of my second chapter.

Besides my advisors, I have been fortunate to have excellent readers and interlocutors at Berkeley. John MacFarlane and Tim Clarke offered incisive remarks on my job market materials and parts of the dissertation, as well as general encouragement throughout the Ph.D. Much earlier in the program, R. Jay Wallace and Daniel Warren served as mentors and qualifying exam advisors, and each helped me to shape interests in philosophical topics outside of my specialization. Finally, it was at Berkeley that I developed a taste for ancient philosophy, Greek, and Latin in no small part due to courses offered by members of the Classics (now DAGRS) department. Thanks are therefore due to A.A. Long, Sara Magrin, G.R.F. Ferrari, Trevor Murphy, and, once more, Tim Clarke. Finally, I am grateful to Ted Slaman for serving as my external committee member and for carefully reading the dissertation.

I would also like to recognize the contributions of the administrative staff in the Philosophy department: Janet Groome, Bill Dellinger, Maura Vrydaghs, and Kathryn Dernham. Every member of the team assisted me in some way during my time at Berkeley. I am certain that they relieved me of many tasks of which I am not even aware! Bill deserves special thanks for helping me to navigate a complicated final year.

Aspects of this dissertation have been greatly improved due to the suggestions of a number of external correspondents and audiences. In particular, Julia Bursten, Tom Ryckman, and Jeremy Avigad offered generous feedback, despite not supervising my work in any official

capacity. I must also thank audiences at the University of Notre Dame, Université de Paris 1 Panthéon-Sorbonne (especially Jean-Pierre Luminet and Andrew Arana), Chapman University (especially Marco Panza), and the 2023 Spring Meeting for the Association for Symbolic Logic/ 2023 Pacific APA (especially Silvia de Toffoli and Neil Barton).

Aristotle tells us that friendship comes about "by sharing in conversation and thought." I'm not sure the directionality is quite right here, but, in any case, I would like to thank my friends who offered discussion and commentary on my work. In particular, thanks are due to Katie "Spanky" Coyne, Ravit Dotan, Madeleine Levac, Russ McIntosh, Dave Gottlieb, and Angela Yeo.

Most importantly, I would like to acknowledge and thank those who have offered me emotional support (and much else) over the course of the Ph.D. I simply would not have made it were it not for Roberto and Izabela (my adoptive parents) and for Jessica. They have made it possible for me to grow in ways beyond the intellectual. Similarly, I would not have made it (perhaps also in a much more real sense) were it not for the support and guidance of my grandmother. She has preserved me from negligence and cruelty over the course of my life. Finally, I must express my thanks and love to my family, Monika, Zoe, and Tule. Monika not only read every last word of the dissertation, but also endured with great patience my endless yammering and perseverating over its details. She, more than anyone else, myself included, understands my virtues, vices, and "moods." I look forward to the coming years with you, whatever they may bring.

1 General Introduction

This dissertation consists of three studies unified by both their methodology and subject matter. Let me begin with a description of the former, since the question of what "philosophical methodology" is and how it differs from that of other disciplines has long been a starting point of philosophical reflection. In what follows, I attempt to practice what might be called "historically and scientifically informed philosophy." By this I mean a type of philosophy that is attentive to three things: (i) the ideas and approaches of philosophers, mathematicians, and scientists of the past;[1] (ii) contemporary technical developments in the sciences;[2] (iii) how scientific practitioners employ these technical aspects. My methodological approach proceeds from the conviction that engagement with these three areas[3] is crucial to worthwhile philosophical research in general. Indeed, I think that this becomes particularly clear in the philosophical study of mathematics and the sciences.

If one only examines contemporary mathematical and scientific theories, viz., independent of their histories and foundational assumptions, it is easy to miss interesting conceptual difficulties that are no longer explicitly addressed. Contemporary theories have a much "smoother" look to them because choices were made in order to deal with a particular problem (or to set it aside entirely). Furthermore, it is often the case that philosophers were involved in this process and that their views directly impacted the work of practitioners. Understanding the dynamics of this process both elucidates new philosophical connections and serves as an effective antidote against the thought that canonical philosophical questions no longer have force.

On the other hand, in order to engage fruitfully with mathematics and the sciences, one must also have some level of technical proficiency. This allows one to exclude questions that are ill-formed, for certainly there are such. It also allows one to appreciate the significance of a particular result, postulate, etc. in the contemporary scientific edifice, and thus suggests promising lines of future research. I attempt to follow this general methodology, and thereby blend the histories of philosophy, mathematics, and the sciences with a close engagement with the technical features of contemporary research.

As for my subject matter, it is the philosophy of mathematics, broadly construed to include excursions into logic, physics, and probability. More precisely, I consider the role played by infinity in shaping our scientific theories and various epistemological questions that arise from the use of infinitary techniques. Infinity is, of course, a perennial source

[1] Aristotle, Kant, Poincaré, Reichenbach, and Weyl loom particularly large.

[2] Throughout the introduction, I use "science" very broadly to include not only the natural sciences but also mathematics and logic. Recent work in number theory, reverse mathematics, cosmology, and probability will feature below.

[3] Or, at least one of them, depending on the nature of one's work.

of both mathematical and philosophical reflection. I have tried to examine it from less typical perspectives. Many discussions of infinity focus on either canonical "paradoxes of the infinite,"[4] such as Zeno's and Russell's, or developments in Cantorian set theory (or both). This is not unjustified: paradoxes serve as a gadfly to philosophy[5] and Cantor's theory of ordinals and cardinals is perhaps the most significant mathematical advance in our understanding of the infinite. Nonetheless, the study of infinity is the domain of neither clever manipulators of paradox nor set theorists alone. Infinity permeates the mathematical and scientific landscape, and thus much of human activity and thought.

There is, however, something puzzling about this state of affairs: it would seem that a famously abstract, difficult, and even indeterminate concept is used to formulate theories about more concrete phenomena. Why is this? What accounts for this "[...] theoretical shaping of the world that presses so far beyond the given" ([Weyl, 2012], 28)? Whence the seeming "mysterious effectiveness of the infinite"?[6] These questions serve as motivation for the following investigations.

It is now worth pausing to ask what is meant by "the infinite." Indeed, one cannot give an univocal definition of the concept, for this is part of its interest, but we can at least provide some disambiguation.[7] One helpful distinction is that between a non-mathematical and mathematical sense of "infinite." In its non-mathematical usage, "infinite" means "unlimited," "endless," "immeasurably great in some respect." In this sense, we might call God or various of God's attributes "infinite." One will also see this non-mathematical sense of the infinite applied to space, time, and the universe, but oftentimes this is superseded by further mathematical descriptions.[8] In its mathematical usage, "infinite" means "having a determinate measure that is not finite." For instance, according to Cantor's theory of infinite cardinals, the "size" of the set of natural numbers is $|\mathbb{N}| = \aleph_0$, the smallest infinite cardinal.

A second distinction, formulated in Aristotle's early and epoch-making investigations into infinity, is that between the "potential" and the "actual" infinite.[9] The potential infinite arises from the indefinite iteration of a process, e.g., the indefinite division of a finite spatial magnitude. As Aristotle says, "In general, the infinite is in virtue of one thing's constantly being taken after another; each thing taken is finite, but it is always one followed by another" (*Physics*, III.6, 206a27-29). On the other hand, the "actual" infinite would result if such a process could be completed "in a flash," resulting in, say, infinitely many actual divisions of a spatial magnitude. Aristotle himself rejects the notion of the actual infinite.[10] Note that Aristotle is working under the auspices of our mathematical notion insofar as he considers the application of infinity to quantifiable magnitudes. Indeed, in contemporary mathematics (and philosophy), his distinction remains very much alive.[11]

In the history of infinity, even these few distinctions intersect one another in interesting

[4] A classic introduction to the infinite, [Moore, 2019], frames most of his discussion in terms of paradoxes.

[5] Indeed, I have been unable to resist it. See Chapter 3.

[6] To paraphrase the title of E. Wigner's famous paper, [Wigner, 1960], concerning the role of mathematics in the natural sciences.

[7] For further discussion, see [Moore, 2019], [Easwaran et al., 2023], and the references therein.

[8] Consider my discussion in Chapter 2.

[9] See *Physics*, III.4-8. I quote from [Hussey, 1983].

[10] As far as I know, there is only one commentator who disagrees with this general statement: see [Rosen, 2021].

[11] For discussion, see, e.g., [Linnebo and Shapiro, 2018].

and complicated ways. I cannot do justice to that story here. Nonetheless, we are now able to narrow our focus somewhat. In what follows, I will be concerned with a strictly mathematical sense of infinity. However, even internal to mathematics, we witness a proliferation of techniques for understanding and measuring the infinite. Indeed, in the two millennia following Aristotle, most philosophers, scientists, and mathematicians considered the potential infinite to be the only notion suitable for scientific purposes.[12] A significant change occurred in the late 19th century as a result of Cantor's groundbreaking mathematical work. His development of transfinite set theory showed that sound mathematical sense could be made of the "actual" infinite.[13] Of course, Cantor's theory was not immediately accepted; far from it. The famous foundational dispute between the "schools" of logicism, intuitionism, formalism, and predicativism soon followed, and this was, at its heart, a dispute over the nature of infinity in mathematics.[14]

As the dust from this dispute began to settle, in part due to the axiomatization of set theory provided by Zermelo and Fraenkel (Zermelo-Fraenkel set theory with Choice; ZFC), concerns about infinity were allayed (or, perhaps, swept under the carpet). However, many questions about the infinite remain. As we shall see, these questions are of great import for the philosophy of mathematics and quickly become intertwined with other philosophical concerns. Given these complicated dependencies, each chapter is not necessarily organized with infinitary concerns foregrounded. However, the motivation and common thread of all chapters resides in the infinite. Here is a small sample of the questions that arise below: What do we gain by proving results about the finite using infinitary resources? What is lost? Can such infinitary proofs be explanatory? If so, why? What is the role of infinity in cosmology? Should it be avoided? What is its connection to other mathematical properties of spacetime? What is the role of infinity in probabilistic reasoning? Should we prefer one criterion for "measuring" infinite quantities over another depending upon our scientific context?

What conclusions can be drawn from such an investigation? Following the methodology sketched above, I did not try to force my studies onto a Procrustean bed but rather allowed myself to be led by the phenomena. The different contexts and ways of thinking about the infinite found below are complex and variegated, and my conclusions reflect this complexity. Nonetheless, I hope to provide my reader with new ways of thinking about the epistemic import of the infinite in mathematized theories. If nothing else, my aim was to map out discrete regions in which infinitary considerations play a central role, viz., number theory, cosmology, and probability theory, and to see whether interesting continuities or discontinuities arise. Ultimately, we find a good deal of discontinuity. In my number-theoretic case, we find that the infinite plays an explanatory role, mediated by (suitably understood) varieties of simplicity and unification. On the other hand, in relativistic cosmology, the infinite produces many difficulties, and thus in the cases I analyze the finite plays an explanatory and simplifying role. Finally, in probability theory, particular kinds of infinitary techniques effectively systematize our intuitions, while others lead to paradox. Let me now conclude

[12]For a philosophical and historical discussion of some important episodes concerning the infinite in the 17th century, see [Mancosu, 1996].

[13]For a discussion of the historical and mathematical developments, see [Ferreirós, 1999].

[14]For reflections on this by two intellectual giants of the time, see [Weyl, 2012] and [Gentzen, 1969]. See also the essays in [van Heijenoort, 1967] and [Mancosu, 1998].

with a chapter-by-chapter summary of these findings.

Chapter 2: I argue for an association between impurity and explanatory power in contemporary mathematics. Roughly, a proof of a theorem is said to be "impure" if it utilizes concepts intuitively "foreign" or "extraneous" to those present in the theorem itself, e.g., the use of infinitary resources in the proof of a finitary theorem. This proposal cuts against a long tradition of philosophers and mathematicians alike (e.g., Aristotle, Bolzano, Hartry Field) who have argued that *purity* and explanatory power go hand-in-hand. Indeed, an association between impurity and explanatory power is rather surprising. If an explanation is to provide the "reason-why" a theorem is true, it is plausible to think that an appeal to different conceptual resources could subtly modify the question we are trying to answer or could cause us to lose essential data for answering the question. However, I show that, provided various conditions hold, these worries are not apropros. My argument proceeds by analyzing a central and deep result of additive number theory, Szemerédi's theorem, and various of its proofs (Chapter 2.3). In particular, I show that it is only by appealing to impure resources, especially those utilizing infinitary techniques, that we obtain an explanatory proof of Szemerédi's theorem. I begin to account for the explanatory power of impurity by articulating a new construal of the content of a mathematical statement, which I call structural content. I argue that the availability of shared structural content both saves intuitive epistemic distinctions made in mathematical practice and simultaneously explicates the intervention of surprising and explanatorily rich resources (Chapter 2.4).

I then proceed to substantiate and explicate the claim that impurity plays an explanatory role. That is, though the above demonstrates how it is *possible* that impurity produces an explanatory proof (via the presence of shared structural content), it does not show, in detail, how this *actually* occurs. My primary contention is that impurity helps to generate varieties of simplicity and unification, and these render the proof explanatory (Chapter 2.5 and 2.7). In particular, following a famous distinction made by Aristotle, I claim that impurity makes the "reason-why" (*to dioti*) a theorem is true particularly explicit and helps us recognize the "reason-why" qua "reason-why." This occurs because infinitary resources help to clear away complicated finitary dependencies that obfuscate the reason-why and its role in the proof of the theorem. This "information management" via infinity in turn helps to produce highly linear and comprehensible proof structures.

Finally, I provide a reassessment of what has come to be known as *Gödel's Doctrine* (Chapter 2.6). In the formulation and proof of his famous Incompleteness Theorems, Gödel had shown the existence of finitary statements that *required* infinitary resources to prove them, viz., the Gödel sentence, G. This led him to postulate that "the unlimited transfinite iteration of the powerset operation is *necessary* to account for finitary mathematics" (Gödel's Doctrine). Indeed, this claim garnered further support over the course of the 20th century because of the production of various other finitary "independence" results (e.g., Paris-Harrington theorem, Goodstein's theorem). Nonetheless, proof-theoretic work by Solomon Feferman and others showed that these finitary results could be proved using much weaker systems, e.g., predicatively justifiable systems, thereby challenging Gödel's Doctrine. I argue that, though these technical results are unimpeachable, their philosophical significance is overstated. In

particular, though "strong" infinitary systems may not be necessary for *proving* these finitary independence results, I argue that such systems are necessary for the production of maximally comprehensible and explanatory proofs. I believe that this then accounts for the fact that mathematicians continue to use "strong" infinitary resources even if these are not strictly necessary. Speaking more broadly, the exclusive focus on axiomatic strength and proof-theoretic necessity has neglected interesting lines of research that are helpful for producing an epistemology of mathematics adequate to how mathematics is actually practiced.

Chapter 3: I begin by considering the fact that modern cosmology faces a radical underdetermination problem. In particular, there are very many cosmological models (models of General Relativity) compatible with our best (and best possible) observational data. At first blush, this may be quite unsurprising given that cosmology deals with physics at extremely large scales. A rather more surprising fact is that, even under strong hypotheses about the global structure of space (the Cosmological Principle), this underdetermination persists; in particular, spatial topology remains observationally underdetermined (Chapter 3.3 and 3.4). Is there any way to break this topological underdetermination? I survey recent work in observational cosmology that has aimed to provide definitive answers on this front and conclude that the prospects for empirically determining the topology of space are not promising (Chapter 3.5). Nonetheless, I argue that we can muster reasons to prefer various topological properties over others. In particular, I favor the adoption of multiply connected universe models on grounds of (i) simplicity, (ii) Machian considerations, and (iii) explanatory power (Chapter 3.6, 3.7, 3.8). Crucially, we are able to appeal to such grounds because multiply connected topologies open up the possibility of finite universe models (consistent with our best data), which in turn avoid thorny issues concerning the postulation of an actually infinite universe.

In light of the above underdetermination, I then consider the puzzling epistemic status of global properties of spacetime, e.g., the topology of space. Indeed, if spatial topology is always underdetermined by observational data, and some ascription of spatial topology is required for the cogency of modern cosmology, what are we to say of it? A natural suggestion is that the topology of space is conventional. After examining work by Poincaré and Reichenbach, I conclude that, though there is a sense in which calling spatial topology "conventional" is correct, conventionality of any stripe does not fully capture its epistemic status (Chapter 3.9). This is because of the foundational role of spatial topology in our cosmological theorizing: it would appear that the topology of space makes possible the application of fundamental physical concepts and subsidiary physical laws. Thus, I turn to Michael Friedman's work on the relativized or constitutive *a priori* and argue that spatial topology is a component of the constitutive *a priori* apparatus of General Relativity (Chapter 3.10). I then discuss how this examination of the epistemic status of spatial topology brings to light various unclarities in philosophical accounts of conventionalism and the constitutive *a priori*. Finally, I conclude by claiming some fusion of these views is required to account for our case. Thus, the epistemic status of spatial topology is extremely subtle and escapes classification by categories commonly employed by philosophers of science.

Chapter 4: I consider several famous paradoxes that arise in infinitary probability theory: the Label Invariance paradox, God's Lottery, and Bertrand's paradox (Chapter 4.3, 4.4). I argue that these paradoxes have been poorly understood by philosophers, mathematicians, and physicists alike because they are not expressly about probability theory. Rather, they concern our intuitions about—and techniques for measuring—infinite sets. I claim that the paradoxes arise for two related reasons. First, they are produced by a conflict between our intuitions concerning finite sets and attempts to generalize these intuitions to the infinite case. When such generalizations occur, only particular intuitions can be preserved, while others must be jettisoned. Second, the dominant generalization thus far has been the notion of Cantorian cardinality and its associated criterion of sameness of "size." That is, two sets have the same cardinality if and only if the elements of the sets can be placed in 1-1 correspondence (**Cantor's Principle**; **CP**). I argue that it is, ultimately, the coarseness of **CP** that produces the probabilistic paradoxes and that finer-grained measures of infinite sets, preserving more delicate part-whole and frequency intuitions, will help to dissolve and diagnose the paradoxes in question. Thus, I provide a unified framework in which to think about these seemingly distinction "probabilistic" paradoxes and suggest how they might be resolved.

Finally, I use this examination to consider our current theories of "infinite counting," i.e., the techniques we possess for "measuring" infinite sets. In particular, I consider the recently developed theory of numerosities, which provides a way to formalize a **Part-Whole** (**PW**) intuition when measuring the "size" of an infinite set (Chapter 4.2). That is, numerosities validate the following intuition for infinite sets: if A is a proper subset of B, then the size of A should be strictly less than the size of B (**PW**). A general theme that emerges from this investigation is that there is an inextricable indeterminacy to our theories of infinite counting. I offer some reflections on this and ultimately conclude that, in particular contexts, this indeterminacy offers a flexibility that allows us to preserve properties suited to the context in question.

2 Szemerédi's Theorem: An Exploration of Impurity, Content, Infinity, and Mathematical Explanation

2.1 Introduction and Argument Outline

This study discusses some aspects of mathematical explanation through a detailed analysis of an important result: Szemerédi's theorem and various of its proofs.[1] By *mathematical explanation* I mean explanation internal to mathematics in which mathematical facts are used to explain other mathematical facts; moreover, I here analyze a "local" conception of mathematical explanation insofar as explanatory power is construed as a property of proofs.[2] The details of the proofs I consider become quite advanced, so I should like to provide a less technical introduction for a general philosophical audience. I do, however, encourage my readers to work through as much of the technical material as possible; the initial statements of the relevant theorems and my summary sections would be helpful places to start.[3]

The main thesis of this chapter is that impurity is not just a central aspect of contemporary mathematics, for that is an obvious descriptive claim, but also a central aspect of *mathematical explanation*.[4] That is, I argue impure methods play a distinctive role in generating explanatory proofs. In defending this claim, I hope to present a synthesis of ideas from philosophy, mathematical practice, and mathematical logic/foundations in order to contribute to a more sophisticated and nuanced philosophy of mathematics. Such a wedding of disciplines is required to do justice to the complexity of my subject matter, viz., mathematics and its philosophical implications. Indeed, it seems unhealthy and intellectually restrictive for there to be distinct "traditions" in the philosophy of mathematics, focusing on either foundations or practice, respectively. Those interested in mathematical practice (what Kitcher, now long ago, called the "maverick tradition") should utilize the

[0]Content from Sections 1-4 of this chapter first appeared in *The Review of Symbolic Logic*. , the author, 2021. Please cite the published version: "Szemerédi's Theorem: An Exploration of Impurity, Explanation, and Content." *The Review of Symbolic Logic*. 16(3): 700-739 (2023). https://www.doi.org/10.1017/S1755020321000538.

[1]See Section 2.2 for criteria of selection. In particular, I consider the combinatorial and ergodic proofs of Szemerédi's theorem.

[2]See the *Stanford Encyclopedia of Philosophy* article by Mancosu for a discussion of both external and internal mathematical explanations as well as the local versus global conception of explanation ([Mancosu, 2018]).

[3]Section 2.3 (Szemerédi's theorem) up to the beginning of Section 2.3.1; Section 2.3.2 (Outline of Proof) and Theorem 2.3.12 (Furstenberg Multiple Recurrence); Section 2.3.3 (Preliminary Mathematical and Philosophical Remarks); Theorem 2.3.37 (Furstenberg Structure Theorem); Section 2.3.4 (Summary). It would also be worthwhile for readers to familiarize themselves with the formal systems defined in Appendix C.

[4]See immediately below for explications of "impurity" and "mathematical explanation."

logical and foundational results at their disposal, and, conversely, the mathematical logician should be attentive to the interesting epistemic features of mathematical practice. Finally, neither group should focus solely on the ontological and epistemological problems raised in Benacerraf's famous set of articles [Benacerraf, 1965] and [Benacerraf, 1973], which have for over fifty years so dominated the landscape of philosophy of mathematics. Indeed, the hope is that close attention paid to both mathematical practice and foundations will shed new light on the (genuinely important) concerns raised in these classic discussions. Similar points have been stressed by both Feferman, at least implicitly, and very explicitly by Mancosu.[5]

Before precisifying my main thesis, a few remarks about the method and scope of this study are in order. First, it is not my aim to enter into a discussion of even a small fraction of the literature on explanation. Instead, I concentrate on a positive proposal: the possibility and utility of explanations via impure techniques. Second, those desirous of a "rational reconstruction" of mathematical explanation will not find it here. I am not sanguine about the prospects for any "theory" of explanation. Explanation, even when restricted to the mathematical context, is an expansive concept, and it is far from clear that there is a singular phenomenon up for analysis.[6] As such, I find it incredibly artificial to lay down *a priori* desiderata, constituting some purported explication of explanation. This runs[7] the risk of becoming quite divorced from the actual mathematics. Rather, it seems much better to proceed from the bottom up by comparing results and techniques that may be classified in some relatively unobjectionable (though perhaps milquetoast) way as explanatory. Once we have such a starting point, we can further examine how these results were proven and hopefully find interesting confluences of conceptual resources that will help to enrich our understanding of what explanation may look like.

My starting point is an ancient one. In *Posterior Analytics*, A.13, Aristotle distinguishes between demonstrations "of the fact" (*ὅτι*; lit."the that") and demonstrations "of the reasoned fact" (*διότι*; lit."the why"). That is, in mathematics[8] we have proofs that show *that* a theorem is true and proofs that show *why* a theorem is true. This is, then, more or less the distinction between non-explanatory, though perfectly cogent, proofs and explanatory proofs. Such a distinction[9] helps to make good sense of mathematics as it is actually practiced, especially the fact that theorems, from the most elementary to the most complex, are

[5]The methodological points made here are, by now, old hat, but they are still important to state at the outset. For a very helpful and detailed analysis of the various traditions in 20th century philosophy of mathematics, see the Introduction of [Mancosu, 2008b].

[6]Some recent papers concerned primarily with Indispensability Arguments make the point that the practice of providing explanations is quite internally disunified and that we must be content with "explanatory mini-projects." See [Baker, 2016].

[7]And has run; see the unificationist approaches of Friedman and Kitcher below.

[8]Aristotle considers not only mathematics but axiomatized sciences in general. I say "axiomatized" because, even though Aristotle employs examples from Greek sciences that were not axiomatized, e.g., astronomy, the theory of explanation provided in the *Posterior Analytics* proceeds from indemonstrable axioms and scientific first principles, or, *ἀρχαί*. Indeed, explaining the inconsistency between this theory and Aristotle's own method has long been a classic problem in Aristotle scholarship. See, for instance, [Barnes, 1969].

[9]I employ the distinction independent of Aristotle's many conditions on when a syllogism counts as a demonstration, e.g., the premises of such a syllogism are immediate, better known than the conclusion, etc. See *An. Post.*, A.2, 71b20 for the statement of these conditions.

often proved multiple times and in various ways.[10] Explanatory proofs are supposed to generate some understanding[11] of the result being proved, while non-explanatory proofs merely give one warrant for the truth of the theorem; explanatory proofs provide "the why," while non-explanatory proofs merely provide "the that." In this paper, the relevant why-question that an explanatory proof should answer will be: Why does a particular pattern occur in sufficiently dense sets of positive integers? The answer turns out to involve a high-level mathematical fact that allows us to exhaustively characterize each set. However, this fact only becomes evident when we appeal to impure techniques. More precisely, *how* this fact materializes and interacts with other elements of proof becomes perspicuous only in the impure setting. Thus, in order to classify a proof as explanatory, it is not quite sufficient to answer the why-question latent in some theorem: one must also show how the answer to the "why" manifests itself.

Let me now introduce what is meant by purity (resp., impurity) in mathematical practice.[12] Mathematicians commonly describe a proof of a theorem as "pure" if it uses only what is "intrinsic" or "close" to the theorem. On the other hand, a proof is impure if it draws on what is "extrinsic," "distant," or "foreign" to the theorem.[13] Much recent work has been done to tease out what, precisely, mathematicians might mean when making purity ascriptions, in large part because there is a great deal of evidence that purity is a highly valued epistemic virtue in mathematical practice. This amounts to making more precise the distance measure implicit in statements about purity. One construal of this measure is that of "elemental closeness:" a proof draws only on what is simpler or more elementary[14] than the theorem. Another is that of "topical closeness:" a proof draws only on what belongs to the content of the theorem or what the theorem is about. Each metric then induces a purity constraint in a straightforward way.[15]

There are intuitive reasons to believe that any such purity constraint yields epistemic dividends, thus accounting for the importance of purity in the history of mathematics.[16] For instance, an elementally pure proof might, "make the most efficient use of the information at the disposal of a given investigator," while a topically pure proof might give one better warrant to believe that the intended statement has been proved, rather than some different, albeit closely related, one ([Arana, 2017], 208). However, surprisingly little has been said about the relationship between explanation and purity in the contemporary literature.[17]

[10] Some examples that immediately spring to mind: the Pythagorean theorem, the Riemann-Roch theorem, the Prime Number theorem, and, of course, the theorem considered in this paper.

[11] By this, I mean something like the conditions of explanation (whatever these may be) logically precede those of understanding.

[12] For a more detailed treatment, see, for instance: [Detlefsen, 2008], [Detlefsen and Arana, 2011], [Arana and Mancosu, 2012], [Arana, 2017], [Arana, 2019].

[13] A simple and intuitive example of impurity would be Descartes' wedding of geometry and algebra in analytic geometry.

[14] Of course, it is then incumbent on one to make precise what is meant by "elementary." *Mutatis mutandis* for topical purity and content.

[15] This suffices for now, but see the end of Section 2.4 where I indicate that the generation of a purity constraint is not so obvious.

[16] Although I consider the mathematical case here, the basic philosophical question is quite general: in order to get an explanation of some phenomenon should one appeal to foreign or endemic conceptual resources? Some appropriate combination thereof?

[17] This is not the case when one considers historical texts. Bolzano argued convincingly for the antithesis

Some recent work has taken this up, but only very casually and in passing.[18] Though it may be the case that considerations of purity are distinct from explanation,[19] it is quite natural to think that pure proofs would be paradigmatically explanatory (though, of course, I resist this natural thought as any sort of meaningful generalization). For instance, if one thinks of an explanation as an answer to a why-question, then an appeal to topical purity would appear to facilitate an explanatory proof by ensuring that one is answering the intended question. However, the case I consider in this paper is one in which *radically* impure techniques, i.e., techniques both elementally and topically impure, provide an explanatory proof.

It is worth mentioning remarks made in a more classical vein by Hartry Field in order to motivate the main thesis of this paper. As is well known, Field is a staunch nominalist and has provided the most systematic attempt to eliminate mathematical entities from scientific explanations.[20] Interestingly, he believes that one should want to be an eliminativist on grounds independent of anti-platonism:

> For *even on the assumption that mathematical entities exist,* there is a *prima facie* oddity in thinking that they enter crucially into explanations of what is going on in the non-platonic realm of matter. It seems to me that the most satisfying explanations are usually 'intrinsic' ones that don't invoke entities that are causally irrelevant to what is being explained. 'Extrinsic' explanations are acceptable [...] but it is natural to think that for any good extrinsic explanation there is an intrinsic explanation that underlies it. [...] I regard the acceptance of an extrinsic explanation as ultimate as at least somewhat odd ([Field, 1989], 18-19).

I hasten to note that Field is considering a rather different context: the use of mathematics in explanations of physical phenomena. In that case, the purported oddity of extrinsic explanations is plausible given that most (platonist and anti-platonist alike) would agree that mathematical entities do not (or could not) participate in the spatio-temporal causal

---

of my claim, i.e., that only by appealing to pure techniques can one get an explanation. See his proof of the Intermediate Value Theorem performed without appeal to "geometric" considerations in "Rein analytischer Beweis..." Bolzano's claim descends in large part from a consequence of Aristotle's theory of demonstration, namely that "[...] it is not possible to prove something by crossing from another genus, e.g., something geometrical by means of arithmetic" (*οὐκ ἄρα ἔστιν ἐξ ἄλλου γένους μεταβάντα δεῖξαι, οἷον τὸ γεωμετρικὸν ἀριθμητικῇ*; translation my own; *An. Post.* A.7, 75a38). I remark below in Section 2.4 on the relationship between my views and Aristotle's. In any case, returning to Bolzano, the great success of the rigorous development of analysis in the 19th century lends considerable weight to the association of purity and explanation. For more concerning Bolzano and mathematical explanation see, for instance, [Kitcher, 1975], [Mancosu, 1999], and [Betti, 2010]. Interestingly, when we pass to mathematics in the 20th century, there is much less emphasis on purity as providing explanations, even though it is still an important epistemic virtue. There seems to be an inextricably historical dimension to the relationship between purity and explanation. Perhaps once the "local foundations" for a subdomain of mathematics are given, i.e., the basic concepts are made sufficiently precise, inconsistencies are removed, etc., purity becomes much less important? But perhaps in providing the "local foundations" purity is essential?

[18] See, for instance, [Lange, 2017].

[19] Namely, purity can stand as a genuine epistemic virtue without appeal to explanation. See a brief discussion in [Lange, 2017], pp. 292-3. Here Lange makes clear that he believes purity and explanation are quite distinct.

[20] See [Field, 1980].

nexus. Thus, his claim that "...the role of mathematical entities, in our explanations of the physical world, is very different from the role of physical entities in the same explanations" is a potentially reasonable one (19). Nonetheless, one might well wonder if there is something odd about extrinsic explanations *in general*, i.e., independent of a causal construal of explanation. Should we think that there is always an intrinsic explanation in the offing, i.e., an explanation appealing only to concepts endemic or "close" to the *explanandum*? In particular, is it odd to think that we might have mathematical explanations of mathematical facts that make essential appeal to impure considerations? Should we think that any such impure explanation has a "better" pure explanation underlying it? Field does not offer any arguments for this general preference, and, obviously, I wish to resist such a move. Nonetheless, it is worth thinking about whether there is an intuitive strangeness to extrinsic (impure) explanations, especially in light of the advantages of pure proofs alluded to above. Part of my task in this chapter is to provide criteria for impure explanations that might allay such concerns[21] and to show that sometimes one must appeal to impure techniques in order to generate an explanatory proof.[22] I also indicate this connection because I tend to think considerations of impurity/purity have much more affinity with a suitably general notion of extrinsic/intrinsic explanation than has been acknowledged.

Another important strand of this study is an exploration of the relationship between infinitary and finitary mathematics[23] and the consequences of such a relationship for an account of mathematical explanation. This relates quite naturally to the forgoing remarks about impurity and explanation: one way of providing an (elementally) impure proof of a finitary theorem is to prove it using infinitary techniques. Furthermore, the fact that infinitary concepts have direct relevance for finitary results is a metamathematically and philosophically fascinating phenomenon that warrants further investigation.[24] Looking ahead: Szemerédi's theorem is a strictly finitary, combinatorial result that involves the additive structure of subsets of the natural numbers. Strikingly, as we shall see, Szemerédi's theorem is equivalent to an infinitary result, and one of its impure proofs (ergodic) makes essential use of a transfinite construction. One moral that I draw from all this is that, somewhat astonishingly, particular results involving the finite and infinite are explained by the same fact (the ubiquitous "dichotomy between structure and randomness"). Another is that, though we may not *require* particular (oftentimes infinitary) mathematical resources to prove a theorem, fixating merely on what one can get away with proof-theoretically obscures the explanatory role of such resources. I believe that one should consider these infinitary and impure resources in some sense necessary if they are the best means of providing an explanatory proof of a given theorem.[25]

A crucial question at this point is the following: how precisely does impurity lead to explanation? I claim that the answer is two-fold. Impurity leads to:

1. greater simplicity in proofs, where simplicity is construed as "conceptual speed-up" (Section 2.5);

[21] See, in particular, Section 2.4 on mathematical content.

[22] Viz., in order to get an explanation of Szemerédi's theorem, one must turn to the impure (ergodic) proof instead of the pure (combinatorial) one.

[23] See Section 2.2 for an explication of the finitary-infinitary distinction.

[24] See [Avigad, 2009] for a related metamathematical investigation.

[25] See my discussion in Section 2.6.

2. unification (of various sorts to be made precise) (Section 2.7).

It has long been claimed that impure techniques are "simpler" than pure ones; recent work by Arana[26] has shown that when one interprets "simplicity" as the proof-theoretic measure of proof length the association is equivocal. However, I demonstrate that a sort of simplicity is achieved by the impure proof of Szemerédi's theorem insofar as the global structure of the proof is made conceptually clearer.[27] I then consider how one should ontologically interpret the mathematical entities essential to an explanatory proof and suggest that they might support a restricted form of platonism via an indispensability argument.[28]

Concerning unification: there is a distinguished tradition in the philosophy of science literature wherein explanation is understood as theoretical unification. The *locus classicus* is Michael Friedman's [Friedman, 1974], although similar views can be found both in Kant and "unofficially" in the Hempelian deductive-nomological model. Friedman's central insight is that any successful account of explanation must be both "objective" and describe how explanation relates to understanding.[29] He believes that explanation as unification meets both criteria because unification, i.e., the subsumption of disparate phenomena under more general regularities or laws, reduces the number of facts an investigator must take as brute, thereby generating greater understanding of the world:

> [...]this is the essence of scientific explanation—science increases our understanding of the world by reducing the total number of independent phenomena that we have to accept as ultimate or given. A world with fewer independent phenomena is, others things equal, more comprehensible than one with more ([Friedman, 1974], 15).

Friedman's formal model for quantitatively measuring such a reduction was shown by Philip Kitcher to suffer from serious difficulties.[30] Nonetheless, his central insight has been incredibly influential. Kitcher himself agrees with Friedman's basic idea that explanation should be understood as theoretical unification and argues at length for this thesis in a number of works.[31] Kitcher believes that one virtue of such an account is that it can be applied uniformly to the sciences and mathematics unlike, for instance, causal theories of explanation which are inapplicable to purely mathematical contexts. Unfortunately, his model, despite its greater philosophical sophistication, has been shown to generate verdicts of explanatoriness contrary to mathematical practice.[32] One of the lessons to take away from this literature is that any "rational reconstruction" of explanation, especially via the construction

[26]See [Arana, 2017].

[27]It is important to note that one usually has to appeal to "soft" measures to make sense of epistemic virtues in mathematics. This will come up later when I consider unificatory models of explanation: it seems that these cannot capture important aspects of mathematical practice when they merely quantify syntactic criteria.

[28]See Section 2.6 below.

[29]In general, I agree that we should have both these criteria. However, Friedman's conception of objectivity is somewhat myopic and should be expanded.

[30]See [Kitcher, 1976].

[31]See [Kitcher, 1981], [Kitcher, 1984], [Kitcher, 1989].

[32]See the case study from real algebraic geometry in [Hafner and Mancosu, 2008].

of a quasi-formal model in which one is counting syntactic features of theories, is unlikely to succeed.[33]

Still, the idea that unification has *something* to do with explanation is intuitive and quite embedded in scientific and mathematical practice. For instance, even in the case study considered in [Hafner and Mancosu, 2008], mathematicians insist upon a sort of unificationist approach. Clearly then, unification (much less explanation!) must be said in many ways. In [Morrison, 2000] Margaret Morrison takes up the task of determining what is meant by "unification" in the natural sciences and shows, quite convincingly, that unification and explanation often come apart. I utilize some distinctions found in her excellent discussion and show that the same conclusion cannot be drawn in the purely mathematical case. Indeed, the impure (specifically, ergodic) proof of Szemerédi's theorem can be understood as unificatory and yet reveals the "reason why" the result holds. There is, then, a significant disanalogy between explanatory mechanisms in the sciences and mathematics. Thus, I argue that Kitcher is partially correct: we can often understand unification as an aspect of explanation in mathematics, but, it would seem, this is less often the case in the natural sciences.

This conclusion must, however, be defended from the mathematical analogue of Morrison's critique of unification in the sciences. One of her main points is that the mathematical apparatus of a scientific theory is often responsible for the unification of the theory, but this fails to actually explain why a particular event/phenomenon occurs. This is because, more often than not, one must investigate the causal behavior of a physical system to provide an explanation, and this information is not encoded by the mathematics of the theory. Initially, one might think that such an issue cannot arise internal to mathematics: there are no causal mechanisms at work. Nonetheless, it is quite reasonable to think that passing to a more abstract, infinitary, unificatory mathematical setting will result in a loss of data germane to the original context. This is certainly the case,[34] but, crucially, I show that the relevant explanatory data is preserved in a detour through the infinitary and impure setting. Another way to put this point is that the explanatory *content* of Szemerédi's theorem is not lost when one proves the theorem via impure techniques. This requires an analysis of mathematical content, i.e., the topic or subject matter of a particular theorem, which I undertake in Section 2.4. I consider a few proposals for what might count as the content of a theorem and attempt to excavate an "intermediate" notion of content that is both faithful to mathematical practice and allows one to account for surprising interventions of impure techniques. Finally, I argue that there must be some content shared by the theorem to be proved and the impure techniques utilized. If this were not the case, then the mathematical analogue of Morrison's thesis would be entirely applicable.

With this (rather long) schematic of my argument in place, let me now turn to the result itself and provide some criteria of selection that, hopefully, will bolster confidence in the philosophical conclusions I have outlined.

[33] I provide a brief survey of some of this material in Section 2.7.

[34] For example, in the case of Szemerédi's theorem, when one passes to the impure (infinitary and ergodic) setting, one loses the ability to compute effective bounds. This loss of information is a quite general phenomenon.

2.2 Criteria of Selection

Before providing my criteria of selection, let me define the distinction between *finitary* and *infinitary* mathematics. Boiled down to a slogan, finitary mathematics can be understood to mean the "theory of finite sets." More precisely, we have the following result:[35]

Theorem 2.2.1. *First-order Peano arithmetic* (PA) *is equivalent to* ZF^-*, i.e., Zermelo-Fraenkel set theory without the Axiom of Infinity and with the negation of the Axiom of Infinity. Here "equivalent" means that* PA *and* ZF^- *are mutually interpretable (and thus equiconsistent).*

The interpretation of PA in ZF^- is quite easy to see: interpret 0 as $\emptyset$ and the successor function as $x \mapsto x \cup \{x\}$. The converse is more surprising but can be done by coding finite sets with natural numbers. In particular, defining the membership relation $\in$ on the natural numbers as

$$n \in m \text{ iff the } n\text{th digit in the binary representation of } m \text{ is } 1 \tag{2.2.1}$$

interprets ZF^- (this interpretation is due to Ackermann[36]). On the other hand, one can understand infinitary mathematics to mean the "theory of infinite sets." This distinction can take on new contours in different contexts, e.g., finitary mathematics in Hilbert's Program is weaker than what I am calling finitary here.[37] When appearing in the mathematical wild, rather than in more precise logical contexts, finitary mathematics employs notions like: the cardinality of finite sets (of course), upper and lower bounds of such sets, and measures of bounded sets. Infinitary mathematics then employs notions like: sequences, measurable sets and functions, general measurability and integrability, convergence, and compactness. This battery of concepts should suffice to make clear when and why a result is called either finitary or infinitary below.[38]

I consider a case study in which infinitary and impure techniques are used to prove results with explicit combinatorial, finitary content: Szemerédi's theorem[39] on arithmetic progressions and a special case of this, van der Waerden's theorem. The former admits of an ergodic proof[40] ([Furstenberg, 1977], [Furstenberg et al., 1982]), the latter of a topo-

[35] See the recent paper by Kaye and Wong [Kaye and Wong, 2007] in which they investigate this result and locate some imprecision in the folklore. They indicate that the commonly cited relationship between PA and set theory is what I have given in Theorem 2.2.1; however, in order to show that the interpretations are inverses of one another (bi-interpretability) one must carefully axiomatize ZF. This is interesting, but does not affect the substance of the distinction made here.

[36] For instance, 55 serves as the code for the finite set $\{0, 1, 2, 4, 5\}$ since $2^5 + 2^4 + 2^2 + 2^1 + 2^0 = 55$.

[37] Once more, see Appendix C for a brief discussion of Hilbert's Program and how various formal systems may relate to it.

[38] For a nice reflection on finitary vs. infinitary mathematics in analysis, see the blog post by Terry Tao here.

[39] There is a recent, albeit brief, discussion of this theorem in the philosophy of mathematical practice literature. See [Arana, 2015]. The main purpose of that paper is to disambiguate various notions of "depth" in mathematical practice taking Szemerédi's theorem as example. Thus, though there is little overlap in content, I take Arana's paper as a nice companion piece that buttresses my selection of the theorem as worthy of philosophical attention.

[40] As well as a Fourier analytic and hypergraph proof.

logical proof ([Furstenberg and Weiss, 1978]), and both have explicit combinatorial proofs[41] ([Szemerédi, 1975], [van der Waerden, 1928], respectively).

Let me now provide some criteria of selection. First, I consider results that do not require[42] infinitary and impure techniques to prove them. This sets them apart from, e.g., the Paris-Harrington theorem, Goodstein's theorem, and Friedman's finitary version of Kruskal's theorem, which I may consider at another time. In attempting to elucidate the explanatory dividends of infinitary and impure methods in mathematics, it is crucial that I examine results that admit of multiple varieties of proof. This allows for a fruitful comparison of techniques, a comparison that (I hope) inclines one think that infinitary and impure methods, though not "required" *per se*, are very much an essential part of mathematical practice and explanation.

Second, these are results for which there is at least some methodological commentary by leading mathematicians.[43] This is of service, as I am of the mind that a successful epistemology of mathematics should be broadly consonant with mathematical practice (at least if we take ourselves to be explicating mathematics as an epistemic enterprise, rather than some philosophical abstraction). Thus, if my philosophical analysis has some support in the mathematical literature, I take this to be an advantage of my approach over those more revisionary.

Third, both results I consider have also been subjected to metamathematical analysis. Van der Waerden's theorem and its topological proof have been examined proof-theoretically by Girard ([Girard, 1987], 4A, 7E). Jeremy Avigad and Henry Towsner have attempted a similar analysis of the ergodic proof of Szemerédi's theorem, though given its greater complexity this analysis is in much earlier stages ([Avigad, 2009], [Towsner, 2008]). Both analyses are descendants of Kreisel's "unwinding" program, the aim of which is to extract the constructive content of *prima facie* non-constructive proofs. In fact, they are perhaps the only clearly successful *mathematical applications* of this program.[44] The metamathematical data available will allow us to make some of our observations more precise, e.g., as I discuss below, the stage of the ergodic proof of Szemerédi's theorem where crucial explanatory work is done turns out to involve axiomatically strong mathematics.[45] It would be very interesting to see

[41] In addition to the usually cited [van der Waerden, 1928], please see the excellent [van der Waerden, 1998] for a less terse discussion.

[42] Indeed, it is likely the case that once much proof-theoretic work has been done, *most* mathematics does not require the full strength of infinitary techniques. We should be careful not to confuse this descriptive claim with a normative one that mathematics *ought* to be this way. Though we can "get away with" proving many theorems from a relatively restricted mathematical universe, a great deal is lost when this is done. See [Feferman, 1964] and [Avigad, 2003] for eloquent articulations of similar views. An interesting question to which I would like to provide a partial answer is: what exactly is lost when we perform proof-theoretic ontological and epistemological reductions? E.g., I am somewhat skeptical of Avigad's strategy in [Avigad, 2009] to provide a purely constructive, finitary proof of the ergodic version of Szemerédi's theorem that is *as perspicuous* as the ergodic proof. Even if an explicit combinatorial version of Furstenberg's proof can be given, I posit that explanatory value will be lost.

[43] For instance, I utilize a number of papers by Terry Tao. I am quite indebted to Tao's excellent exposition of many of the following results.

[44] See [Feferman, 1996]. Another example worth mentioning is Dirichlet's Theorem on Arithmetic Progressions; see [Avigad, 2003], pp. 267-269. Finally, refer to [Kohlenbach, 2008] for a contemporary discussion of "unwinding" and "proof-mining" techniques in mathematics.

[45] See [Towsner, 2008] and [Avigad, 2009]. For a brief discussion of these and related formal systems, see

if this phenomenon is more widespread. In particular, one lesson of the reverse mathematical program is that surprisingly diverse mathematical results cluster around various levels of the reverse mathematical hierarchy. I am interested in whether we could find explanatory clusters in the hierarchy, i.e., whether the resources needed to provide explanatory proofs of diverse theorems accumulate at particular levels of axiomatic strength.

Finally, at least the initial presentation of these results should be relatively comprehensible to many readers, given that the original theorems are "number-theoretic" (about the behavior of certain sets of natural numbers or integers) and may be stated in finitary terms. Though the infinitary methods I survey require a good deal more mathematical background, I hope to make the important points about them tolerably clear for the working philosopher.[46] It should also be remarked that the apparently elementary nature of these theorems does not at all impugn their centrality or depth.[47] To my mind, many philosophical analyses of mathematics are unsatisfying or fail outright because they choose examples that have no obvious mathematical significance; this should then lead one to doubt the significance of the philosophical conclusions drawn. That is, if one is attempting to provide a philosophy *of mathematics*, the examples analyzed should be of acknowledged mathematical import.[48] Van der Waerden's and Szemerédi's theorems, on the other hand, have been loci of work in number theory and combinatorics for almost a century, and, as we shall see, gave rise to entirely new sub-fields of mathematics. Finally, the conceptual approaches used to prove these theorems have shown astonishing applicability in the solution of other, long-standing questions, e.g., whether the prime numbers contain arbitrarily long arithmetic progressions.[49] Indeed, the full significance of these theorems is far from understood.

2.3 Szemerédi's Theorem

Let us begin with a crucial definition:

Definition 2.3.1. (Arithmetic Progression) For $k, r \geq 1$, a *k-length arithmetic progression*, written as $a, a+r, \ldots, a+(k-1)r$, is a sequence of k integers such that each element of the progression differs from its predecessor by precisely r. We call a the *base point* and r the *radius* of the arithmetic progression.

Example 2.3.2. The set of numbers $\{1, \ldots, 9\}$ is an extremely trivial example of an arithmetic progression. We have $a = 1$, $r = 1$, and $k = 9$.

Appendices B and C.

[46] See my recommendations above (fn. 3) for the "theoretical minimum."

[47] Again, see [Arana, 2015].

[48] This is not to say that simpler cases are of no use; indeed, in his [Lange, 2017], Marc Lange considers nothing beyond the reach of an undergraduate in mathematics and in so doing assembles an important collection of phenomenological data. However, at some point, philosophers have to engage with the "embarrassment of riches" offered by contemporary mathematics ([Mancosu, 2008a]). I think the philosophical gains will be similarly rich, and I hope that this case study inclines my reader to think this as well.

[49] Answered affirmatively by Ben Green and Terry Tao in [Green and Tao, 2008]. This result is now known as the Green-Tao theorem. As they note, it is very interesting that their entirely finitary argument is conceptually closer to techniques in infinitary ergodic theory than it is to techniques in quantitative analysis. Indeed, they rely upon methods akin to those employed in the ergodic proof of Szemerédi's theorem. This indicates the importance and depth of these methods.

An arithmetic progression of interest to computational mathematicians is one of the longest known arithmetic progression of primes:[50]

Example 2.3.3. Take $a = 56211383760397$, $r = 44546738095860$ and $k = 0, 1, \ldots, 22$ for an arithmetic progression of primes given by $56211383760397 + 44546738095860k$.

Note that oftentimes we deal with *arbitrarily long* arithmetic progressions, where this is commonly taken to mean finite arithmetic progressions of arbitrary length, i.e., without the specification of some k.

In 1975 Szemerédi proved the following remarkable theorem, providing an answer to a long-standing conjecture of Turán and Erdős:[51]

Theorem 2.3.4 (Szemerédi; infinitary)**.** *Let A be a subset of the integers $\mathbb{Z}$ with positive upper density, i.e.,*

$$\delta(A) := \limsup_{N\to\infty} \frac{|A \cap [-N, N]|}{2N + 1} > 0. \tag{2.3.1}$$

Then A contains arbitrarily long arithmetic progressions.

This is equivalent via a routine compactness argument to the strictly finitary statement:

Theorem 2.3.5 (Szemerédi; finitary)**.** *For every $k \geq 1$ and real number $0 < \delta \leq 1$, there is an integer $N(k, \delta) \geq 1$ such that for every $N \geq N(k, \delta)$, every set $A \subset \{1, \ldots, N\}$ of cardinality $|A| \geq \delta N$ contains at least one arithmetic sequence of length k.*

Szemerédi's theorem is striking given its extreme generality. We lay down no conditions on the set A under consideration, except that it is quite large, and discover that any such A necessarily has a particular structure. Indeed, this is the sort of result that seems to call out for explanation, namely, an account of *why* the theorem holds. Discovering "the why" of Szemerédi's theorem turns out to be a difficult affair. Szemerédi's original proof was entirely combinatorial and finitary; indeed, it is paradigmatic of what one might call a "pure" proof. It is, however, incredibly intricate and has been judged, even by the very best professional mathematicians working in this field, as "remarkably subtle" ([Tao, 2006]). Indeed, it is quite rare, except in the most difficult and involved of proofs, to include a schematic diagram outlining the proof steps, but this is just what Szemerédi does in [Szemerédi, 1975] (see Appendix A).

Our aim will be to understand the strategy of the ergodic proof of Szemerédi's theorem and in so doing call attention to questions significant for an epistemology of mathematics, especially questions concerning the nature of mathematical content, mathematical explanation, and the role of infinitary reasoning. These questions will then be analyzed further in the strictly philosophical sections below. I will also attempt to provide some of the main

[50] Note that the result of Green-Tao in [Green and Tao, 2008] is non-constructive and thus merely proves the *existence* of arbitrarily long arithmetic progressions of primes.

[51] Both Szemerédi's theorem and the Green-Tao theorem on primes are special cases of the most general (and still open) Erdős-Turan conjecture for arithmetic progressions: Suppose that $A = \{a_1 < a_2 < \ldots\}$ is an infinite sequence of integers such that $\sum 1/a_i = \infty$. Then A contains arbitrarily long arithmetic progressions.

steps of the combinatorial argument to give a flavor of how it works, but only after the ergodic proof of the theorem. Indeed, I want to suggest that this way is best for generating any sort of understanding of *why* the theorem holds.

At the broadest level of analysis, the ergodic proof of Szemerédi's theorem consists of two main stages: (i) the proof of an equivalence between the original combinatorial statement of the theorem (Theorem(s) 2.3.4 and 2.3.5) and an ergodic analogue (Theorem 2.3.12; correspondence stage); (ii) the proof of the ergodic analogue via a "structure" theorem (Theorem 2.3.37; structure stage). I would like to note at the outset that the pure (combinatorial) proof of Szemerédi's theorem in [Szemerédi, 1975] also appeals to its own structure theorem (a combinatorial one; see Lemma 2.5.3 and Lemma A.1), but obviously bypasses the detour through ergodic theory. This detour is explanatorily significant primarily because the ergodic structure theorem is much conceptually cleaner than its combinatorial analogue and generates significant simplification of the entire proof. As a result, we are able to see *why* Szemerédi's theorem holds. All this will emerge in the discussion below. Let us now turn to the first stage of the ergodic proof, the correspondence stage, which, though less crucial for my analysis of mathematical explanation, is epistemologically significant in its own right.

2.3.1 Introduction to Correspondence Principles via van der Waerden's Theorem

Prima facie ergodic theory and combinatorics are quite unrelated to one another: the former involves the study of measure-preserving dynamical systems (infinitary and sometimes highly uncomputable objects), while the latter involves studying various properties of finite structures. Remarkably, Furstenberg, et al. demonstrated techniques by which one might understand and prove a combinatorial theorem in a dynamical setting. Techniques that "convert" a theorem from conceptual setting X to conceptual setting Y are known generally as *correspondence principles*. I believe correspondence principles have much to do with the uniqueness of mathematical epistemology and are worthy of careful philosophical study. In particular, these function somewhat like bridge laws/principles familiar from the philosophical literature. These "mathematical" bridge laws are important for my purposes because they facilitate the move from one theorem stated in a particular context to another statement in a different context equivalent over some stronger mathematical theory; this then allows for an expansion of explanatory resources for proof. Let me provide some examples and give a flavor of how these principles function.

Ultimately, we will be concerned with *Furstenberg Correspondence*; this is the first step in the ergodic proof of Szemerédi's theorem wherein combinatorial, finitary content of the above theorem is converted to a problem about recurrence patterns in dynamical systems. Before analyzing this correspondence principle, I will examine a simpler case: namely, the correspondence (Theorem 2.3.10) involved in van der Waerden's theorem, which is a special case of Szemerédi's theorem. This should help us to understand the more complicated case later on. Let's lay out the basic definitions we will need to make sense of these correspondence principles.

Definition 2.3.6 (Dynamical System)**.** A *dynamical system* is a pair (X, T) where X is a set (or abstract space) and $T : X \to X$ is an invertible map operating on the elements of X.

By $T^n x$ we understand T applied n times to some $x \in X$. Intuitively, we study the evolution of X as it is transformed by T over time.

For example, we might consider the finite system (X, T) where X is a finite set and T a permutation of the elements of X. Similarly, we might understand a dynamical system as the action of a group G on a set X. Thus far, at this level of abstraction, little interesting can be said. However, applying a little more structure to dynamical systems will yield surprising results. Specializing some (X, T) to the case where X is a compact metric space and T a homeomorphism[52] will provide a proof of the following:

Theorem 2.3.7 (van der Waerden)**.** *For any finite coloring of the integers, there are arbitrarily long arithmetic progressions.*

Like Szemerédi's theorem, this too is equivalent to a strictly finitary theorem:

Theorem 2.3.8 (van der Waerden; finitary)**.** *Let $k, m \geq 1$. Then any m-coloring of $\mathbb{N}$ contains a monochromatic progression of length k.*

Furstenberg showed that this could be understood in the setting of a topological dynamical system insofar as it is equivalent to the following recurrence theorem:

Theorem 2.3.9 (Multiple Recurrence in Open Covers)**.** *Let (X, T) be a topological dynamical system, i.e., X is a compact metric space and $T : X \to X$ is a homeomorphism.*[53] *Let $(U_\alpha)_{\alpha \in A}$ be an open cover of X. Then there is some U_α such that for every $k \geq 1$, we have $U_\alpha \cap T^n U_\alpha \cap \cdots \cap T^{(k-1)n} U_\alpha \neq \emptyset$ for some $n > 0$.*

Theorem 2.3.10 (Combinatorial-Topological Correspondence Principle)**.** *van der Waerden's theorem on arithmetic progressions and Multiple Recurrence in Open Covers are equivalent.*

I will prove neither van der Waerden's theorem nor Multiple Recurrence nor the attendant correspondence principle here. One reason for this is that the explanatory gains of "going infinitary and impure" in this case are less striking (though not absent) than for Szemerédi's theorem, and so I am inclined to skimp on the details. In particular, the combinatorial and topological proof strategies are very similar. Nonetheless, the simpler case of van der Waerden's theorem is still quite instructive, so let us examine the basic steps in the proof.

The combinatorial proof of van der Waerden's theorem essentially proceeds by double induction on k and m. For an excellent and careful build-up to the general argument, see [Katz and Reimann, 2018]. Tao's paper presents the proof via a "color focusing" strategy, which I find a bit clearer, though the essential ideas are the same.[54]

In any case, the strategy of the topological proof of van der Waerden's theorem is analogous to the combinatorial proof: it proves Multiple Recurrence on a simple sort of topological

[52] A *homeomorphism* is a map $f : X \to Y$ between topological spaces that is bijective, continuous, and has a continuous inverse.

[53] One may write X merely as a topological space and T as merely a continuous map, though it makes only a small difference in the proof.

[54] One might immediately wonder whether proofs by induction are explanatory or not. Indeed, this is a rather natural question and quite a lot of literature addresses it. See the many references provided in [Mancosu, 2018].

dynamical system[55] and then proceeds by double induction. However, there are some important epistemic advantages of the topological setting. Indeed, Tao makes the explicit comparison:

> By invoking this correspondence principle [that between combinatorics and topological dynamics] one leaves the realm of number theory and enters the infinitary realm of abstract topology. However, a key advantage of doing this is that we can now manipulate a new object, namely the compact topological space X. Indeed, the proof proceeds by first proving the claim for a particularly simple class of such X, the *minimal* spaces X, and then extending to general X. This strategy can of course also be applied directly on the integers, without appeal to the correspondence principle, but it becomes somewhat less intuitive when doing so...([Tao, 2007], 150-1).

The idea seems to be the following: "moving infinitary" allows us to abstract from many features of the combinatorial proof that render this proof difficult to understand, e.g., the need to keep track of many different parameters at once, while retaining the features that show why the result is true. Indeed, we see that topological spaces X encode at least some of the relevant finite coloring information expressed in van der Waerden's theorem. Tao gives us a nice and very simple example of this: if the coloring considered is such that one never sees a red integer immediately after a blue integer, this fact will be picked up by the ambient topological space X, i.e., X is disjoint from the set of points $\{(x_n), n \in \mathbb{Z} : x_0 \text{ is blue}, x_1 \text{ is red}\}$. I think this gets to the heart of what is fascinating about mathematics: we are able to move from concepts of type X to a *prima facie* different set of concepts Y and find that various results about these are in fact *equivalent* to results involving only type X. How can we make sense of this? Certainly some information is lost in the transition from combinatorics to topology, but nothing *essential* to the theorem we seek to prove: in both settings we can show that, with minimal assumptions on the ambient set or space, some pattern will always occur, viz., an arithmetic progression or recurrence in a dynamical system.

There are, then, several things we should notice. First, the notion of a correspondence principle requires further analysis and should be related to questions of "mathematical content," namely what are the various combinatorial theorems really *about* given that they are shown to be equivalent to seemingly foreign statements about dynamical systems? On an intuitive reading of mathematical content, subsets of $\mathbb{Z}$ are neither topological nor ergodic; however, are other, potentially more useful notions of content available (Section 2.4)? Second, what are the actual epistemic advantages of moving to a more "abstract" setting? In the case of van der Warden's theorem, the explanatory gains are not so clear given the similarity of the combinatorial and topological proofs; however, in the Szemerédi case, the similarities between proofs becomes much less explicit and the infinitary, ergodic setting is *much* conceptually cleaner than the finitary, combinatorial setting. I consider this at greater length in the section on impurity, simplicity, and explanation (Section 2.5).

Before moving on, let me provide a general schema for how proofs via correspondence

[55] An important difference between the dynamical proofs of van der Waerden's and Szemerédi's theorems is that in the latter case the "simple sort" of system we wish to consider does not exhaust the field of candidates. This is one way to see why van der Waerden is a special case of Szemerédi's.

principles are to work. This will help focus the arguments of the next section, which grow more complex.

1. Take a theorem of type X for which we may have a proof utilizing concepts and techniques endemic to type X, e.g., van der Waerden's and Szemerédi's theorems with their combinatorial proofs.

2. Find (or construct) a statement of type Y which seems to mirror the desired result of type X. This will then be the theorem to be proven.

3. Prove that there is an equivalence between the two theorems of types X and Y, i.e., a correspondence principle.

4. Prove theorem of type Y with resources endemic to type Y, thereby proving the theorem of type X.

With this preliminary information in hand, let me now explicitly turn to the ergodic proof of Szemerédi's theorem.

2.3.2 Ergodic Proof of Szemerédi's Theorem

Outline of the Proof

Here I provide a sketch of the ergodic proof.[56] This should be sufficient for most readers and will contain the main conceptual moves that I will discuss in my philosophical analysis. We begin with the correspondence stage of the proof: we have a problem about arithmetic progressions in the integers and wish to convert it to an ergodic setting. The basic idea here is that we can identify subsets $A \subset \mathbb{Z}$ with subsets of the ambient set or space X in the dynamical system. Similarly, we can identify functions on the integers with functions on the dynamical system. After this is done, rather surprisingly, the task of finding an arithmetic progression in some $A \subset \mathbb{Z}$ is in fact equivalent to finding an arithmetic progression in the subset of X identified with A. Let me try to make this a bit more precise (though saving definitions for the following section). We can show, for our dynamical system[57] (X, T), that there is a probability measure μ on X such that $\mu(E) = \mu(T^n E)$ for all $n \in \mathbb{Z}$ and measurable sets $E \subset X$ with E of positive measure. Then, the existence of some k length arithmetic progression in $A \subset \mathbb{Z}$ (the assertion of Szemerédi's theorem) is equivalent to the existence of some $x \in X$ such that a recurrence pattern $x, T^n x, \ldots, T^{(k-1)n} x$ is in E.

In particular, we can show that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu\left(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E\right) > 0 \tag{2.3.2}$$

[56] See the expository paper [Zhao,] for a very helpful and detailed analysis of the ergodic proof. I have benefitted a great deal from reading this.

[57] To be specified as a measure-preserving system.

for $\mu(E) > 0$. This implies the recurrence claim above, i.e., $x, T^n x, \ldots, T^{(k-1)n} x$ is in E, since it shows that the intersection of sets $E \cap T^n E \cap \cdots T^{(k-1)n} E$ is non-empty for some n. This concludes the correspondence principle stage of the proof. As I tried to draw out above, this is already quite remarkable. We have an equivalence between two results in two entirely different domains of mathematics and, furthermore, one domain deals with explicitly computational entities, while the other deals with highly abstract and infinitary entities. In addition to the questions I raised above, this move might also lead us to propose that some variety of unification is occurring. We see that the more abstract dynamical setting provides us with a conceptual framework to understand recurrence properties of various maps T, but also a specific kind of structure on the integers. Understanding precisely what kind of unification this is requires some work. However, this suggests the possibility that unification and explanation in mathematics might go together very often, unlike, for instance, in the case of the physical sciences.[58]

The second stage of the proof involves showing 2.3.2 holds. I believe this is where the ergodic proof yields great explanatory dividends. Let me first outline the situation we find ourselves in when trying to prove 2.3.2. In short, any proof of 2.3.2 requires decomposing the system under consideration into a structured component and a random component, which, by the Correspondence Stage of the proof, is equivalent to decomposing any $A \subset \mathbb{Z}$ into a structured set and a random set. Tao presents the situation with usual adeptness as a "fundamental dichotomy between structure and randomness." He continues,

> the reason for the truth of Szemerédi's theorem is very different in the cases when A is random, and when A is structured. These two cases can then be combined to handle the case when A is (or contains) a large (pseudo-)random subset of a structured set. Each of the proofs of Szemerédi's theorem[59] now hinge on a *structure theorem* which, very roughly speaking, asserts that *every* set of positive density is a large pseudo-random[60] subset of a structured set ([Tao, 2006], 583).

I will have more to say about the usage of structure theorems below. Returning to the proof sketch, it turns out that in the ergodic setting the requisite structure theorem and the dichotomy between structure and randomness are particularly explicit. This fact and the simplification it effects in proving Szemerédi's theorem renders the ergodic proof explanatory. Let's sketch how the structure theorem is applied in the ergodic setting.

The dichotomy presents itself as that between the periodicity of E under transformation T (structured) or the mixing of E under transformation T (random). If E is periodic, then this simply means $T^l E = E$ for some $l > 0$. Establishing 2.3.2 is then trivial as the summand will be $\mu(E)$ whenever n is a multiple of l. Even in the less well-behaved case

[58] See my remarks in the Introduction and Section 2.7.

[59] Of which, at the moment, (and not including generalizations) there are four: (i) combinatorial, (ii) ergodic, (iii) Fourier analytic, (iv) hypergraph.

[60] Just as with "random" and "structured," one must make explicit what is meant by "pseudo-random" in a particular context. Very generally, we can say that a pseudo-random set of integers resembles a random set of integers with similar density in terms of particular arithmetic statistics. For instance, in [Green and Tao, 2008], Green and Tao construct a superset of the primes (the "almost primes"), which is pseudorandom in the sense that it satisfies a "linear forms condition" and a "correlation condition." These conditions correspond very closely to the ergodic notion of weak-mixing (defined below).

of near-periodicity, we can show 2.3.2 in a very similar fashion. When all E for a space X are near-periodic, then we say the dynamical system is *compact*. On the other hand, we get the random scenario when E has a "mixing" property. Intuitively, this can be thought of as follows: if E is some event in the probability space X, then T "mixes" the space randomly such that all events $\{E, T^1E, T^2E, \ldots\}$ are independent of one another. This yields 2.3.2 trivially as then each recurrence pattern is just $\mu(E)^k$. Similar to the structured case, even when we relax the mixing to the case of "weak mixing," i.e., for sufficiently large intervals of time, E and T^nE become uncorrelated,[61] 2.3.2 can be proven. When all E in X have this weak-mixing property, we say that the system is a weak mixing system. In either the compact or weak mixing cases, we can prove our desired result. I discuss this in greater detail below.

However, not all systems are either compact or weak-mixing.[62] Nonetheless, provided that X is not completely weak-mixing (i.e., totally random), we can always find some structured component Y of X, which will itself give rise to a compact system. We can then analyze the map from $X \to Y$, called an *extension map*, which itself has structured or random behavior. If $X \to Y$ is a weak-mixing or compact extension, then, even if X is neither of these, we can project the problem down to Y. However, like systems, not all extensions are either weak-mixing or compact. If $X \to Y$ is some such intermediate case, then we can find some intermediate extension $X \to Y_1 \to Y$ such that $Y_1 \to Y$ is a compact extension. We can then pass to Y and show 2.3.2 holds. Thus, we must show that if 2.3.2 holds in Y_1 it holds in X. This process can be continued (transfinitely) by constructing a *tower* $X \to Y_\alpha \to \cdots \to Y_1 \to Y$ where each step $Y_{n+1} \to Y_n$ is a compact extension and $X \to Y_\alpha$ is weak-mixing. We can then lift property 2.3.2 from Y all the way up the tower to X itself. This is the essence of the Furstenberg Structure Theorem (Theorem 2.3.37).

Thus, the ergodic approach to Szemerédi's theorem yields a highly perspicuous proof structure, which, in every case considered, will us tell *why* a system (and thus subset of integers) has the recurrence pattern (arithmetic progressions) it does:[63]

1. Correspondence Principle Stage
   (a) Convert Szemerédi's theorem to a theorem about recurrence in a measure-preserving system, i.e., Furstenberg Multiple Recurrence (Theorem 2.3.12 below).
   (b) Prove that these theorems are equivalent.
2. Structure Theorem Stage
   (a) Prove Furstenberg Multiple Recurrence.
   (b) This must be done for any measure-preserving system for arbitrary k.
   (c) Classify each system as either structured (compact), random (weak-mixing), or neither. If the system is structured or random, we are done.

[61] Informally: knowing the initial state is no guide to the final state.

[62] This differs from the topological proof of van der Waerden's theorem where all spaces under consideration will either be minimal or not.

[63] Note, however, that the ergodic proof is *not* a proof by cases; I discuss this further below. See Section 2.3.4.

(d) If the system is neither, we can construct a tower via the Furstenberg Structure Theorem, and eventually reduce to either the compact or weak mixing case. Q.E.D.

Furstenberg Correspondence

Let me now spell out in greater mathematical detail how the Correspondence Principle stage of the proof is to work. Again, we claim that Szemerédi's theorem is equivalent to a recurrence theorem on measure-preserving systems (Furstenberg correspondence principle). To that end, define:

Definition 2.3.11. A *measure-preserving system* is a quadruple $(X, \mathscr{B}, \mu, T)$ where X is a set (or abstract space), $\mathscr{B}$ is a σ-algebra of subsets of X, $\mu : \mathscr{B} \to [0,1]$ is a probability measure, namely, μ is additive and $\mu(X) = 1$, and $T : X \to X$ is a measurable transformation such that $\mu(T^{-1}(E)) = \mu(E)$ for all $E \in \mathscr{B}$.

Then we have

Theorem 2.3.12 (Furstenberg Multiple Recurrence)**.** *Let $(X, \mathscr{B}, \mu, T)$ be a measure-preserving system. Then for any $E \in \mathscr{B}$ of positive measure and any $k \geq 1$, there exists some $n > 0$ such that*

$$\liminf_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n} E) > 0. \tag{2.3.3}$$

We now show that Furstenberg Multiple Recurrence is in fact equivalent to Szemerédi's theorem, thus establishing our desired correspondence principle. There are many ways to go about this; the most natural is to exploit a correspondence between sets in the system and sets in the integers.

First, we identify sets E in the measure-preserving system with sets A in the integers and likewise functions f on the system and functions F on the integers. Take $E \in \mathscr{B}$ and some $x \in X$. We then define the *recurrence set* $A_{x,E} \subset \mathbb{Z}$ as

$$A_{x,E} := \{n \in \mathbb{Z} : T^n x \in E\}. \tag{2.3.4}$$

Likewise, given some function $f : X \to \mathbb{R}$ on $(X, \mathscr{B}, \mu, T)$ and $x \in X$, we define an associated sequence $F_{x,f} : \mathbb{Z} \to \mathbb{R}$ as

$$F_{x,f}(n) := f(T^n x). \tag{2.3.5}$$

Now we show that some set in the measure-preserving system will yield a recurrence set $A_{x,E}$ of positive upper density, as in the combinatorial version of Szemerédi's theorem.

Lemma 2.3.13. *Let $(X, \mathscr{B}, \mu, T)$ be a measure-preserving system, let $E \in \mathscr{B}$ be of positive measure. Then there is some set F of positive measure such that the recurrence set $A_{x,E}$ has positive upper density for all $x \in F$.*

Now we can show the equivalence we desire.

Theorem 2.3.14 (Furstenberg Correspondence)**.** *The Furstenberg Multiple Recurrence theorem and Szemerédi's theorem are equivalent.*

Proof. **Szemerédi** $\Rightarrow$ **FMR**: Take the set of all k-term arithmetic progressions in $\mathbb{Z}$ and call it $\mathscr{A}_k$. Let X_a be the set of points $x \in X$ such that $T^{a_j}x \in E$ for each $j = 1, 2, \dots k$. Then let

$$X_k := \bigcup_{a \in \mathscr{A}_k} X_a \tag{2.3.6}$$

which is the set of all points $x \in X$ such that our recurrence set $A_{x,E} := \{n \in \mathbb{Z} : T^n x \in E\}$ contains some arithmetic progression.

Now we use set F constructed in the above Lemma. By Szemerédi's theorem, $A_{x,E}$ contains a k-term arithmetic progression for all $x \in F$. Hence, $F \subset X_k$ and $\mu(X_k) \geq \mu(F) > 0$. $\mathscr{A}_k$ is countable and so $\mu(X_a) > 0$ for any arithmetic progression $a \in \mathscr{A}_k$. Thus,

$$T^l X_a \subset E \cap T^n E \cap \cdots \cap T^{(k-1)n} E \tag{2.3.7}$$

for some $l \in \mathbb{Z}$ and $n > 0$. But T is measure-preserving and X_a has positive measure, so $T^l X_a$ must be also. Thus, we have Furstenberg Multiple Recurrence.

FMR $\Rightarrow$ **Szemerédi**: The idea here is to apply Furstenberg recurrence to a rather unconventional dynamical system. This will then yield a number-theoretic result. One can think about this as starting with a set in the integers and generating its dynamical analogue to which FMR is then applied. We proceed by explicitly constructing a T-invariant measure μ such that $\mu(E) > 0$ where E is some subset of our unconventional dynamical system (X, T).

This proof follows [Tao, 2007] closely. Suppose for contradiction that Szemerédi's theorem does not hold. Then there is some $k \geq 1$ and some $A \subset \mathbb{Z}$ without an arithmetic progression of length k and some sequence N_i of integers such that $\liminf_{i \to \infty} \frac{|A \cap [-N_i, N_i]|}{2N_i + 1} > 0$.

Now let us construct our unconventional dynamical system. Let Λ be some finite set and form $\Omega := \Lambda^{\mathbb{Z}}$, i.e., the set of all sequences with entries from Λ. Let X be any closed T-invariant subset of Ω, where T is the shift-map, i.e., for some $x \in \Omega$, $Tx(n) \mapsto Tx(n-1)$. Then (X, T) is the dynamical system we desire. In particular, we can identify any subset of $\mathbb{Z}$ with elements of $X := \{0, 1\}^{\mathbb{Z}}$.

Now for each i indexing N_i, we consider the random set $A_i := A + x_i$ where x_i is a randomly chosen integer from $[-N_i, N_i]$. We can then think of A_i as a random variable taking values in X which, on the σ-algebra $\mathscr{B}$ of X, we identify with a probability measure μ_i on X. In essence, we have now constructed our measure-preserving system to which FMR can be applied. Using properties of X (separable compact Hausdorff space) we get that the μ_i for

each random variable weakly converge to some probability measure μ, i.e.,

$$\lim_{i\to\infty}\int_X f\,d\mu_i = \int_X f\,d\mu \tag{2.3.8}$$

for any continuous f on X. Now let $E \in \mathscr{B}$ be given by $E := \Big\{(x_n)_{n\in\mathbb{Z}} \in \{0,1\}^{\mathbb{Z}} : x_n = 1\Big\}$. E is both open and closed so

$$\lim_{i\to\infty}\mu_i(E) = \mu(E). \tag{2.3.9}$$

Finally, given the way in which we constructed A_i and E,

$$\mu_i(E) = \frac{|A \cap [-N_i, N_i]|}{2N_i + 1} \tag{2.3.10}$$

and this is of positive measure. Since we selected T to be the shift-map, i.e., $T(x_n)_{n\in\mathbb{Z}} := (x_{n-1})_{n\in\mathbb{Z}}$, we have that $\lim_{i\to\infty}\mu_i(TE) - \mu_i(E) = 0$. Thus, in general, for any finite boolean combination of E and its shifts, we have that μ is T-invariant.

By hypothesis, A contains no arithmetic progressions of length k, so after taking limits,

$$\mu(E \cap T^n E \cap \cdots \cap T^{(k-1)n}E) = 0, \tag{2.3.11}$$

for any $n > 0$. But this contradicts FMR. Thus, Szemerédi's theorem holds. □

2.3.3 Proof of Furstenberg Multiple Recurrence via the Furstenberg Structure Theorem

Preliminary Mathematical and Philosophical Remarks

Now consider the Structure Theorem stage of proof that provides the reason why Szemerédi's theorem holds. But what, precisely, do I mean by this? Recall that, on the number-theoretic version of Szemerédi's theorem, we consider arbitrary subsets $A \subset \mathbb{Z}$ and show that they contain arbitrarily long arithmetic progressions. Unsurprisingly, given the lack of explicit constraints on A, one can find such arithmetic progressions in various subsets for very different reasons.[64] For example, consider some *random* subset A of $\mathbb{Z}$, where this means each integer n in A has an independent probability of δ with $0 < \delta < 1$. It is now quite simple to show that A in fact has infinitely many arithmetic progressions of length k because any such progression will have a probability of δ^k of lying in A.[65] On the other hand, were we to consider some "structured" set, we would get the same answer for an entirely different reason. For example, consider the *Bohr set* $\left\{n \in \mathbb{Z} : ||n\alpha||_{\mathbb{R}/\mathbb{Z}} \leq \delta\right\}$ with δ as above, $\alpha \in \mathbb{R}$, and $||\cdot||$ yielding the distance between the argument and the nearest integer. For any α,

[64] See [Tao, 2006] for an extended discussion.

[65] Take arbitrary k-length arithmetic progression $a, a+r, \ldots, a(k-1)r$. Each term contributes an independent probability of δ and we have k terms, thus the total probability of finding the arithmetic progression is δ^k.

each αn can be made arbitrarily close to n and so each n is correlated with periods of α. The sequence of such periods will then just be the arithmetic progression we seek.[66]

As I will sketch in greater detail below, the ergodic situation precisely mirrors the number-theoretic situation. Namely, we will obtain Furstenberg Multiple Recurrence in an arbitrary measure preserving system because systems exhibit either sufficient randomness (weak mixing) or sufficient structure (compactness). In the weak mixing case, on average, the events $T^n E, T^{2n} E, \ldots, T^{(k-1)n} E$ are uncorrelated such that the measure of E is approximately $\mu(E)^k$ fairly often (thus satisfying 2.3.2). As in the random number-theoretic case, since the systems under consideration are sufficiently mixing (random), sooner or later we will find the desired recurrence property by sheer luck. We will also find the recurrence property in sufficiently structured systems (compact) because these systems will have events $T^n E$ recur at regular intervals, i.e., the intersection of E and $T^n E$ over k-many iterations of T^n will be non-empty. Again, this is analogous to the number-theoretic case described above: sufficient structure in the original set (or system) will guarantee the existence of the desired arithmetic progression (or recurrence property).

The moral here is that we can find arbitrarily long arithmetic progressions in random sets, structured sets, and various sets in between; however, the reason we can do this always turns upon the classification of the set itself. This would seem to dash any hopes of a general reason for a result like Szemerédi's theorem; namely, the proof of the theorem would have to amount to checking the existence of an arithmetic progression for each A, given some classification of A. Such a task seems hopeless and I daresay would not produce an explanation of why the theorem is true. One could not then even state the theorem as such; rather, we would have some conjunction of independent theorems, each showing that some set A_i has an arbitrarily long arithmetic progression given the classification of set A_i. In order to avoid this situation, one would hope for a result that would allow us to group *prima facie* very different subsets A into one class that shares some feature F, which in turn allows us to understand why arithmetic progressions occur in each A. And this is just what the Furstenberg Structure Theorem provides for the ergodic equivalent of Szemerédi's theorem. Indeed, it should be noted that *every* proof of Szemerédi's theorem (combinatorial, ergodic, Fourier analytic, hypergraph) makes essential use of some structure theorem that asserts every set of positive density is or contains a large (pseudo-)random subset of a structured set. Thus, the reason why we can always find the desired arithmetic progressions for any large subset $A \subset \mathbb{Z}$ is that every subset can be decomposed into components that we know how to handle, thereby obviating the need to classify each A and check for arithmetic progressions.

One might, at this point, complain that I have ignored essential details about the subsets A under consideration. One might think that the real reason why some A has the arithmetic progressions it does is because of the kind of set A is, e.g., highly structured, highly random, etc., and that the kind of high-level structural result I am gesturing at abstracts from these details. The way to actually *explain* why a particular set has the arithmetic progressions it does requires a careful analysis of *that* set in particular. There may be something to this complaint; indeed, I think this may be the mathematical analogue to Morrison's contention that techniques of unification invariably abstract from "mechanisms" of physical systems

[66] [Tao, 2006] gives many other examples, but I choose this one because it is entirely analogous to the example of a structured system I consider below, the Circle Rotation system.

which explain the behavior of the system. I will consider this in greater detail below (Section 2.7). However, though this kind of criticism may be fairly levelled at particular high-level structural results, it does not really hold here. As we shall see, the Furstenberg Structure Theorem gives an exhaustive means of characterizing each system (and so set) without, as it were, digging into the nitty gritty details of the particular system or set. When we apply the structure theorem to a particular system, even though the system might lie anywhere along the spectrum of structure and randomness, we will be able to classify the system precisely. This classification does rely on our understanding the two ends of spectrum, i.e., why a random system and why a highly structured system will have the recurrence pattern. However, this is the only requisite contribution of "lower-level" or, if I can be allowed the expression, "mechanistic" content. Once we know how to deal with the extremes and have the structure theorem in hand, the actual classifications of the various systems (sets) do not really matter and certainly do not contribute to the explanation of why Szemerédi's theorem holds. We bypass examining each particular system because each system will have the pattern it does in virtue of the same reasons that compact and weak mixing systems have this pattern; however, the only reason we can assert this is because of the Furstenberg Structure Theorem. Let me now turn to this.

Introducing the Structure Theorem

There is an enormous amount of detail involved in explicitly proving the Furstenberg Structure Theorem and thus Furstenberg Multiple Recurrence (and thus, by Furstenberg Correspondence, Szemerédi's theorem). I cannot enter into all these details here, nor would it be terribly helpful for the philosophical aims of this paper.[67] [Furstenberg, 1977] contains Furstenberg's original ergodic proof of Szemerédi's theorem, though the proof strategy is less clear-cut than the 1979 and 1982 papers, precisely because the original proof does not make essential use of the (full) Furstenberg Structure Theorem. [Furstenberg and Katznelson, 1979] and [Furstenberg et al., 1982] do so and, in fact, prove an even stronger result. I primarily follow [Furstenberg et al., 1982] in my presentation below. I have also referred to the very helpful expository article by Zhao [Zhao,].

It is convenient to state Furstenberg Multiple Recurrence (Theorem 2.3.12) in terms of functions. Namely, we are interested in showing that the following property holds for every measure-preserving system:

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f\, T^n f\, T^{2n} f\, T^{(k-1)n} f\, d\mu > 0, \tag{2.3.12}$$

for $f \in L^\infty(X), f \geq 0$, f not a.e. zero. The formulation of (2.3.2; Theorem 2.3.12) and 2.3.12 are in fact equivalent because one can set $f = \chi_E$, i.e., f is simply the indicator function of some set E of positive measure in the σ-algebra $\mathscr{B}$ of the measure-preserving system, and because, if $\int_X f\, d\mu > 0$, we have the bound $f \geq c\chi_E$ for some $c > 0$. Following

[67] See [Furstenberg, 1977], [Furstenberg and Katznelson, 1979], and [Furstenberg et al., 1982] for full details. [Tao, 2007] is also elucidating, but he does not offer explicit proofs; rather, he provides motivation for various crucial moves involved in proving the Structure Theorem. Bryna Kra's lecture notes are also excellent and I have made use of these. See [Kra, 2007].

the literature, we say that any measure preserving system that satisfies 2.3.12 and thus Furstenberg Multiple Recurrence is SZ (for "Szemerédi").

We seek to prove that all measure-preserving systems are SZ, and thus, by Furstenberg Correspondence, that Szemerédi's theorem holds. We proceed by examining the various kinds of measure-preserving systems and ask whether each kind is SZ. It is trivial to prove this for any measure-preserving system when $k = 1, 2$. The case of $k = 3$, in combinatorial dress, is known as Roth's theorem and was established in 1953 by K.F. Roth. I will not consider these cases here, but rather the general claim for arbitrary k over all measure-preserving systems.[68]

This in turn can be done by considering in more detail the "dichotomy between structure and randomness." I examine two kinds of measure-preserving systems: compact (structured) and weak mixing (random), which can be proven to be SZ in a relatively straightforward fashion. The ultimate point will be that these extremes do not exhaust the field of candidate measure-preserving systems. This is *unlike* van der Waerden's theorem, in which there is only one kind of structured topological system, the minimal systems, and transferring the desired recurrence property from minimal systems to arbitrary topological systems is trivial using Zorn's lemma. Because of the additional complexity of the ergodic case, we require the Furstenberg Structure Theorem, the key to proving Furstenberg Multiple Recurrence and Szemerédi's theorem.

The strategy for the remainder of this section is as follows: Show that all compact and all weak mixing systems are SZ. Then, show that any system X is either weak mixing or has a nontrivial compact component. Next, show that this property is preserved under all compact extensions, weak mixing extensions, and limits between components. Finally, show that *any* measure preserving system can be decomposed into a transfinite sequence of factors (intuitively, subsystems of the original system) such that the "base" factor is trivial (and compact), each extension is either compact or weak mixing, and thus that any measure preserving system is SZ. This last step is the content of the Furstenberg Structure Theorem.

Weak Mixing and Compact Systems

I attempted to describe above, in an intuitive fashion, why weak mixing (random) and compact (structured) systems will be SZ. Let me now formulate these types of systems in a bit more detail.

Definition 2.3.15. We say that a system $(X, \mathscr{B}, \mu, T)$ is *mixing* if the following condition holds:

$$\lim_{n\to\infty} \mu(T^n E_1 \cap E_2) = \mu(E_1)\mu(E_2) \tag{2.3.13}$$

for all $E_1, E_2 \in \mathscr{B}$. In essence, this tells us that for sufficiently large n, the probability of being in E_2 at time t_0 and in E_1 at time t_n is the product of the individual probabilities. Namely, over sufficiently large periods of time, events are uncorrelated. Thus, a mixing

[68] It is interesting to note that Szemerédi proved the (combinatorial) case of $k = 4$ in 1969. Proving the theorem for arbitrary k then took six years, an indication of its great difficulty.

system is a highly random system in the sense that events on the probability space are uncorrelated.

As noted in the proof outline, any mixing system will trivially satisfy Furstenberg Multiple Recurrence. Consider then the weaker randomness condition:

Definition 2.3.16. We say that a system is *weak mixing* if the following condition holds:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} |\mu(T^n E_1 \cap E_2) - \mu(E_1)\mu(E_2)| = 0 \tag{2.3.14}$$

for all $E_1, E_2 \in \mathscr{B}$.

Weak mixing may be formulated in the following ways (at least):

Proposition 2.3.17. *Let $(X, \mathscr{B}, \mu, T)$ be a system. Then the following are equivalent:*

1. *The system is weak mixing;*
2. *There is a set $S \subset \mathbb{N}$ of density zero such that, for all E_1, E_2,*

$$\mu(T^n E_1 \cap E_2) \to \mu(E_1)\mu(E_2), \tag{2.3.15}$$

   as $n \to \infty$ and $n \notin S$.

3. *The system $(X \times X, \mathscr{B} \times \mathscr{B}, \mu \times \mu, T \times T)$ is ergodic, where ergodicity is the property that all T-invariant sets of a system, i.e., all $E \in \mathscr{B}$ with $TE = E$, have measure 0 or 1.*

Consider the following example of a weak mixing system.

Example 2.3.18. (Bernoulli System) We can construct a weak mixing measure preserving system from the previously mentioned "unconventional" dynamical system called a *Bernoulli system*. Consider the space X of all sequences $\{x_n\}_{n\in\mathbb{Z}}$ taking values in some finite set, say $\Gamma := \{1, \dots, r\}$. We can then formulate the σ-algebra $\mathscr{B}$ for X by letting $\mathscr{B}$ be the minimal σ-algebra for which each evaluation $x \mapsto x_n$ is measurable. The probability measure μ on $\mathscr{B}$ will then just be the product measure

$$\mu(x_{i_1} = j_1, x_{i_2} = j_2, \dots, x_{i_n} = j_n) = p_{j_1} \cdots p_{j_n} \tag{2.3.16}$$

where we equip the finite set Γ with a probability distribution $p_1, \dots, p_r$. The measure preserving transformation T is just the shift transformation from above.

Proposition 2.3.19. *Let $(X, \mathscr{B}, \mu, T)$ be the Bernoulli System. Then*

$$\lim_{n\to\infty} \mu(E_0 \cap T^n E_1 \cap T^{2n} E_2 \cap \cdots \cap T^{(k-1)n} E_{k-1}) \to \mu(E_0)\mu(E_1) \cdots \mu(E_{k-1}). \tag{2.3.17}$$

Proof. (Sketch) It suffices to prove this for any $E' \in \mathscr{B}$ that takes its values on a finite fixed set of terms. These are dense in $\mathscr{B}$, and so any measurable E can be approximated

arbitrarily closely to some E'. Thus, we can obtain the proposition from considering any such E' simply by taking limits. But the proposition is then quite trivial, since, as $n \to \infty$, the terms defining each $E_0, T^n E_1$, etc. will be disjoint, and so the events in the probability space are all independent. Thus, the measure of the intersection of all such events is just the product of the measure of the events. □

Thus, the system will be SZ since

$$\lim_{n\to\infty} \mu(E_0 \cap T^n E_1 \cap T^{2n} E_2 \cap \cdots \cap T^{(k-1)n} E_{k-1}) = \mu(E)^k > 0. \tag{2.3.18}$$

In fact, this is true for all weak mixing systems.

Theorem 2.3.20. *If a measure preserving system $(X, \mathscr{B}, \mu, T)$ is weak mixing, then it is SZ.*

Proof. See [Furstenberg et al., 1982], Section 3. The theorem follows from a number of classical results in analysis, but the central insight is the same as in the example above: for weak mixing systems $\mu(\bigcap_{j=0}^{k-1} T^{jn} A)$ is approximately $\mu(A)^k$ for most n. □

Consider next an example of a highly structured system. Analogous to the "extreme" randomness exhibited by a mixing system, an "extreme" case of structure is exhibited by a periodic system, i.e., the case where T is periodic such that $T^n A = A$. Obviously, this will be SZ. Let us then weaken periodicity to "almost periodicity." This is easily grasped by examining the following example.

Example 2.3.21. (Circle Rotation) Consider $(X, \mathscr{B}, \mu, T)$ where X is the circle $\mathbb{R}/\mathbb{Z}$ with σ-algebra $\mathscr{B}$, uniform probability measure μ (here Lebesgue measure), and the shift/rotation transformation $T : x \mapsto x + a$ with $a \in \mathbb{R}$. There are two ways in which this system might evolve under the action of T: (i) a rational; (ii) a irrational. If a is rational, then we get the trivial periodic case. If irrational, T will be "almost periodic," in the sense that T^n can be made arbitrarily close to the identity transformation, since any na can be made arbitrarily close to any integer.[69] Thus, the transformation in question will also preserve the structure of the system as the transformation is either periodic or almost periodic.

More precisely, we can prove that the Circle Rotation system is SZ.

Proposition 2.3.22. *Let $(X, \mathscr{B}, \mu, T)$ be the Circle Rotation system. This system satisfies Furstenberg Multiple Recurrence, i.e., for any $E \in \mathscr{B}$, we have*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(E \cap T^n E \cap T^{2n} E \cap \cdots \cap T^{(k-1)n} E) > 0. \tag{2.3.19}$$

Proof. Let $E \in \mathscr{B}$ with $\mu(E) > 0$. Since $\int_E \chi_E(x + y) d\mu(x)$ is a continuous function of y, for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\mu(E \cap (E - y) \cap (E - 2y) \cap \cdots (E - (k-1)y)) > \mu(E) - \epsilon \tag{2.3.20}$$

[69] Compare to the example of the Bohr set given in Section 2.3.3.

whenever $|y| < \delta$. Select $\epsilon = \frac{1}{2}\mu(E)$. Then

$$\mu(E \cap T^n E \cap T^{2n} E \cap \cdots \cap T^{(k-1)n} E) = \mu(E \cap (E+na) \cap (E+2na) \cap \cdots \cap (E+(k-1)na))$$
$$> \mu(E) - \epsilon = \frac{1}{2}\mu(E) > 0.$$

□

Just as the Bernoulli system was an instance of the general class of weak mixing systems, the Circle Rotation system is an instance of *compact* systems. We find that all compact systems also satisfy Furstenberg Multiple Recurrence. It is here that the move from the set formulation to the function formulation in 2.3.12 becomes convenient.

Definition 2.3.23. A system $(X, \mathscr{B}, \mu, T)$ is *compact* if for every $f \in L^2(X, \mathscr{B}, \mu)$ the closure in $L^2(X, \mathscr{B}, \mu)$ of the orbit $\{f, Tf, T^2 f, \ldots, T^n f, \ldots\}$ is compact. The closure and compactness in question occur in the norm topology of $L^2(X, \mathscr{B}, \mu)$.

More intuitively, this means that any such orbit is totally bounded; namely, given $\epsilon > 0$, there is an n such that every $T^j f$ is ϵ away from $\{f, Tf, T^2 f, \ldots, T^n f, \ldots\}$ in the $L^2(X, \mathscr{B}, \mu)$ norm.

Compact systems are "structured" in the sense that each application of T to some $E \in \mathscr{B}$ will return sufficiently close to E, and so each iteration of T will overlap somewhat, thereby satisfying Furstenberg Multiple Recurrence. That is

Theorem 2.3.24. *If $(X, \mathscr{B}, \mu, T)$ is compact, then for any $f \in L^\infty(X, \mathscr{B}, \mu)$, $f \geq 0$, f not a.e. 0, we have*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int_X f\, T^n f\, T^{2n} f \cdots T^{(k-1)n} f \, d\mu > 0. \tag{2.3.21}$$

That is, every compact system is SZ.

Proof. Again, see [Furstenberg et al., 1982], Section 4. The proof more or less comes down to the fact that, for any $f \in L^\infty(X, \mathscr{B}, \mu)$ such that $T^n f$ approximates f,

$$\int_X f\, T^n f\, T^{2n} f \cdots T^{(k-1)n} f \, d\mu \approx \int_X f^k \, d\mu > 0. \tag{2.3.22}$$

□

Dealing With Other Cases: Factors and the Structure Theorem

As I have indicated repeatedly, the compact (structured) systems and weak mixing (random) systems do not exhaust all cases for which we must prove Furstenberg Multiple Recurrence. We require, then, some way to analyze the other types of systems. Somewhat miraculously, if we consider these two extremes of structured and random systems, we can prove Furstenberg Multiple Recurrence for some "component" of an arbitrary system. Finally, by constructing

the tower of extensions mentioned above, we can prove that successively larger factors are SZ until we arrive at the arbitrary system under consideration also being SZ. This completes the proof.

Let me begin by defining "components" or *factors* of systems:

Definition 2.3.25. For some measure preserving system $(X, \mathscr{B}, \mu, T)$, a *factor* is a T-invariant sub-algebra of $\mathscr{B}$. That is, some σ-algebra $\mathscr{B}' \subseteq \mathscr{B}$ such that $TE, T^{-1}E \in \mathscr{B}'$ for any $E \in \mathscr{B}'$. More generally, we call a factor a subsystem $X' := (X, \mathscr{B}', \mu, T)$. A factor will be *trivial* if the measure of any $E \in \mathscr{B}'$ is either 0 or 1. It will be *compact* if X' is itself a compact measure preserving system.

This definition is a special case of the following more general definition:

Definition 2.3.26. (Extensions) Let $X := (X, \mathscr{B}, \mu, T)$ and $Y := (Y, \mathscr{C}, \nu, S)$ be measure preserving systems. An *extension map* or *factor map* $\pi : X \to Y$ is a measure preserving map, i.e., for any $C \in \mathscr{C}$, then $\pi^{-1}(C) \in \mathscr{B}$ and $\mu(\pi^{-1}(C)) = \nu(C)$, that is also shift compatible, i.e., $\pi \circ T = S \circ \pi$. For any extension map $X \to Y$, we say Y is a factor of X, and X is an extension of Y. Certainly, when we have some factor $(X, \mathscr{B}', \mu, T)$ of $(X, \mathscr{B}, \mu, T)$ where $\mathscr{B}' \subseteq \mathscr{B}$, then $\pi : (X, \mathscr{B}, \mu, T) \to (X, \mathscr{B}', \mu, T)$ is measure preserving and shift compatible.

One can then show the following:

Theorem 2.3.27 (Koopman-von Neumann Dichotomy Theorem)**.** *For some measure preserving system $(X, \mathscr{B}, \mu, T)$, exactly one of the following holds: either the system is weak mixing or the system has a nontrivial compact factor.*

This then allows us to apply either Theorem 2.3.20 or Theorem 2.3.24 to some factor of an arbitrary system. Let me now sketch the proof of the Dichotomy Theorem; I do this in greater detail than much of the above because this theorem is the simplest example of the absolutely crucial "dichotomy between structure and randomness."

Proof. Assume first that $(X, \mathscr{B}, \mu, T)$ is weak mixing. Then we wish to prove that this system has no nontrivial compact factor. This reduces to showing (i) that there are no (nonconstant) $f \in L^2(X, \mathscr{B}, \mu)$ with an orbit of compact closure, i.e., $\overline{\{T^n f\}}_{n\in\mathbb{N}} \subset L^2(X, \mathscr{B}, \mu)$. From the compactness of the orbit closure, we can find a finite subset $\{T^{n_1}f, T^{n_2}f, \ldots, T^{n_k}f\} \subset \{T^n f\}$ such that $||T^{n_i}f - T^{n_j}f|| \geq \epsilon$ for any $\epsilon > 0$. We can extend this such that, for any n, $\{T^{n+n_1}f, T^{n+n_2}f, \ldots, T^{n+n_r}f\}$ also has the property $||T^{n+n_i}f - T^{n+n_j}f|| \geq \epsilon$ and has the same cardinality of the initial finite subset. Thus, for any j, $||T^{n+n_j}f - f|| < \epsilon$. Therefore, for any $\epsilon > 0$, any subset $S \subset \mathbb{N}$ of positive lower density, and $n \in S$, we have $||f - T^n f|| < \epsilon$ in the $L^2(X, \mathscr{B}, \mu)$-norm.

Further, the definition of weak mixing gives, for any $f, g \in L^2(X, \mathscr{B}, \mu)$ and $\delta > 0$

$$\left| \int f\, T^n g\, d\mu - \int f\, d\mu \int g\, d\mu \right| < \delta \tag{2.3.23}$$

for n a member of a set of density 0. Indeed,

$$\left| \int f\, T^n f\, d\mu - \left(\int f\, d\mu \right)^2 \right| < \delta \tag{2.3.24}$$

for such a set. Then, for some such n

$$\left| \int f^2\, d\mu - \left(\int f\, d\mu \right)^2 \right| < \delta + \epsilon ||f||_{L^2(X,\mathscr{B},\mu)}. \tag{2.3.25}$$

But ϵ, δ were arbitrary, so

$$\int f^2\, d\mu = \left(\int f\, d\mu \right)^2, \tag{2.3.26}$$

which for real-valued f tells us that f is constant a.e. This finishes (i).

Now assume that $(X, \mathscr{B}, \mu, T)$ is not weak mixing. Recall Proposition 2.3.17 above: from the equivalence of the conditions that $(X, \mathscr{B}, \mu, T)$ is weak mixing and that $(X \times X, \mathscr{B} \times \mathscr{B}, \mu \times \mu, T \times T)$ is ergodic, it follows that there is a nontrivial invariant function on $X \times X$. Furthermore, there is a nonconstant almost periodic[70] function $f \in L^2(X, \mathscr{B}, \mu, T)$. This is because there is always such a function f when $T \times T$ is not ergodic, i.e., when the system is not weak mixing. Using these facts, we can always construct a non-trivial compact factor of a weak mixing system. See ([Furstenberg et al., 1982], 540) for this construction. This completes the proof of the Dichotomy Theorem. □

We are now in striking distance of proving Furstenberg Multiple Recurrence and thus Szemerédi's theorem. Our situation is the following: if we have a weak mixing system $(X, \mathscr{B}, \mu, T)$, we are done by Theorem 2.3.20. If we have a system that is not weak mixing, then we know by the Dichotomy theorem that it will possess a nontrivial compact factor. This factor then satisfies Furstenberg Multiple Recurrence by Theorem 2.3.24. Of course, what we now need is a way to "fill in" whether Furstenberg Multiple Recurrence (Theorem 2.3.12) holds for the other factors of a given system; that is, we seek some way to generalize the Dichotomy theorem.

To that end, we "relativize" the properties of compactness and weak mixing systems to the action of T on the various factors (more precisely: the action of T on the sub-σ-algebras of factors). That is, we speak of a factor X'' of a system X being either *relatively weak mixing* or *relatively compact* with respect to some other factor X' where the σ-algebra $\mathscr{B}'$ of X' is contained in the σ-algebra $\mathscr{B}''$ of X''. Ordering the factors of a system by inclusion, we show that if Furstenberg Multiple Recurrence holds for the elements in the σ-algebra of a smaller factor, then it will hold for larger factors. Finally, if there is some proper factor $\mathscr{B}'$ of the system $(X, \mathscr{B}, \mu, T)$ such that the system is not relatively weak mixing with respect to the proper factor, then there is another factor $\mathscr{B}''$ that is relatively compact with respect

[70] That is, the orbit closure of f $\overline{\{T^n f\}} \subset L^2(X, \mathscr{B}, \mu, T)$ is compact.

to $\mathscr{B}'$. Thus, analogous to the Dichotomy theorem, we can prove that factors are either relatively compact or relatively weak mixing with respect to one another. More precisely

Theorem 2.3.28. *If an arbitrary measure preserving system $(X, \mathscr{B}, \mu, T)$ is not weak mixing relative to a proper T-invariant factor of the system $\mathscr{B}'$, then there is a T-invariant factor $\mathscr{B}''$ such that $\mathscr{B}' \subsetneq \mathscr{B}'' \subset \mathscr{B}$ and $(X, \mathscr{B}'', \mu, T)$ is compact relative to $(X, \mathscr{B}', \mu, T)$.*

One might also formulate this in terms of extension maps:

Theorem 2.3.29. *Let $\pi : X \to Y$ be an extension. If $X \to Y$ is not weak mixing then there exists a non-trivial compact extension $X' \to Y$ such that $X \to X' \to Y$.*

The details of the "relativization" process are terribly germane for us, but here is a quick overview. We saw above that one way of characterizing the weak mixing property is in terms of the "product" system $X \times X$ being ergodic. We can then characterize weak mixing extensions (i.e., factors that are weak mixing with respect to another factor) by looking at a variety of relative product, called a *fiber product*.

Definition 2.3.30. (Fiber Product) Take measure preserving systems $X := (X, \mathscr{B}, \mu, T)$ and $Y := (Y, \mathscr{C}, \nu, S)$ and consider two extensions $\pi : X \to Y$ and $\pi' : X' = (X', \mathscr{B}', \mu', T') \to Y$. Then the *fiber product* of X and X' relative to Y is $X \times_Y X' = (X \times_Y X', \mathscr{B} \times_Y \mathscr{B}', \mu \times_Y \mu', T \times_Y T')$. The underlying space is $X \times_Y X'$, the set theoretic fiber product, which is defined as

$$X \times_Y X' := \{(x, x') \in X \times X' : \pi(x) = \pi'(x') \in Y\} = \bigcup_{y \in Y} \pi^{-1}(y) \times (\pi')^{-1}(y) \subset X \times X'.$$

The σ-algebra $\mathscr{B} \times_Y \mathscr{B}'$ is the restriction of $\mathscr{B} \times \mathscr{B}'$ from $X \times X'$ to $X \times_Y X'$. The measure $\mu \times_Y \mu'$ is given by its disintegration as $(\mu \times_Y \mu')_y = \mu_y \times \mu'_y$ for $y \in Y$, supported on $\pi^{-1}(y) \times (\pi')^{-1}(y)$. Finally, the action $T \times T'$ is given by $T \times T'(x, x') = (Tx, Tx')$.

Definition 2.3.31. (Relatively Weak Mixing/Weak Mixing Extension) Let $X := (X, \mathscr{B}, \mu, T)$ be a measure preserving system and let $Y := (Y, \mathscr{C}, \nu, S)$ be a factor. Then we say that X is weak mixing relative to Y, or, equivalently, that $X \to Y$ is a weak mixing extension if $X \times_Y X$ is ergodic, i.e., the fiber product of X and X relative to Y (by Proposition 2.3.17 above).

Lemma 2.3.32. *(SZ lifts through weak mixing extensions) Let $X \to Y$ be a weak mixing extension. If Y is SZ, then X is SZ.*

Equivalently, in Furstenberg's original formulation

Lemma 2.3.33. *Let $(X, \mathscr{B}, \mu, T)$ be a relatively weak mixing extension of $(Y, \mathscr{C}, \nu, S)$. If the action of S on $\mathscr{C}$ is SZ, then so is the action of T on $\mathscr{B}$, and thus both systems are SZ.*

Proof. See [Furstenberg et al., 1982], Theorem 8.4 and preceding lemmas. □

We now formulate the analogous definition and lemmas in the compact case.

Definition 2.3.34. (Relatively Compact/Compact Extension) An extension $X \to Y$ of measure preserving systems is compact if the set of almost periodic functions in $L^2(X, \mathscr{B}, \mu)$ is dense in $L^2(X, \mathscr{B}, \mu)$. A function $f \in L^2(X, \mu, T)$ is said to be almost periodic relative to factor Y if for every $\delta > 0$ there are functions $g_1, \ldots, g_n \in L^2(X, \mathscr{B}, \mu)$ such that for every $j \in \mathbb{Z}$

$$\inf_{1 \leq s \leq n} ||T^j f - g_s||_{L^2(\mu_y)} < \delta \tag{2.3.27}$$

for almost all $y \in Y$, where μ_y is the measure on the fiber $\pi^{-1}(y) \subset X$.

Lemma 2.3.35. *(SZ property lifts through compact extensions) Let $X \to Y$ be a compact extension. If Y is SZ, then X is SZ.*

Again, in Furstenberg's formulation

Lemma 2.3.36. *Let $(X, \mathscr{B}, \mu, T)$ be a compact extension of $(Y, \mathscr{C}, \nu, S)$. If the action of S on $\mathscr{C}$ is SZ, then so is the action of T on $\mathscr{B}$.*

Proof. Proving that the property SZ lifts through compact extensions is quite complicated. See [Furstenberg et al., 1982], Theorem 9.1 for the full proof. □

We are now in position to prove Theorem(s) 2.3.28 and 2.3.29. Once we have done this, we are essentially done. Let me summarize. We have seen that SZ holds for all weak mixing and compact systems and that, by the Dichotomy Theorem, any measure preserving system is either weak mixing or contains a nontrivial compact factor. We now know that the property SZ lifts through weak mixing and compact extensions. Thus, by proving Theorem(s) 2.3.28 and 2.3.29, we will "relativize" the dichotomy to extensions, and so, no matter what measure preserving system presents itself, we know that it will be SZ! Consider the tower of factors, descending from the "full" system X:

$$X \to Y_n \to \cdots \to Y_2 \to Y_1 \to Y_0 \tag{2.3.28}$$

If X is weak mixing, we are done, since weak mixing systems are SZ. If not, then it has a nontrivial compact factor, say Y_0, which is also SZ. But we then know that for each compact $Y_n \to Y_{n-1}$, the property SZ at the bottom of the tower lifts. If at some point we reach $X \to Y_n$ which is not compact, then by the relativized dichotomy theorem, it must be a weak mixing extension. But Y_n is SZ, since Y_{n-1} was and SZ lifts through weak mixing extensions too. Thus, X will also be SZ.

What if we never reach this stopping point? This too does not pose a problem, since we can keep extending the intermediate factors of X into the transfinite,[71] and so, again, X will be SZ. Thus, we finally have

Theorem 2.3.37 (Furstenberg-Zimmer Structure Theorem)**.** *Let $(X, \mathscr{B}, \mu, T)$ be any measure preserving system. Then there is an ordinal α and a transfinite increasing sequence of factors $(\mathscr{B}_\beta)_{\beta \leq \alpha}$ such that:*

[71] Indeed, it has been explicitly shown that the last (commonly called "maximal distal") factor in the tower can extend arbitrarily far into the countable ordinals. See [Beleznay and Foreman, 1996].

1. *$\mathscr{B}_0$ is trivial (i.e., $\{\emptyset, X\}$);*
2. *For each successor ordinal $\beta+1 < \alpha$, $(X, \mathscr{B}_{\beta+1}, \mu, T)$ is compact relative to $(X, \mathscr{B}_\beta, \mu, T)$;*
3. *For each limit ordinal $\lambda \leq \alpha$, $\mathscr{B}_\lambda = \bigcup_{\beta<\lambda} \mathscr{B}_\beta$;*
4. *$(X, \mathscr{B}, \mu, T)$ is weak mixing relative to $(X, \mathscr{B}_\alpha, \mu, T)$.*

Finally, since we deal with limit ordinals in the Structure Theorem, we require that SZ is preserved through limiting stages. To that end:

Proposition 2.3.38. *Let $(\mathscr{B})_{\beta\in B}$ be a totally ordered chain of factors of arbitrary measure preserving system X, and let X be the limit of $(\mathscr{B})_{\beta\in B}$. Then, provided each $\mathscr{B}_\beta$ is SZ, X is also.*

This completes the ergodic proof of Szemerédi's Theorem.

2.3.4 Summary

We began with Szemerédi's theorem on the integers: any sufficiently dense $A \subset \mathbb{Z}$ will contain arbitrarily long arithmetic progressions (Theorem 2.3.4). This fact will always turn upon the classification of A. In particular, it is easy to see why A has arithmetic progressions when A is either highly structured or highly random. The difficulty is that A may not be either, and so any proof of the theorem requires checking the existence of arithmetic progressions for each A. One might then appeal to the Szemerédi Regularity Lemma (Lemma(s) 2.5.3 and A.1) in order to provide a structural result that classifies all subsets under consideration; however, this lemma does not show in a perspicuous way how such a classification occurs (see my remarks in Section 2.5).

Turn then to the ergodic approach. Szemerédi's theorem is equivalent to Furstenberg Multiple Recurrence (Theorem 2.3.12), a theorem which asserts the existence of a particular recurrence pattern in any measure preserving system. We are then in a situation entirely like the number-theoretic one above: it is easy to show that highly random (mixing, weak mixing) and highly structured (compact) systems exhibit this recurrence pattern (Theorem(s) 2.3.20 and 2.3.24, respectively). However, there are many other ways a system might be, and so we seem resigned to scattered results about particular systems.

Luckily, any measure preserving system can be decomposed into structured and random pieces we know how to handle. The Koopman-von Neumann Dichotomy Theorem asserts that a system is either weak mixing or has a nontrivial compact factor (Theorem 2.3.27). This dichotomy can then be relativized to maps between factors of a system, i.e., each extension of a factor will be either weak mixing or compact (Theorem(s) 2.3.28 and 2.3.29). Therefore, there exists a tower of extensions

$$X \to Y_\alpha \to \ldots \to Y_2 \to Y_1 \to Y_0, \tag{2.3.29}$$

indexed by ordinals, where each step is either weak mixing or compact (Furstenberg Structure Theorem; Theorem 2.3.37). We also know that the recurrence property we desire lifts through weak mixing extensions, compact extensions, and limits (Lemma 2.3.32 and Lemma 2.3.35

and Proposition 2.3.38). Thus, via transfinite induction, we can show the recurrence property lifts all the way to X. X was an arbitrarily chosen measure preserving system, and so Furstenberg Multiple Recurrence holds (Theorem 2.3.12), and thus Szemerédi's theorem (Theorem 2.3.4) holds.

Before entering into the strictly philosophical analysis, I would like to note that, in a very straightforward sense, the ergodic proof of Szemerédi's theorem is explanatory. In short, if one were to ask, "Why does that particular pattern occur in measure preserving systems," this proof gives an entirely perspicuous means of answering the question, i.e., "Because (i) any such system can be analyzed exhaustively into two cases and (ii) we know how to analyze each case." Taking this unobjectionable, but perhaps not terribly helpful account of explanation on board, let us now see what this case of mathematical explanation involves. First, we can certainly say that explanatory work is being done by impure techniques; after analyzing Szemerédi's Regularity Lemma and its role in the combinatorial proof I would like to claim further that the ergodic setting is the proper one for providing an explanation.

2.4 Conceptual Convergences and Mathematical Content

Before demonstrating how impurity leads to explanation, let me turn to an investigation of mathematical content. I proceed in this fashion because the latter will help to deepen our understanding of the former; in particular, a sufficiently fine-grained analysis of the content of a theorem will help us to see how it is possible that impure techniques do not necessarily result in a loss of data required for an explanatory proof. Such an investigation will also yield philosophical dividends of its own.

The Furstenberg Correspondence Principle (Theorem 2.3.14) and the ergodic proof of Szemerédi's theorem immediately raise questions about mathematical content:[72] what, precisely, is Szemerédi's theorem *about*? What does it mean and how are we to understand it?[73] How do we account for such surprising convergences of mathematical domains, which, as Rota claims,[74] constitute the "essence" of mathematics? One straightforward—though perhaps unintuitive—answer might involve saying that particular number-theoretic facts,

[72] See [Martin, 1998] for a brief, but helpful, discussion of the relationship between mathematical evidence and the content of mathematical propositions.

[73] Wittgenstein raises this question in his *Remarks on the Foundations of Mathematics*. I suppose I am in broad agreement with his claim that: "[T]he understanding of a mathematical proposition is not guaranteed by its verbal form, as is the case with most non-mathematical propositions" ([Wittgenstein, 1978], 147). Unfortunately, the way in which he elaborates on this is hardly serviceable. He proposes that the sense of a mathematical statement is gotten from its proof. But how can this be? How can one even set out to prove a statement that has no initial sense? This embroils us in a sort of Meno problem, the only escape from which requires that we know *something* initially. I propose that this will be the "intuitive content" of a mathematical statement. See my discussion of [Arana and Mancosu, 2012] below for an explication of intuitive content.

[74] As he says, "The mystery as well as the glory of mathematics lies not so much in the fact that abstract theories turn out to be useful in solving problems but in the fact that—wonder of wonders—a theory meant for one type of problem is often the only way of solving problems of entirely different kinds, problems for which the theory was not intended. These coincidences occur so frequently that they must belong to the essence of mathematics. No philosophy of mathematics shall be excused from explaining such occurrences" ([Rota, 1997], 114).

viz., Szemerédi's theorem, are ergodic in nature.[75] That is, one might suppose that facts about integers and natural numbers have "hidden" or "implicit" infinitary and ergodic content. This kind of suggestion has an impressive pedigree. Bourbaki claims something in this vicinity,[76] though I take the ultimate point to be more subtle:

> Where the superficial observer sees only two, or several, quite distinct theories, lending one another 'unexpected support' through the intervention of a mathematician of genius, the axiomatic method teaches us to look for the deep-lying reasons for such a discovery, to find the common ideas of these theories, buried under the accumulation of details properly belonging to each of them, to bring these ideas forward and out them in their proper light ([Bourbaki, 1950], 223).

Should we then accept our intuitive, "first glance" characterization of Szemerédi's theorem as number-theoretic, properly belonging to a domain of mathematics genuinely distinct from ergodic theory, or, once the various proofs of the theorem have come to light, should we expect that such intuitive boundaries are not the final word on the matter?

This question might seem more pressing in cases where the theorem under consideration *only* has, as of yet, a proof that utilizes intuitively "foreign" or "extraneous" elements (the classic example being Fermat's Last Theorem[77]) or in cases where a sentence is unprovable in the system in which it is stated (e.g., Gödel sentences and the Paris-Harrington sentence are unprovable in PA even though they are truths expressible in $\mathsf{L_{PA}}$). In order to prove these, a detour through the infinite and the impure is unavoidable. However, Bourbaki had less exotic cases[78] in mind, and the question of content remains an interesting one even when we have both impure and pure proofs at our disposal (as with Szemerédi's theorem). This is because, at the very least, the detour through the impure is left quite unaccounted for: it occurs *in spite of* the fact that a pure proof is available, thereby suggesting there is more to the theorem than initially meets the eye. Thus, I would like to outline two proposals:

1. There is a notion of mathematical content that will help to account for *prima facie* surprising convergences of mathematical domains.

2. This new notion of content will help us to understand the relationship between impure techniques of proof and explanation.

Let us begin, then, by characterizing the notion of mathematical content in greater detail.

2.4.1 Intuitive and Formal Mathematical Content

A natural starting point is the analysis of mathematical content by Arana and Mancosu in [Arana and Mancosu, 2012]. Here they draw an important distinction between "informal"

[75] Or, if considering van der Waerden's theorem, topological in nature.

[76] In [Detlefsen and Arana, 2011], Detlefsen and Arana interpret the above quote (along with McLarty) as claiming that, e.g., particular arithmetical statements have topological content. Incidentally, the discussion of [Detlefsen and Arana, 2011] involves another result due to Furstenberg: his topological proof of the infinitude of primes.

[77] See my brief discussion of this in Section 2.6.

[78] Indeed, in [Bourbaki, 1950], the focus is on very central mathematical concepts (groups, topologies). One might think that Gödel sentences and the like are rather more pathological.

or "intuitive" content and "formal" or "axiomatic" content. Loosely, the former notion amounts to what someone with a casual acquaintance with mathematics would understand by some statement. For instance, Szemerédi's theorem is about a particular kind of regularity in sufficiently dense subsets of the integers (infinitary formulation) or in sufficiently large subsets of the natural numbers (finitary formulation). On the other hand, the formal notion of content amounts to the "inferential role of [a] statement within an axiomatic system" ([Arana and Mancosu, 2012], 327). This distinction is drawn in the context of their discussion of Desargues' theorem. One of the primary questions under investigation is whether the spatial proof of this (ostensibly) planar result is to be judged as impure (following Hilbert) or pure (following Michael Hallett) in light of metamathematical data from Hilbert's *Grundlagen der Geometrie*. In particular, Arana and Mancosu are concerned with articulating a notion of mathematical content that "...can support an adequate account of talk of purity in mathematical practice" ([Arana and Mancosu, 2012], 324). They conclude that only intuitive content is able do this work. I take no issue with this conclusion; indeed, my assertion that the ergodic proof of Szemerédi's theorem is impure relies upon the availability and cogency of intuitive content. However, it is worth considering whether we might utilize other, more nuanced, articulations of mathematical content in cases that warrant them. The theorem considered here is one such case: intuitive content cannot help us make sense of why it is that the ergodic setting is adept at modeling (and even clarifying) the combinatorial proof of Szemerédi's theorem. I outline a third kind of content below that I believe is up to the task; however, before turning to this positive proposal, let me explicate and offer some reflections on the less familiar notion of formal content.

Arana and Mancosu consider Michael Hallett's claim[79] that the spatial proof of the planar Desargues' theorem is pure because the theorem itself has "tacit spatial content." This notion of "tacit" or "hidden" content descends from Isaacson's influential discussion of "hidden higher-order concepts" in [Isaacson, 1996]. The main thesis of this essay is that some mathematical truths expressible in arithmetic, e.g., Gödel sentences, contain "hidden higher-order concepts," and thus first-order Peano arithmetic (PA) is actually complete with respect to *genuinely arithmetical* sentences. By higher-order,[80] Isaacson means:

> [T]he standard usage for quantification over sets of individuals in distinction to first-order quantification over the individuals themselves. But I also mean to include in this phrase something of the notion of the infinitary, in the sense of presupposing an infinite totality [...] ([Isaacson, 1996], 210).

According to Isaacson, these higher-order concepts are implicit in arithmetical propositions via the coding of various syntactic properties and relations by properties and relations of the natural numbers (as in Gödel's proof of the Incompleteness Theorems). Put more precisely, a Gödel sentence G can be shown in PA to be equivalent to a sentence expressing (by coding) a metamathematical property of PA, e.g., Con(PA). This metamathematical property is unprovable in PA, but can be proven using higher-order (in the sense above) concepts. Thus, the equivalence reveals the implicit higher-order content of G.

[79] See [Hallett, 2008].

[80] See the beginning of Section 2.2 for my explication of the the finitary-infinitary distinction which maps quite closely on to Isaacson's own.

Isaacson does not say anything further about the exact nature of the implicit higher-order content of G, but presumably this is gotten from the arithmetical coding of transfinite ordinals required for the proof of Con(PA). In particular, we know that a finitistically acceptable theory (e.g., PRA) augmented with a principle expressing (by coding) transfinite induction along a well-ordering of the ordinals up to a particular ordinal, ϵ_0, proves Con(PA).[81] This principle must, then, provide the content of Con(PA) (and thus of G) that is higher-order and non-arithmetical; however, the exact nature of this content is not so straightforward given that, for example, Gentzen's proof of Con(PA) does not *explicitly* rely upon transfinite ordinals. Rather, it relies upon induction along a well-ordering of *ordinal notations* (coded by natural numbers), and this well-ordering is isomorphic to the well-ordering of transfinite ordinals up to ϵ_0. Nonetheless, the idea seems to be that the use of this additional machinery in developing ordinal notations suffices to impart infinitary higher-order content to G. This is because of Isaacson's understanding of what counts as a genuinely arithmetical truth:

> [A] truth expressed in the (first-order) language of arithmetic is arithmetical just in case [(i)] its truth is directly perceivable on the basis of our (higher-order) articulation of our grasp of the structure of the natural numbers *or* [(ii)] directly perceivable from truths in the language of arithmetic which are themselves arithmetical ([Isaacson, 1996], 217).

Thus, for Isaacson, there is an innocuous sort of higher-order content (given by (i)), the use of which does not render truths expressible in $\mathsf{L_{PA}}$ non-arithmetical. For instance, if we come to perceive the truth of some φ expressible in $\mathsf{L_{PA}}$ from our understanding of Dedekind's second-order, categorical characterization of the natural numbers, then φ is still arithmetical. On the other hand, the Gödel sentence G, although expressible in $\mathsf{L_{PA}}$, is rendered non-arithmetical because: (a) it is equivalent to a sentence ψ expressible in $\mathsf{L_{PA}}$ that encodes Con(PA); (b) the proof of Con(PA) requires induction on the well-ordering of ordinal notations for ordinals $< \epsilon_0$. The question of G's higher-order content thus reduces to that of whether this inductive principle can be justified arithmetically. Isaacson claims that it is "reasonably evident" that G is not arithmetical,[82] and so the answer to this question is "no." I am inclined to agree with his analysis; however, I am not sure that the reasons supporting it are so straightforward. First, we would need to show that there is no arithmetical means by which the well-foundedness of the ordering of the codes of the ordinals $< \epsilon_0$ could be understood. I find it plausible that this could be arithmetically justified for sufficiently small transfinite ordinals, perhaps, e.g., $\omega + \omega, \omega + \omega + \omega, \ldots, \omega \cdot \omega$. However, for the purposes of obtaining a proof of Con(PA), we must be able to perceive arithmetically the truth of induction along the well-ordering of codes of ordinals *up to* ϵ_0, and this is a much more imposing requirement.[83] The nature of the structure of this well-ordering seems sufficiently complicated to render it higher-order and thus non-arithmetical. I take it that something along these lines is what Isaacson means by the implicit higher-order content of G.

[81] The ordinal ϵ_0 is the limit ordinal of the increasing sequence $\omega < \omega^\omega < \omega^{\omega^\omega} < \cdots$. Also, note that the ordinals $< \epsilon_0$ must be written in a particular way, i.e., Cantor normal form.

[82] At least, in the sense of (i), since it is possible that new proofs will emerge that proceed from "recognizably arithmetical" truths.

[83] It is clear that the innocuous higher-order content (Dedekind's categorical characterization of the natural numbers) cannot justify the resources needed to perceive the truth of Con(PA). That is, there is no way to proceed from Dedekind's characterization to $\mathsf{TI_{PR}}(\epsilon_0)$.

According to Arana and Mancosu, the case of Desargues' theorem is analogous to that of the Gödel sentence because it can serve as a spatial incidence axiom in Hilbert's axiomatic system of the *Grundlagen*, i.e., the planar Desargues' theorem can play the same inferential role as an explicitly spatial mathematical proposition.[84] They summarize the situation in the following way:

> [I]n both cases, we have sentences whose ordinary understanding indicates that it has content of one type (e.g., arithmetical or planar), but that an analysis of the sentences' inferential roles reveals that these sentences have tacit content of another type (e.g., infinitary or spatial, respectively) ([Arana and Mancosu, 2012], 333).

Our question, then, is: can Szemerédi's theorem be understood in the same way?

There are some similarities between our case study and the phenomenon of hidden content at work in Gödel sentences and Desargues' theorem. Our ordinary or intuitive understanding of Szemerédi's theorem indicates that it has finitary, combinatorial content; however, it is equivalent to an ostensibly infinitary result (Furstenberg Multiple Recurrence), and its ergodic proof makes essential use of a transfinite construction (the tower of extensions in the Furstenberg Structure Theorem). This initial assessment of the situation suggests that the notion of formal content may be of use in explicating this surprising confluence of mathematical resources. However, closer examination reveals that formal content is not really available to us. Let us write the formal content of the Gödel sentence G as:

$$\mathscr{F}_G := \{\psi : \vdash_{\mathsf{PA}} G \leftrightarrow \psi\}\,. \tag{2.4.1}$$

That is, the formal content of G is the class of statements such that each statement is provably equivalent to G in PA. Thus, as it should be, some ψ expressing by coding $\mathrm{Con}(\mathsf{PA})$ is included in the formal content of G. However, when we think of formal content in the above fashion, there is a rather restricted set of circumstances in which this will make sense, viz., when we have an independence result. In the case of G, while we can prove $G \leftrightarrow \psi$ in PA, we cannot prove the property that ψ expresses, i.e., $\mathrm{Con}(\mathsf{PA})$, in PA by Gödel's Second Incompleteness Theorem. Thus, we have the requisite independence. In the case of Szemerédi's theorem, we could try to write its formal content as:

$$\mathscr{F}_{\mathsf{SZ}} := \{\psi : \vdash_{\mathsf{ZFC}} \mathsf{SZ} \leftrightarrow \psi\}\,. \tag{2.4.2}$$

Letting some $\psi :=$ Furstenberg Multiple Recurrence we would register the "hidden" infinitary ergodic content of Szemerédi's theorem. However, Furstenberg Multiple Recurrence is not independent of ZFC, the system in which we prove the equivalence. This has the effect of saying that Szemerédi's theorem and Furstenberg Multiple Recurrence have the *same* formal content; indeed, all theorems of ZFC would then have the same formal content, rendering this notion entirely trivial.

Another way of seeing the unavailability of formal content in our case is via the absence of a coding phenomenon. According to Isaacson,"The relationship of coding constitutes a rigid

[84] See [Arana and Mancosu, 2012], Section 4 for more details.

link between the arithmetical and the higher-order truths, which pulls the ostensibly arithmetical truth up into the higher-order" ([Isaacson, 1996], 221). There is no such mechanism at work in the Furstenberg Correspondence Principle.[85] Rather, as I have shown above, we have an identification of sets and functions in the combinatorial and ergodic contexts, along with a clever choice of "unconventional" dynamical system that effects this identification.[86] This should be thought of as a "modeling" of the combinatorial phenomenon in the ergodic setting, rather than indicating the presence of coded or "hidden" content.[87]

Of course, these considerations need not rule out the possibility of using formal content to explicate Furstenberg Correspondence in the future. This would depend upon a reverse mathematical analysis of both Szemerédi's theorem and Furstenberg Multiple Recurrence and the existence of an independence result. That is, we would need to show that some system T proves Szemerédi's theorem and its equivalence with Furstenberg Multiple Correspondence but not Multiple Correspondence itself. It is very likely that Szemerédi's theorem can be proved in RCA_0 and perhaps even weaker systems, although this would require some rather tedious work to verify.[88] Furthermore, the metamathematical analysis of the infinitary content of Furstenberg Correspondence in [Avigad, 2009], Section 5, indicates that it is reinterpretable in computational or combinatorial terms. Thus, given this metamathematical data, it would appear that the independence result required for formal content to make sense is unlikely to materialize.

It should be noted that, even though the Correspondence Stage of the ergodic proof may be computationally reinterpretable, this is not true of the Structure Theorem Stage of the proof. Indeed, significant infinitary and nonconstructive content enters the scene when one builds extensions $X \to Y_\alpha \to \cdots \to Y_1 \to Y$. This is because Theorem 2.3.27 (The Dichotomy Theorem of Koopman and von Neumann) and its relativized version, Theorem 2.3.28, rely upon understanding the limiting behavior of various dynamical systems, and it turns out that this limiting behavior is in general not computable.[89] The degree of nonconstructivity then increases given the transfinite iteration in the Furstenberg Structure Theorem (Theorem 2.3.37). This is quite interesting in its own right, since I claim that this part of the proof generates many of the explanatory advantages of the ergodic setting. And this is because the use of quite strong infinitary techniques most cleanly and explicitly draws out the structural feature that gives the reason for the truth of Szemerédi's theorem. In short, the most radically nonconstructive, infinitary content is found in the stage of proof that does the explaining. It would be an interesting exercise to see if such an analysis can be given for other results.[90]

Finally, it is worth considering whether a looser notion of hidden higher-order content

[85] Arana and Mancosu note that there is no *explicit* coding in the case of Desargues' theorem either, but surmise that it is given by the "...algebra of segments permitted by Desargues' theorem in the presence of axioms I 1-2, II, and III, or more directly the alleged spatial content it inherits from its stereometrical proof" ([Arana and Mancosu, 2012], 333).

[86] See my brief discussion after Theorem 2.3.12 above.

[87] One might also say that we adopt a new sort of perspective on the combinatorial results by considering them *qua* ergodic (in a sense to be made precise). I return to this idea in the final section when I discuss the Szemerédi Regularity Lemma (Theorem 2.5.3) and the Furstenberg Structure Theorem (Theorem 2.3.37).

[88] See my discussion in Section 2.5.

[89] See [Avigad et al., 2010] and [Avigad and Simic, 2006].

[90] See my remarks in Section C.

might be of interest. This would then no longer involve Arana and Mancosu's gloss of hidden higher-order content as *formal content*, and thus the need for an independence result would no longer be present. Isaacson himself suggests something along these lines when dealing with "in principle" provability. That is, there might be cases where a theorem is provable in PA, but this proof would be "infeasibly long" and thus "the higher-order perspective is essential for *actual* conviction as to the truth of an arithmetically expressed sentence" ([Isaacson, 1996], 221). I am rather sympathetic to the idea that various statements in some L_T could be said to have higher-order content in light of the unsurveyability of their proofs[91] in the system T. Indeed, this kind of thought is closely related to my claims in Section 2.6 that, even if infinitary resources are not proof-theoretically *required*, they are still in some sense necessary for intelligible mathematics (or, in Isaacson's words, necessary for "actual conviction"). It is implausible that Szemerédi's theorem has hidden higher-order content in this sense: its proof is very *difficult*, but by no means unsurveyable. However, the other example considered in Section 2.6 is Fermat's Last Theorem (FLT), and this seems a good candidate for an ascription of "loose" higher-order content. Few seem to doubt that it is "in principle" provable in PA; however, actually producing and understanding such a proof appears to be nearly impossible. It is then reasonable to ascribe *some sort* of hidden higher-order content to FLT.

2.4.2 A Further Refinement of Mathematical Content

Thus, though the notion of formal content (and looser versions of it) may be useful in particular circumstances, it cannot help us to understand the case at hand. Should we then evaluate Szemerédi's theorem in terms of intuitive content alone? Unfortunately, as I have indicated, this coarse articulation of content fails to capture crucial data about the theorem, data that would be desirable to have for philosophical discussions relating content, purity, explanation, mathematical evidence, and likely other topics. In particular, we should like to say something about the relationship between the combinatorial and ergodic settings in light of both the equivalence between Szemerédi's theorem and Furstenberg Multiple Correspondence and the epistemic dividends generated by the ergodic proof. I find it quite reasonable that this connection be captured in our understanding of the content of Szemerédi's theorem. In short, then, intuitive content is not sufficiently fine-grained to capture mathematical information of interest, while formal content is unavailable to us and, even if it were, is neglectful of important epistemic distinctions.[92] Is there then another, "intermediate" notion of content that might make sense of the ergodic proof of Szemerédi's theorem? I believe that there is and attempt to use the somewhat sketchy suggestions of Bourbaki to develop this

[91] Isaacson notes that this looser notion of higher-order content would not affect his thesis that PA is complete with respect to genuinely arithmetical truths because any expansion in the class of statements possessing higher-order content would be accompanied by a narrowing of the domain of the arithmetically true.

[92] For instance, Arana and Mancosu reject formal content as useful in making purity ascriptions because it has the following consequences: (i) The very question of content cannot be articulated using only formal content (as I have noted at the outset); (ii) Someone without beliefs concerning space could not understand Desargues' theorem; (iii) Purity ascriptions would be entirely trivial as every theorem would have a pure proof; (iv) The content of statements would be radically contextualized, e.g., it would depend entirely upon the axiomatic context.

idea. Call this third sort of content *structural content*[93] to mirror the "deep-lying reasons" indicated in the above quotation by Bourbaki.

Bourbaki describes the "axiomatic method" as a "systematic study of the relations existing between different mathematical theories" which takes as its central concept that of the mathematical structure ([Bourbaki, 1950], 222). Consider Bourbaki's description of a mathematical structure:

> The common character of the different concepts designated by this generic name [mathematical structure], is that they can be applied to sets of elements whose nature has not been specified; to define a structure, one takes as given one or several relations, into which the elements enter [...]; then one postulates that the given relation, or relations, satisfy certain conditions (which are explicitly stated and which are the axioms of the structure under consideration) (*ibid.*, 226).

I propose that the structural content of a theorem is the instantiation of a particular fundamental mathematical structure by the entities intuitively involved in the theorem. Let me fix intuitions with a simple example: the realization that $(\mathbb{R}, +)$ is a particular instance of a group structure[94] contributes additional content to a theorem involving addition on the reals. This content cannot be what we are calling intuitive: it seems doubtful that a mathematical novice proving a theorem about $(\mathbb{R}, +)$ that is not explicitly group-theoretic would grasp the group axioms. Nor would a theorem involving $(\mathbb{R}, +)$ gain a contribution of strictly formal content from the realization that $(\mathbb{R}, +)$ is the instantiation of a fundamental structure; the understanding that $(\mathbb{R}, +)$ is a group does not involve seeing that it can *serve* as an axiom in a system, but rather the understanding that $(\mathbb{R}, +)$ *satisfies* particular axioms. Furthermore, this realization will induce a relational fact about the particular entity in question, $(\mathbb{R}, +)$, and other groups, i.e., the fact that they all instantiate the group structure. This fact contributes content that is of a much more global character than intuitive content, while not being merely inferential. We might then fruitfully compare various instances of group structures once we realize that they can all be identified as such.

Bourbaki is careful to note that this description of the mathematical universe via structures is "...schematic, and idealized as well as frozen" and is to be understood as a "supple and fertile research instrument" (*ibid.*, 229; 231). As such, it is open to revision, and, in particular "[the] definition of structures is not sufficiently general" (*ibid.*, 226, fn). I want to suggest that we can proceed by taking a mathematical structure to be a somewhat more flexible notion than a concept defined by explicit axioms. Of course, once we broaden the notion of mathematical structure to include more than explicit axiomatizations, some further attempt at characterization must be made. A natural suggestion is that we appeal to compositional facts about mathematical objects: for some object X in (conceptual) domain D write

$$X \text{ can be identified as } Y \sim Z \tag{2.4.3}$$

[93] I had also proposed the rather more clumsy, "Bourbakiste content," since the word "structure" has so many associations. However, I opted for the more elegant term in the end.

[94] Where this structure is defined by the group axioms: closure, associativity, existence of identity and inverse.

for more restricted objects Y, Z in the same domain under appropriate relation $\sim$ (this may be a cross product, direct product, set-theoretic union, etc.). It is not difficult to find such results, especially in algebraic contexts. Indeed, there are vast swathes of mathematics concerned with "structure theorems": structure theory of countable Abelian groups, structure theory of semisimple Lie algebras, etc. I mention these other examples to convince my reader that the compositional facts I advert to are deeply embedded in mathematical practice, and, as such, provide a reasonable starting point for my intermediate notion of content.[95]

Strikingly, these compositional facts are not restricted to algebraic contexts. We have seen that subsets of integers of positive density can be decomposed in such a fashion by appeal to a Structure Theorem, which expresses the dichotomy between structure and randomness.[96] Thus, according to my content proposal, Szemerédi's theorem, simply by virtue of the fact that it involves sufficiently dense subsets of $\mathbb{Z}$, has as part of its content the dichotomy between structure and randomness. Even though the theorem is not obviously about this dichotomy, as an intuitive reading or the "verbal form" of the mathematical proposition[97] would suggest, I believe we should consider the instantiation of this dichotomy to be included in its content. However, and here is the crucial point, this structural fact about how the objects in question may be decomposed need not be (and is not) endemic to the integers or the naturals alone. One can similarly see the dichotomy at work in the ergodic setting: measure-preserving systems also exhibit this structural fact as expressed by the Furstenberg Structure Theorem. We might think of the different mathematical objects in question (sets of integers, measure-preserving systems) as multiple realizations of this same high-level structural fact.[98] Thus, by appealing to structural content, we allow that intuitively different mathematical domains retain some degree of conceptual independence, facilitating various practical distinctions like purity ascriptions, but gain a principled way to talk about surprising confluences of such domains. Both the ergodic and number-theoretic settings exhibit the structural dichotomy in question, i.e., they both encode the same sort of information (the confluence), but each context obviously involves different objects that serve to made the dichotomy precise in different ways (the independence).

A little more explicitly: the proofs of Szemerédi's theorem (T_C; Theorem 2.3.4) and Furstenberg Multiple Recurrence (T_E; Theorem 2.3.12) rely on their respective Structure Theorems, the Szemerédi Regularity Lemma (S_C; Theorem 2.5.3), and the Furstenberg-Zimmer Structure Theorem (S_E; Theorem 2.3.37). Thus, we have a class of theorems $\mathscr{T}$, of which T_C, T_E are members, defined by the fact that all theorems in the class rely on Structure Theorems (theorems showing how a particular entity can be decomposed into a structured

[95] I do not claim such structural results function in the same way as the dichotomy between structure and randomness in Szemerédi's theorem. Such a claim would inevitably involve a careful analysis of particular algebraic structure theorems. It is worth noting in this connection that a particular result in the structure theory of Abelian groups (Ulm's theorem) has been shown to require the strength of Π^1_1-CA_0. See [Simpson, 1999] for details. It is possible, pending further mathematical research, that the ergodic proof of Szemerédi's theorem in which the full strength of the Furstenberg Structure Theorem is used, will also require Π^1_1-CA_0. See Appendix B for discussion.

[96] See Tao's quote in Section 2.3.2.

[97] See Wittgenstein's remark in fn. 73.

[98] This idea is common in the philosophy of science literature. Multiple realizability is, more or less, the idea that there can be heterogeneous "realizers" of "upper-level" properties and generalizations. That is, these generalizations characterize the same behavior in physically distinct systems.

and random part) for their proof. The presence of these Structure Theorems in these proofs then indicates that there is a general structural fact in the offing, i.e., the dichotomy between structure and randomness. And so, we have found an incredibly important higher level property of the theorems in $\mathscr{T}$, a property that is not captured by the intuitive content of the theorem. Crucially, the fact that $T_C, T_E \in \mathscr{T}$ allows us to make sense of the surprising intervention of the Furstenberg Structure Theorem in the proof of Szemerédi's theorem and makes the ergodic context suitable to prove a number theoretic result. This is because the objects manipulated in each context (subsets of integers and measure-preserving systems) can be decomposed into structured and random components.[99] The fundamental dichotomy between structure and randomness allows us to make sense of the confluence of these *prima facie* very different domains, and, as we shall see, its clear expression in the ergodic context renders the ergodic proof of Szemerédi's theorem explanatory.

Before closing this section, let me respond to a natural objection. One might think that in offering this notion of structural content I have tried to have my cake and eat it too. I have tried to provide a notion of content that retains intuitive ascriptions of purity/impurity while also explicating the intervention of impure resources. This has been done, more or less, by carving out a genus of mathematical results, a sort of mathematical natural kind if you will, under which both Szemerédi's theorem and Furstenberg Multiple Recurrence[100] fall. However, one might think that this then requires saying that the ergodic proof of Szemerédi's theorem is pure. This objection would be analogous to that leveled by Arana and Mancosu against proponents of formal content:[101] privileging formal content "...threatens to trivialize purity, in making it the case that *every* theorem has a pure proof, when the content of that theorem is fully understood" ([Arana and Mancosu, 2012], 336). Similarly, one might say that every theorem has a pure proof once investigators have dug "deep enough" and uncovered the requisite structural similarities. Another way of putting this point would be to appeal to the topical content of a theorem and observe the purity metric it induces. For example, under my construal of structural content, both Szemerédi's theorem and Furstenberg Multiple Recurrence are *topically close* insofar as both subsets of integers and measure-preserving systems instantiate the dichotomy between structure and randomness. This topical closeness then indicates that the ergodic resources used to prove Szemerédi's theorem do not involve an appeal to impure techniques.

In response, I am inclined to restrict the ways in which the topical content of a theorem induces a purity constraint.[102] If we are interested in philosophically explicating mathematics as practiced, then the purity constraints we delineate should map closely onto the activity

[99] Of course, here we do not have axioms defining a structure which particular entities instantiate. Nonetheless, at least in the case under consideration, both "structure" and "randomness" can be made perfectly precise in each context, viz., combinatorial and ergodic.

[100] As well as the analogous results in Fourier analysis and hypergraph theory and special cases of all these theorems. Thus, this is quite a robust genus.

[101] This easily translates into Aristotelian terms as well. We have not in fact "crossed genera" when we utilize structural content. See fn. 17 above.

[102] Thus, my claim concerning purity in the Introduction should be modified; the distance metric given by the topical content of a theorem does not necessarily generate a purity constraint in a "straightforward" way. One might, on the other hand, simply say that structural content is a distinct notion of content from the topical kind, but this seems quite counter-intuitive to me. I believe it is correct to say that both Szemerédi's theorem and Furstenberg Multiple Recurrence are, in some sense, *about* the dichotomy in question.

of mathematicians. There is a good deal of evidence that, when speaking of a "pure" proof, mathematicians are operating with an intuitive notion of content. Thus, intuitive content is what philosophers should appeal to in making claims about the purity or impurity of proofs. Nonetheless, I believe the arguments provided above support the existence of structural content. Structural content is intelligibly part of what a theorem is about precisely because the entities *intuitively* involved in the theorem instantiate particular structural features. However, we should block the move from the presence of structural content to a structural purity constraint because the notion of purity generated is not acknowledged by the mathematical community and, like formal content, would trivialize ascriptions of purity made in practice. A structural purity constraint would simply be an idle postulation.

Furthermore, from a strictly conceptual perspective (independent of a methodological interest to generate philosophical theses consonant with mathematical practice), structural content enjoys a crucial advantage over formal content: it is able to coexist with intuitive content. If both notions were not simultaneously available to us, then we would be in no better a position than the proponent of formal content. Structural content, however, keeps the original mathematical concepts in view, whereas formal content ignores these concepts altogether by privileging the axiomatic role of a statement as the determinant of the statements's content. Indeed, the ascription of structural content *requires* that we recognize the intuitive content of a theorem, for it is the entities intuitively recognized that instantiate the structural features of interest. We must initially acknowledge that two domains of mathematics are ostensibly unrelated but interact in surprising ways. Only then can we acknowledge and begin to uncover the "deep-lying reasons" for this interaction, reasons which are not about only one domain or the other but rather encompass both. Thus, we can retain a purity constraint generated by intuitive content: ergodic theory and combinatorics involve intuitively different concepts and movement between these domains should be considered an instance of impurity. However, these domains fall within a larger structural genus, which elucidates an important commonality between them without obliterating their intuitive differences (as is done by formal content). We can then comprehend the possibility of these domains interacting in an intelligible way given their membership in a common genus.

2.4.3 Summary

Let me summarize the dialectic thus far. The equivalence of Szemerédi's theorem and Furstenberg Multiple Recurrence is surprising and may suggest that Szemerédi's theorem has hidden higher-order content: in this case, hidden infinitary and ergodic content. That is, in some sense, Szemerédi's theorem is about both subsets of integers and infinitary measure-preserving systems. In order to better understand this, I examined Arana and Mancosu's discussion of intuitive and formal content. We saw that the notion of formal content operative in other discussions of hidden higher-order content is unavailable to us and agreed with the argument of [Arana and Mancosu, 2012] that intuitive content is essential for making purity ascriptions. Nonetheless, intuitive content fails to explicate the relationship between combinatorics and ergodic theory in the ergodic proof of Szemerédi's theorem.

Thus, I have tried to excavate a third kind of mathematical content: structural content. The ultimate point of this third kind is that it occupies the middle ground between intuitive and formal. Intuitive content is important and useful in mathematical practice, yet it does

not help us to make sense of surprising convergences of mathematical results; formal content is useful only in very restricted contexts and may trivialize important epistemic distinctions. Structural content, on the other hand, is more fine-grained than intuitive content but still occurs within mathematical practice. I have described it as the instantiation of particular structural facts by intuitively recognized entities that induce higher-level dependencies between theorems, linking them in a conceptual nexus that crosses over intuitively distinct domains. Of course, unrestricted appeal to *any* property shared by various theorems (e.g., all theorems involving prime $p = 5$) will not be useful; this may be a case of "gerrymandering" mathematical properties. I have tried to avoid such an issue by suggesting an important class of compositional properties. In our case, it is clear that both Szemerédi's theorem and Furstenberg Multiple Recurrence require Structure theorems in their respective proofs, i.e., the instantiation of the dichotomy between structure and randomness is apparent. Thus, the ascription of structural content is reasonable.

Though I have argued for my notion of content via a particular case (albeit a very robust one), I believe that the account may be generalized. Of course, this will require similar analyses of theorems proved via intuitively impure techniques. I will not undertake this here but would like to mention an exemplary case consistent with my philosophical discussion: the complex analytic (impure) and arithmetical (pure) proofs of the Prime Number Theorem.[103] The Prime Number Theorem (PNT) asserts that $\pi(x) \sim x/\log(x)$ as $x \to \infty$ where $\pi(x)$ counts the number of primes $p \leq x$. This was independently proved by Hadamard and de la Vallée Poussin in 1896 via appeal to complex function theory.[104] Indeed, they showed that the PNT is in fact equivalent to the non-existence of zeroes of the Riemann zeta function $\zeta(s)$ in complex variable $s = \sigma + it$ for $\mathrm{Re}(s) = 1$. It was long thought that an " 'elementary' proof of the PNT, not depending on analytical ideas remote from the problem itself," would be desirable ([Ingham, 2008], 651). This was achieved, again independently, by Selberg and Erdős whose analyses were entirely elementary and arithmetical. Many of the details are not germane to our discussion here; however, it is crucial to note the following structural similarity between the complex-analytic and arithmetical proofs as brought out by Ingham's incisive review of the Selberg and Erdős papers. Ingham extracts four analytical properties of $f := -\zeta'/\zeta$ which "embody the essential analytical fact on which previous proofs of the PNT [impure, complex-analytic ones] have been based." He goes on to note that

> What Selberg and Erdős do is to deduce the PNT directly from the *arithmetical counterparts* of (i), (iii), (iv), without the explicit intervention of the analytical fact. In principle this opens up the possibility of a new approach, in which the old logical arrangement is reversed and the analytical properties of $\zeta(s)$ are deduced from the arithmetical properties of the sequence of primes ([Ingham, 2008], 654, emphasis my own).

Arana reads this as indicating the presence of implicit complex-analytic content in explicitly arithmetical statements ([Arana, 2019]). My position should, by now, be quite predictable. What we have instead is the instantiation of higher-level structural properties by analytic functions and prime numbers, respectively. Selberg and Erdős eliminate essential

[103] This has been extensively analyzed by Arana. See, for instance, [Arana, 2019].
[104] See [Hadamard, 1896] and [de la Vallée Poussin, 1896].

appeal to strictly complex-analytic content and pass to these structural features. Thus, we need not relinquish the intuitive impurity of the complex-analytic proof of PNT by appeal to mysterious "hidden" content; we have instead the presence of structural content.[105]

Finally, this discussion of content will play a role in the following sections concerning explanation. The basic point is quite simple: the use of more abstract, infinitary, and impure techniques may not always preserve data relevant to providing an explanatory proof of a theorem. This is the mathematical analogue of Morrison's worry[106] about mathematized scientific theories and might incline one to think that pure proofs are more explanatory than impure ones. I believe that my notion of structural content provides one reason to think that this worry may not always be apposite. If one can show that particular structural facts crucial to the proof of a theorem T in context X also occur in context Y, then it is entirely reasonable to think that a proof of T in context Y will be a candidate explanatory proof. As I discuss in the following sections, this is precisely what happens in the ergodic proof of Szemerédi's theorem. We see the instantiation of the dichotomy between structure and randomness in both the number-theoretic and ergodic contexts, but because of the greater conceptual clarity in the ergodic context, which occurs in part because the infinitary nature of the objects smooths out many inessential details, the impure proof is explanatory, while the pure proof is not.

2.5 Impurity, Simplicity, and Explanation

2.5.1 Introduction

In this section, I consider the rather widespread claim that impure proofs are often simpler than pure proofs.[107] There has been some very nice recent work by Arana on this topic ([Arana, 2017]). The central idea of this paper is that various impure techniques do not univocally simplify proofs, where "simplification" is identified with the common complexity-theoretic metric of proof length. It is perhaps unsurprising that impure considerations do not yield this kind of simplification; however, is it plausible to say that some sort of simplification occurs? And if so, what kind? I will argue that the ergodic proof is simpler than the combinatorial proof of Szemerédi's theorem under a suitable construal of simplicity and that this is one way to understand the explanatory power of impure techniques. Part of my analysis of this claim will involve a brief examination of the combinatorial Structure Theorem (Szemerédi's Regularity Lemma; Lemma 2.5.3) and its role in the combinatorial proof, thus fulfilling the "comparison of techniques" promissory note from the introduction.

2.5.2 Does Impurity Yield Simplicity as Proof Length?

Let me provide a summary of the argument in [Arana, 2017]. After considering historical evidence for the claim that impure techniques generate simplicity, Arana notes that there

[105] This is not to say that the presence of these higher-level structural facts is not mysterious in itself.
[106] See both the Introduction and Section 2.7 below.
[107] See [Arana, 2017] for historical and contemporary evidence for the claim's pervasiveness.

are at least two kinds of simplicity at work.[108] The first is *verificational simplicity*: this measures the "...simplicity of determining whether a given proof is a proof at all; thus it measures the simplicity of confirming the validity of the deductions of a given proof" (*ibid.*, 4). The second is *inventional simplicity*: this measures "the simplicity of discovering a proof of a given statement" (*ibid.*, 4). His investigation focuses on the former and thus seeks to evaluate the claim that impure proofs are generally simpler to verify than pure proofs of the same statement.

As noted above, Arana evaluates the verificational simplicity of impure proof by appeal to proof-theoretic techniques.[109] In particular, the length of a proof in a formal theory is taken to be the measure of verificational simplicity. Consider a base theory T with conservative extensions. An extension T' of T is conservative over T iff for every sentence $\varphi \in \mathsf{L}_{\mathsf{T}}$ such that $\vdash_{\mathsf{T}'} \varphi$ we have $\vdash_{\mathsf{T}} \varphi$. The basic idea is then to compare the length of proofs of φ in T with those of φ in T'. We take Primitive Recursive Arithmetic (PRA) as our base theory. This is gotten from first-order Peano Arithmetic (Z_1; PA) by adding symbols and defining equations for all primitive recursive functions and restricting induction to quantifier-free formulas.

Following a classificatory scheme of Ignjatović[110] we have conservative extensions of PRA of two different types: arithmetical and conceptual. Arithmetical extensions typically add more induction. For example, one can obtain $I\Sigma_1$ as an arithmetical extension of PRA by restricting the induction schema of PA to Σ^0_1-formulas rather than quantifier-free formulas.[111] On the other hand, conceptual extensions add elements of an ostensibly different conceptual type: sets and set-theoretic principles. The reverse mathematical subsystems of second-order arithmetic mentioned above would then count as such conceptual extensions of PRA. In particular, Arana considers the following chain of extensions

$$\mathsf{PRA} \subset \mathsf{RCA}_0 \subset \mathsf{WKL}_0 \subset \mathsf{WKL}_0^+. \tag{2.5.1}$$

Now that we have our base theory and its conservative extensions fixed, we must ascertain whether these extensions generate impure proofs of theorems of PRA. Arana argues that conceptual extensions are topically impure[112] insofar as PRA utilizes facts only about natural numbers and makes no appeal to set-theoretic resources. The most important thing to note here is that, though PRA does use functions, these can be understood algorithmically and thus independently of set-theoretic formulations.[113] He also argues that arithmetical extensions are impure, but elementally so, as arithmetical extensions add induction principles stronger than the quantifier-free induction of PRA.

[108] These were first discussed in [Detlefsen, 1990].

[109] The reader should, once more, refer to Section C for formal definitions and brief discussions of the systems and results employed in this section.

[110] See [Caldon and Ignjatovic, 2005].

[111] See [Arana, 2017], p. 214 for more details concerning $I\Sigma_1$.

[112] See the introduction to this chapter for the distinction between topical and elemental impurity.

[113] See [Arana, 2017] for an extended defense of the impurity of conceptual extensions. In particular, one might argue that proofs of an $I\Sigma_1$-theorem in RCA_0 are topically impure. However, unlike PRA and its conceptual extensions, $I\Sigma_1$ and RCA_0 are mutually interpretable. Thus, it may be reasonable to conclude that the deployment of set-theoretic resources to prove $I\Sigma_1$-theorems in RCA_0 is "a mirage." However, is the mutual interpretability of theorems φ and ψ tantamount to their having the same meaning? This is a controversial semantic thesis that renders constraints of mathematical practice otiose, thereby impairing our ability to understand mathematics, and thus should be rejected.

Finally, Arana appeals to "speed up" results for conservative extensions of PRA. Here I go into slightly more detail following [Caldon and Ignjatovic, 2005]. For some proof p, we write the total length of the proof as $\ell(p)$, which counts the total number of symbol occurrences in p. Define

$$2_m^n = \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}}_{m}, \tag{2.5.2}$$

i.e., a stack of m twos with the final two having exponent n. A function f has *Kalmár elementary growth* if there is a natural number m such that $f(x)$ is eventually majorized by 2_m^x. A function has *roughly super-exponential growth rate* if it does not have a Kalmár elementary growth rate, but rather, for some polynomial $P(x)$ in natural coefficients, $f(x)$ is dominated by $P(2_x^x)$. Finally, $f(x)$ has *polynomial growth rate* if it is eventually dominated by a polynomial $P(x)$ in natural coefficients.

Definition 2.5.1. Let T' and T be two theories such that $\mathsf{T} \subset \mathsf{T}'$ and let Γ be a subset of theorems of T. Then

1. T' has *roughly super-exponential speed-up* over T with respect to Γ if there is a sequence of formulas $\{\gamma_i\}_{i\in\omega}$ in Γ such that if p_n^{T} and $p_n^{\mathsf{T}'}$ are the shortest proofs of φ_n in T and T' respectively, then:
   (a) No function f with Kalmár elementary growth rate satisfies $\ell(p_n^{\mathsf{T}}) < f(\ell(p_n^{\mathsf{T}'}))$ for all n;
   (b) There is a function f with a roughly super-exponential growth rate such that the inequality holds for all n.
2. T' has at most *polynomial speed-up* over T if there is a polynomial $P(x)$ in natural coefficients such that for any theorem φ of T we have: if p^{T} and $p^{\mathsf{T}'}$ are the shortest proofs of φ in T and T' respectively, then $\ell(p_n^{\mathsf{T}}) < P(\ell(p_n^{\mathsf{T}'}))$ for all n.

Note that roughly super-exponential speed-up results in a significant shortening of proofs. Thus, if any of the conservative extensions has a roughly super-exponential speed-up over PRA, it is reasonable to conclude that, for that extension, impurity and simplicity are indeed associated. On the other hand, two procedures are usually determined to belong to the same efficiency class if, for the same input, the number of steps needed to execute one procedure is less than or equal to the value of a polynomial evaluated at the number of steps in the other procedure. Thus, polynomial speed-up does not generate a significant shortening of proofs.

[Caldon and Ignjatovic, 2005] shows that $I\Sigma_1$ has roughly super-exponential speed-up over PRA with respect to Π_1^0 theorems of PRA, and so elementally impure proofs in $I\Sigma_1$ do effect simplification. Given the mutual interpretability of RCA_0 and $I\Sigma_1$, RCA_0 has at most polynomial speed-up over $I\Sigma_1$. Finally, with respect to Π_1^1-theorems, WKL_0 has at most polynomial speed-up over RCA_0 and similarly for WKL_0^+ over WKL_0. Thus, conceptually impure proofs in $\mathsf{RCA}_0, \mathsf{WKL}_0$, and WKL_0^+ do not produce any further significant simplification. From these findings, it is reasonable to conclude that, at best, the association of

impurity and simplicity is equivocal, at least insofar as we agree with Arana's arguments for the impurity of arithmetical and conceptual extensions and the identification of simplicity with proof-length.

Attempting such formal analyses is oftentimes a helping starting point for philosophical reflection, but I think we should be unsurprised by the fact that a relatively coarse formal measure does not perfectly map onto distinctions made in mathematical practice. Mathematicians rarely think about purely syntactic criteria like proof length. Indeed, if this were what they commonly meant when adverting to simplicity, it would be easy enough for them to say this, rather than making less precise claims about elegance or clarity or naturality. Philosophers and logicians are certainly live to this point. For example, Avigad writes

> [proof length] has something to do with explaining how infinitary methods can make a proof simpler and more comprehensible. But the advantages of working in a conservative extension seem to have as much to do with the perspicuity and naturality of the notions involved, and using the number of symbols in an uninterpreted derivation as the sole measure of complexity is unlikely to provide useful insight ([Avigad, 2003], 276n18).

But perhaps there are other candidate notions of simplicity up for analysis? And perhaps these might support the consistent association of impurity and simplicity in mathematical practice? Indeed, Caldon and Ignjatović remark

> ...in the above considerations, the speed-up is measured only in terms of lengths of proofs: this does not rule out a "conceptual speed-up," i.e., a formulation of the proof which uses concepts that make the proof easier to grasp, even if there is no speed-up in terms of length of formal proofs ([Caldon and Ignjatovic, 2005], 781).

These remarks are interesting and suggestive, but certainly in need of further clarification. In particular, there seems to be a local-global ambiguity at work; one might think that: (i) individual concepts formulated in an extension are clearer (local); or (ii) particular segments of the proof are made clearer in virtue of these concepts; or (iii) the entire proof is made clearer because of the way that the concepts in the extension interact (global). In the following section, I interpret Szemerédi's theorem in light of these remarks and argue that something like (iii) occurs in the ergodic proof.[114] Thus, I provide an interesting and important example in which impurity yields simplicity construed as "conceptual speed-up." Finally, I argue that the simplicity achieved by the impure proof also generates explanatory power.

2.5.3 Does Impurity Yield Simplicity as Conceptual Speed-up?

Arana's examination of the relationship between impurity and simplicity in a formal, proof-theoretic context has its advantages: for example, we can say that the move from PRA to $I\Sigma_1$ (elemental impurity) produces some speed-up; however, proofs in *conservative* conceptual

[114] Both the remarks by Avigad and Caldon and Ignjatović appeal to the perspicuity of concepts formulated in conservative extensions. We do not have conservativity in our case (see below).

extensions of PRA (topical impurity) do not shorten proofs any further. On the other hand, this method "...may distort the phenomena being measured," and/or, as I noted above, may simply fail to correspond to distinctions that interest us. And this may occur because, once we formalize theorems of ordinary mathematics in, say, PRA, the information we wish to analyze is not faithfully represented.[115] As such, it seems to me that an evaluation of the relationship between impurity and simplicity is more fruitfully performed internal to "ordinary" mathematical practice. However, before doing this, let us see if anything interesting can be extracted from a formal analysis of Szemerédi's theorem.

Unfortunately, little appears to be known definitively about its strength.[116] It has been shown, via a method of Shelah, that van der Waerden's theorem (a special case of Szemerédi's theorem) is provable[117] in RCA_0. Thus, it seems natural to ask whether Szemerédi is provable in, say, ACA_0, or, more optimistically, in RCA_0. If one examines Szemerédi's original proof in [Szemerédi, 1975], one will see that it should be formalizable in RCA_0 as it has elementary bounds and the argument is entirely combinatorial. Furthermore, utilizing bounds by Gowers in [Gowers, 2001], it is likely that this can be gotten down to EFA. I hasten to note that these statements should be understood as "confident claims" based upon an examination of the bounds and resources in the arguments. However, in order to provide rigorous justification, one would have to encode carefully all definitions involved in the language of second-order arithmetic, which I have not done.[118]

Thus, given that the weakest natural subsystem of Z_2 in which Szemerédi's theorem can be proven is not known, we do not have a base theory from which to begin a formal analysis in Arana's sense. But let us see what can be done if we assume RCA_0 as our base theory. It is known that the Furstenberg Structure Theorem, the key to the ergodic proof of Szemerédi's theorem, can be formalized in Π^1_1-CA_0. It is also conjectured that the reversal holds (over ACA_0).[119] Thus, we would be interested in the following conceptual extension

$$\mathsf{RCA}_0 \subset \Pi^1_1\text{-}\mathsf{CA}_0\,. \tag{2.5.3}$$

More cautiously, the paper [Avigad and Towsner, 2010] shows that Furstenberg's original proof in [Furstenberg, 1977] requires the strength of ACA_0 (over RCA_0) where we prove Furstenberg Multiple Recurrence (Theorem 2.3.12) for each k via the Furstenberg Structure Theorem (Theorem 2.3.37). We require slightly more than ACA_0 if we are proving Furstenberg Multiple Recurrence for all k. Thus, we would then be interested in the conceptual

[115] It is also the case that metamathematical results required to facilitate such an analysis may not be available.

[116] See Simpson's survey article [Simpson,]. I would find it surprising, if the strength of Szemerédi's original proof had been established, that this was not mentioned in, e.g., [Avigad, 2009], which deals with the theorem directly.

[117] Shelah shows that the Hales-Jewett theorem implies van der Waerden's theorem. Then he shows that the Hales-Jewett function has primitive recursive upper bounds, and thus the van der Waerden function does. See [Shelah, 1988]. Recently, it has been shown by Matet in [Matet, 2007] that something much weaker than RCA_0 is needed: merely super-exponential function arithmetic.

[118] See Appendix B for further discussion.

[119] The equivalence of the Structure Theorem with Π^1_1-CA_0 over ACA_0 was claimed in [Avigad, 2009], but now Avigad is not confident in this claim. Again, see Appendix B.

extension

$$\mathsf{RCA}_0 \subset \mathsf{ACA}_0, \tag{2.5.4}$$

or perhaps $\mathsf{RCA}_0 \subset \mathsf{ACA}_0{}^+$. It seems entirely possible that as we move to subsystems of Z_2 stronger than those considered in [Arana, 2017] and [Caldon and Ignjatovic, 2005], i.e., stronger than RCA_0, WKL_0, and WKL_0^+, we might get a significant speed-up (better than polynomial) and thus a significant shortening of proof length.

However, we also need a conservativity result in order to apply Arana's analysis. That is, we wish to find conservative extensions of RCA_0 that offer more than polynomial speed-up (and where these extensions add set-existence principles, i.e., are "conceptual" extensions). Preliminary results by Yokoyama in [Yokoyama, 2010] suggest that what we are after will not materialize; in particular, he proves the existence of a maximal Π^1_2-axiomatizable[120] conceptual extension of RCA_0 and conjectures that no such extension offers more than polynomial speed-up. As Arana notes, all known methods of producing conservative extensions of RCA_0 rely upon the Π^1_2-axiomatizability of the extension. Thus, if Yokoyama is able to prove his conjecture (as of yet, this has not been done) or if no new methods for producing conservative extensions of RCA_0 are found, then we should not expect any significant speed-up results for conceptual extensions of RCA_0.[121] To summarize: a number of technical results would have to fall very precisely in place for an analysis of simplicity in terms of proof-length to apply to Szemerédi's theorem, and, even if this were to happen, I have already noted that the dividends may be of questionable philosophical value.

In any case, let me turn to the details of our case study to see if these can yield further insight into the relationship between impurity and simplicity. Like Arana, I am here interested in *verificational simplicity*, i.e., the simplicity of confirming the deductions of a given proof. The impurity in question is, again, both elemental and topical: the resources of the ergodic proof are both more computationally complex than the combinatorial theorem in question (elemental) and involve intuitively different concepts (topical). The topical impurity proceeds from my argument in the above section. I claimed that, despite the surprising equivalence of Szemerédi's theorem and the ergodic Furstenberg Multiple Recurrence, we should retain a notion of intuitive content, which can then be supplemented by structural content. Thus, our question is whether the radical impurity of the ergodic proof of Szemerédi's theorem is verificationally simpler than the combinatorial proof insofar as the impure proof yields conceptual speed-up.

Before an explicit comparison of techniques, consider again the local-global ambiguity in Caldon and Ignjatović's discussion of conceptual speed-up. *Prima facie*, if one is thinking locally in terms of particular concepts, it appears obvious that some impure techniques are not simpler than pure ones. Take the case at hand: the basic objects[122] of study in Szemerédi's theorem are sufficiently dense sets of integers (or even just subsets of $\mathbb{N}$). This seems entirely comprehensible to a mathematical novice; one needs only to understand the integers as the counting numbers along with their additive inverses (and zero). We then have

[120] Note that both WKL_0 and WKL_0^+ are Π^1_2-axiomatizable.

[121] This might change if new methods were found of producing conservative extensions of RCA_0 not relying on Π^1_2-axiomatizability.

[122] I use "object" here as a term of art; it is not intended to convey any sort of ontological commitment.

the only slightly more difficult notion of density, which is a particular measure of the size of an infinite set relative to an ambient set. On the other hand, the basic objects of the ergodic proof are measure preserving systems $(X, \mathscr{B}, \mu, T)$. Such an object is straightforwardly more complex: there are more distinct entities comprising the system (a set, a σ-algebra, a measure, a measure-preserving transformation), whereas in the combinatorial case there are merely numbers. Furthermore, the entities of a measure-preserving system are less familiar to those without a more advanced mathematical education (even though they can be intuitively understood). Thus, we should not think that simplicity as conceptual speed-up occurs at the local level. However, to move from this point to the claim that, globally, an impure proof is more complex is to commit a compositional fallacy. It is entirely possible that the way in which the less familiar objects interact in the proof is more perspicuous than the interaction of the more familiar objects. Let me now demonstrate that this is the case in a (brief) comparison of the combinatorial and ergodic proofs of Szemerédi's theorem.

2.5.4 Comparison of Techniques

The fundamental point of this comparison can—once more—be profitably understood through Aristotle's *Posterior Analytics*. In *Post. An.* A.2, Aristotle gives an explication of scientific knowledge[123] as follows:

> We think that we have scientific knowledge of each thing without qualification, and not in the sophistical manner accidentally, whenever we think that we know the cause on account of which the object is, that it is the explanation [or cause] of that [object], and that it is not possible for this [the object] to be otherwise.[124] (translation my own; 71b9-12)

Accordingly, there are three obvious ways[125] in which we might fail to have scientific knowledge (or understanding) of some fact X: (i) by getting the cause/explanation[126] of X wrong, e.g., thinking it is Z instead of Y; (ii) by knowing the explanation Y of X, but failing to register it *qua* explanation[127]; (iii) by thinking that X is not necessary. Remaining agnostic concerning this last condition, my claim is that the pure proof of Szemerédi's theorem fails to generate understanding, i.e., is not explanatory, because (ii) occurs. We have at our disposal the crucial structural fact that provides the reason why for the theorem, but it is not recognized as such in the pure proof. This is because the role of this structural fact is lost amid

[123] "Scientific knowledge" translates *epistēmē* (*ἐπιστήμη*). It is translated by some as "understanding". See, for instance, [Barnes, 1993].

[124] The Greek of [Ross and Minio-Paluello, 1964] reads: *Ἐπίστασθαι δὲ οἰόμεθ' ἕκαστον ἁπλῶς, ἀλλὰ μὴ τόν σοφιστικὸν τρόπον τὸν κατὰ συμβεβηκός, ὅταν τήν τ' αἰτίαν οἰώμεθα γινώσκειν δι' ἣν τὸ πρᾶγμά ἐστιν, ὅτι ἐκείνου αἰτία ἐστί, καὶ μὴ ἐνδέχεσθαι τοῦτ' ἄλλως ἔχειν.*

[125] Note that Aristotle's gloss on scientific knowledge here precedes (and prepares the way for) his theory of demonstration, which provides *further* conditions that guarantee an Aristotelian syllogism can produce scientific knowledge. I do not wish to enter into a discussion of these further conditions and merely wish to point out that Aristotle's discussion of knowledge early on in the *Posterior Analytics* remains quite instructive.

[126] Translating *αἰτία*.

[127] See his famous discussion of the syllogisms involving the non-twinkling and nearness of the planets in *Post. An.* A.13.

the welter of delicate combinatorial manipulations required to carry out the proof. Namely, and somewhat ironically, our efforts to establish the truth of the theorem in this manner obscure the reason for the truth. On the other hand, the impure ergodic proof effects a sort of global conceptual simplification and renders the explanatory structural fact evident. This simplification occurs in part because of the infinitary nature of the ergodic objects; it leads to an explanatory proof because structural content[128] is shared between the combinatorial and ergodic domains, which ensures we do not lose the putative reason why. Thus, we see that the relationship between impurity, explanation, and simplicity is quite complicated: the (topical) impurity has to be sufficiently constrained, i.e., it must be an appropriate setting in which to prove the original result, but if this is the case, then various other aspects of the impurity (e.g., its infinitary aspects) serve to generate significant conceptual advantages. It may be significant that both varieties of impurity (elemental and topical) are present in this case. It would be interesting to see if my association of impurity with explanatory power holds when only one variety is present.

I will now try to give a sense for the conceptual obstructions present in the pure proof of Szemerédi's theorem and how these are overcome in the ergodic proof.[129] It will be helpful to consider remarks on the matter by Tao:

> Szemerédi's original proof of this theorem is a remarkably intricate piece of combinatorial reasoning. Most proofs of theorems in mathematics—even long and difficult ones—generally come with a reasonably compact "high-level" overview, in which the proof is (conceptually, at least) broken down into simpler pieces. There may well be technical difficulties in formulating and then proving each of the component pieces, and then in fitting the pieces together, but usually the "big picture" is reasonably clear. [...] In contrast, the pieces of Szemerédi's proof are highly interlocking, particularly with regard to all the epsilon-type parameters involved; it takes quite a bit of notational setup and foundational lemmas *before the key steps of the proof can even be stated, let alone proved* (Blog post here; emphasis my own).

He goes on to note:

> Many years ago I tried to present the proof [of Szemerédi's original paper], but I was unable to find much of a simplification, and my exposition is probably not that much clearer than the original text. Even the use of nonstandard analysis, which is often helpful in cleaning up armies of epsilons, turns out to be a bit tricky to apply here. (In typical applications of nonstandard analysis, one can get by with a single nonstandard universe, constructed as an ultrapower of the standard universe; but to correctly model all the epsilons occuring in Szemerédi's argument, one needs to repeatedly perform the ultrapower construction to obtain a (finite) sequence of increasingly nonstandard (and increasingly saturated) universes, each one containing unbounded quantities that are far larger than any

[128] Again, the dichotomy between structure and randomness encoded in the respective Structure Theorems.

[129] I claim no originality in my presentation. Indeed, I have adhered closely to Tao's own presentation of these obstructions here. I have tried to map what he says there onto the original theorems, facts, lemmas, etc. in [Szemerédi, 1975].

> quantity that appears in the preceding universe [...] This sequence of universes does end up concealing all the epsilons, but it is not so clear that this is a net gain in clarity for the proof.

Given the first quote, it may be a fool's errand to try to indicate, even in outline, the main steps of Szemerédi's theorem, but I would like to provide my reader with some further evidence for my claims. The second quote is interesting for a few reasons. One is that the pure proof, after nearly 50 years, still has no significant (combinatorial) simplification. This is something of a rarity and again indicates the great complexity of the theorem.[130] Another is that other infinitary techniques (non-standard universes) have not provided any real simplification of the proof. I would like to suggest (tentatively) that this gives us further warrant for thinking the ergodic setting is the *correct* or *appropriate* infinitary setting for proving Szemerédi's theorem in an explanatory fashion.

Let us now take a look at some of the key ingredients of the pure proof of Szemerédi's theorem. Consider first the combinatorial Structure Theorem, i.e., the Szemerédi Regularity Lemma (see Lemma A.1 for the original statement). Here I present a more modern formulation. Begin with the definition:

Definition 2.5.2. (ϵ-regularity) Let $G = (V, E)$ be a graph with A, B sets of vertices in V. Let

$$e(A, B) := |\{(a, b) \in A \times B : \{a, b\} \in E\}|,$$

i.e., the number of (a, b) such that ab is an edge of the graph. Then the *density* of A, B is

$$d(A, B) := \frac{e(A,B)}{|A||B|}.$$

Finally, we call the pair A, B ϵ-regular if

$$|d(A', B') - d(A, B)| \leq \epsilon$$

for all subsets $A' \subset A$ and $B' \subset B$ such that $|A'| \geq \epsilon|A|$ and $|B'| \geq \epsilon|B|$.

The crucial idea of this definition is that a pair (A, B) is regular if it resembles a random graph with edge-probability d. We then have:

Lemma 2.5.3. *(Szemerédi Regularity Lemma; modern version) Let $G = (V, E)$ be a graph, and let $\epsilon > 0$. Then there is a positive integer K_0 such that, for any graph G, V can be partitioned into sets $V_1 \cup \cdots \cup V_K$ with $K \leq K_0$, with sizes differing by at most 1, such that all but at most ϵK^2 of the pairs (V_i, V_j) $(1 \leq i \leq j \leq K)$ are ϵ-regular.*

Thus, this asserts, roughly, that every graph can be decomposed into a few pieces, all of which resemble a random graph.[131] Random graphs of given edge density d are generally much easier to handle than all graphs of edge density d. Thus, we can carry over results

[130] For instance, it is interesting that the Green-Tao theorem, which uses Szemerédi's theorem as one of its main steps, has already been significantly simplified by Conlon, Fox, and Zhao.

[131] More precisely, the "pieces" are the ϵ-regular pairs and these behave like random bipartite graphs. A *bipartite* graph is a graph $G = (V, E)$ such that we can partition $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$, and $E \subseteq V_1 \times V_2$ such that every $e \in E$ has one endpoint in V_1 and the other in V_2.

that are trivial for random graphs to all graphs. This is a structural result very much like the Furstenberg Structure Theorem.[132] Recall in this case it was easy to see why recurrence patterns hold in random (weak mixing) and structured (periodic) systems. Because of the Furstenberg Structure Theorem, we then know that any measure-preserving system will be broken down into components we can handle easily and thus prove the general recurrence result.

So much for the combinatorial structure theorem. The other key ingredients of Szemerédi's proof are van der Waerden's theorem (Theorem(s) 2.3.7 and 2.3.8) and a very tricky analysis of the densities of sets $A \subset \mathbb{Z}$ along what are typically called *generalized arithmetic progressions*. This analysis amounts roughly to the following: we are interested in finding some arithmetic progression $P_1 := a, a + r, \ldots, a + (k-1)r$ (rank 1, k length arithmetic progression) in every sufficiently dense $A \subset \mathbb{Z}$. Importantly, one can construct an arithmetic progression of arbitrary rank. So, for example, a rank 2 arithmetic progression will be of the form $P_2 := P_1, P_1 + r_2, \ldots, P_1 + (k_2 - 1)r_2$, i.e., an arithmetic progression of arithmetic progressions P_1 of rank 1. Szemerédi's basic strategy is to construct a massive arithmetic progression of high rank D containing many elements of $A \subset \mathbb{Z}$ and then winnow down this huge generalized arithmetic progression until one arrives at a rank 1 arithmetic progression containing *only* elements of A. Note that Lemma 5, Lemma 6, and Fact 12 of [Szemerédi, 1975] involve generalized arithmetic progressions in this way.[133]

Let me attempt to describe how this winnowing process works. In order to terminate the process at a k length rank $D = 1$ arithmetic progression with all elements in A, we must choose our generalized arithmetic progression to be of suitable rank. This turns out to be quite large: $2^k + 1$ (I remark on this below). The next step is to locate a generalized arithmetic progression of rank $2^k + 1$ that is "saturated" by elements of A.[134] More or less, this amounts to the fact that each rank $D - 1$ arithmetic progression comprising our massive generalized arithmetic progression has an "almost maximal" number of elements of A. Once this is done, we hope to find a family of arithmetic progressions of rank $D-1$ inside the generalized arithmetic progression of rank D, which has sufficiently nice arithmetical properties in order to invoke a weak mixing assumption. This assumption is applied to sets $A_i := \{a \in P_{D-1} : a + ir_D \in A\}$ with $i = 0, \ldots, k_D - 1$. As noted above, weak mixing is tantamount to saying that A_i behaves like a random set of integers. Explicitly, this weak mixing assumption looks like

$$|A_i \cap E| \approx \delta\sigma|P_{D-1}| \tag{2.5.5}$$

where $P_D := P_{D-1}, P_{D-1} + r_D, \ldots, P_{D-1} + (k_D - 1)r_D$, E is an arbitrary subset of P_{D-1} with density approximately σ and δ is the density of A_i in P_{D-1} with $i = 0, \ldots, k_D - 1$. We must now justify and significantly strengthen this weak mixing assumption. This is done via a sort of double counting argument and invokes both van der Waerden's theorem and, most importantly, the Szemerédi Regularity Lemma. In essence, one must justify the weak mixing assumption for a very large number of subsets E_j; in particular, the number of E_j for which

[132] See Theorem 2.3.37 above.

[133] Note that there is a slight error in Szemerédi's diagram included in the appendix. We should have $L_6 \to F_{12}$, i.e., Lemma 6 implies Fact 12, not the other way round.

[134] See [Szemerédi, 1975], p. 208, for the technical definition of saturated.

the weak mixing assumption holds cannot be much smaller than the length of the arithmetic progression being used to establish the assumption (in this case length $k_D - 1$). This is precisely the role of the Regularity Lemma: it allows us to show that sufficiently many of the E_j have this weak mixing (random) behavior. This induces a condition like the above.[135] We then wish to apply this strong weak mixing condition to the family of generalized arithmetic progressions of rank $D-1$ sitting in the original generalized arithmetic progression of rank D. This can only be done if the family is structured in a very precise way.

Even after all this has been done there is a very significant obstruction. We wish to apply the analogous weak mixing condition to appropriate index $i = 0, \ldots, k_D - 1$, but getting this to hold for all i simultaneously is exceedingly difficult. Furthermore, Tao describes it as a "Tower of Hanoi" situation wherein once the condition is gotten to apply to some i, there is no guarantee that it holds for previous values of i. This partially accounts for the huge rank $D = 2^k + 1$ of the generalized arithmetic progression as we must constantly move back and forth between indices to get the correct behavior of the progressions and subsets such that we eventually find a rank 1 arithmetic progression with all elements in our original A.

A few remarks are now in order. First, this outline is the coarsest exegesis of the most basic steps of the proof and abstracts away from many details of Tao's own outline, which itself gives only the "slightest hint" as to how everything fits together. One should note especially that many details have been left imprecise (mere approximation $\approx$ rather than providing explicit error terms, adverting to "nice" properties of order, etc.). This should be contrasted to the outline of the ergodic proof provided in Section 2.3.2, which, I hope, leaves one much less in the dark as the main moves of the proof can be stated with much less set up. Second, especially given the issues of moving back and forth between indices (the Tower of Hanoi), the combinatorial proof of Szemerédi's theorem is highly *non-linear*. Again, contrast to the ergodic proof, which, though complex component-wise, can be set forth in a linear fashion. The non-linearity and linearity respectively are perfectly *objective* features of the proofs and, at the "global" level of overall proof structure, the linear ergodic proof is much more perspicuous than the non-linear combinatorial one.

Third, this kind of simplification of the overall proof structure can then be said to generate explanatory power. An explanation, taken to be a syntactic or semantic object (an argument of some sort), should *prima facie* be of an asymmetric character.[136] We proceed from the *explanans* to the *explanandum* and not vice versa. To do so would be to deny the possibility of explanation whatsoever. Having our proof structured in a linear way makes this explanatory asymmetry apparent: once we have proved the Correspondence Principle (Theorem 2.3.14), the ergodic proof essentially reduces to relativizing the Koopman-von Neumann Dichotomy Theorem (Theorem 2.3.27) to obtain the Furstenberg Structure Theorem (Theorem 2.3.37). Once this theorem has been gotten, we proceed directly from it to the Correspondence Principle to Szemerédi's theorem (Theorem 2.3.4). On the other hand, even though the Regularity Lemma (Lemma 2.5.3) can licitly be said to be the reason why Szemerédi's theorem holds, *how* it interacts with the other features of the combinatorial proof is quite opaque. In particular, justifying why we need the weak mixing assumptions (especially in its

[135] See Lemma 4 and Lemma 5 in [Szemerédi, 1975].

[136] I take it that since we here are talking of proofs, i.e., a kind of argument, we can say that we are interested in finding the subset of proofs that are explanations.

strongest form) and thus why we need the Regularity Lemma is a much more difficult task than justifying why we need the Furstenberg Structure Theorem. Furthermore, even once we have used the Regularity Lemma to obtain our desired weak mixing property, actually getting this property to hold for various indices is very difficult. Given this opacity, the pure combinatorial proof fails to yield any sort of understanding as to why Szemerédi's theorem is true, and thus fails to be explanatory. I would like to stress that I do not think this is a strictly pragmatic or epistemic feature of the situation. Rather, it is because of an objective feature of the pure proof that one fails to understand the result: we are not able to locate the structural and explanatory fact *qua* explanation internal to the nexus of reasoning that constitutes the proof.[137] As I stated in the introduction, in order to have a truly explanatory proof, one must see both the putative reason why and how this reason interacts with other aspects of the proof. This further requirement allows one to register the explanation *qua* explanation, and this only happens in the impure proof.

2.5.5 Summary

Thus, I believe this analysis supports the following chain of inferences: Impurity ⇒ Simplicity ⇒ Explanatory Power. The impure ergodic proof of Szemerédi's theorem generates simplicity construed as "conceptual speed-up." This conceptual speed-up does not occur at a "local" level: as I have noted the concepts involved in the ergodic proof are, taken on their own, more difficult to understand. However, once we pass to the "global" level of overall proof structure, the way in which the ergodic concepts interact does generate a significant simplification of the proof, in particular, a more direct and linear one. This occurs in large part because of the infinitary nature of the ergodic entities. We are, for example, able to work with the behavior of measure-preserving systems in the limit, which significantly reduces the difficulties of ascertaining error terms and delicate density calculations required in the combinatorial setting. This is all well and good, but why would one expect this simplification to lead to an explanatory proof? Indeed, would it not be the case that crucial explanatory data would be lost? This is where we can invoke the structural content shared by Furstenberg Multiple Recurrence and Szemerédi's theorem. Both belong to the mathematical genus defined by the "dichotomy between structure and randomness," and so both theorems are about this dichotomy in some sense. Furthermore, this structural fact is the reason why each theorem holds: decomposing either $A \subset \mathbb{Z}$ or arbitrary measure-preserving system X into structured and random pieces allows one to find the desired arithmetic progression or recurrence pattern. Thus, the crucial explanatory fact is present in both the combinatorial and ergodic domains. However, the combinatorial proof, given its many computational detours, does not show that this structural fact is the reason why. Not so in the ergodic setting: the advantages of the impure and infinitary setting can be exploited to bring the structural fact to light.

[137] That is, option (ii) from the beginning of Section 2.5.4 occurs.

2.6 The Necessity of Infinity and A Bridge to Platonism

My analysis of the explanatory power of the ergodic proof of Szemerédi's theorem can be put into fruitful dialogue with Feferman's discussion of the purported "necessity" of the infinite in [Feferman, 1987]. After examining the main points of this paper, I will draw some epistemological and metaphysical conclusions about the impure and infinitary entities that have featured in the above discussion. I argue that particular infinitary resources *are* in some sense necessary[138] and that these resources might be understood platonistically (contra Feferman) via a form of indispensability argument.

2.6.1 In What Sense is Infinity Necessary?

Feferman calls *Gödel's Doctrine* the claim that

> the "true reason" for the incompleteness phenomena is that "the formation of ever higher types can be continued into the transfinite," both in systems explicitly using types and systems of set theory such as ZF for which the (cumulative) type structure is implicit in the axioms ([Feferman, 1987], 190; quoting [Gödel, 1931], fn. 48a).

Informally, this means that, given some system S_0 whose axioms are thought to be true, we adjoin an extension S_1 that decides previously undecidable propositions of S_0; however, the incompleteness simply reappears in the form of undecidable propositions of S_1, requiring the addition of yet another extension, S_2, etc. Specifying Gödel's general claim to the case where the "undecidable propositions" are of finitary character, Gödel's Doctrine can be glossed as: "...the unlimited transfinite iteration of the power-set operation is necessary to account for finitary mathematics" ([Feferman, 1987], 190). Feferman also considers a formulation of Gödel's Doctrine in terms of systems of set theory; however, this formulation does not directly concern us here, and so I will focus upon that in terms of finitary propositions and the powerset.

As is well known, $\mathsf{Con}(\mathsf{PA})$ is a Π^0_1 statement not provable in PA. In order to prove it, one provides an extension of PA involving higher-order quantification, full induction, and a comprehension principle (Π^1_1-CA, though Feferman shows Δ^1_1-CA suffices[139]). This extension allows one to formulate a definition of formal truth for PA and then prove the statement expressing that "every statement provable in PA is true." *A fortiori* this establishes $\mathsf{Con}(\mathsf{PA})$, since, for the formal truth predicate $\mathsf{Tr}_{\mathsf{PA}}(\cdot)$, we have

$$\neg\mathsf{Tr}_{\mathsf{PA}}(\ulcorner\bot\urcorner) \rightarrow \mathsf{Con}(\mathsf{PA}). \tag{2.6.1}$$

Thus, the adjunction of higher-order principles allows us to prove statements with explicit, finitary content, viz., $\mathsf{Con}(\mathsf{PA})$. Of course, we must be careful to distinguish between potentially innocuous higher-order principles[140] and those that rely upon the full powerset. It is

[138] Viz., necessary for maximally intelligible mathematics.

[139] See Section C for definitions and brief discussions of many of the technical terms employed in this section. Feferman's proof sketch occurs on pp. 191-192 of [Feferman, 1987]. Actually, even less than Δ^1_1-CA is required since $\mathsf{ACA} \vdash \mathsf{Con}(\mathsf{PA})$, but Feferman does not consider this here.

[140] Recall that an analogous point occurred in my discussion of Isaacson's notion of higher-order content.

this latter sort that principally concerns Feferman:

> Gödel's doctrine can be challenged when it is read as asserting that the platonist view of the determinateness of the power set operation and its iteration through all the ordinals is *necessary* for the derivation of *previously* undecidable but true Π^0_1 statements ([Feferman, 1987], 193).

Thus, he thinks we proceed from (i) platonism to (ii) the "determinateness" or meaningfulness of the powerset to (iii) the requisite comprehension principles that allow for the proof of finitary results. Note that, in showing the proof of Con(PA) requires only Δ^1_1-CA, Feferman has *already* undermined Gödel's Doctrine and its attendant platonism. This is because Δ^1_1-CA only requires a small fragment of the powerset of the naturals and consequently does not require platonism to justify belief in the determinateness of the full powerset. Indeed, [Feferman and Sieg, 1981] shows that Δ^1_1-CA is predicatively justifiable.[141] There is, however, one conceptual worry present here: it is not entirely obvious why Feferman associates the meaningfulness of the full powerset with *platonism alone*. It is conceivable that other philosophical positions could justify, say, mathematics utilizing impredicative comprehension principles or even much more.[142] This association is, however, crucial to the integrity of Feferman's argument since he seeks to undermine Gödel's Doctrine given his discomfort with platonism as the "medieval metaphysics of mathematics" (following Weyl).

In any case, he provides a few interesting ways in which one might reinterpret the "true reason" for the incompleteness phenomena, none of which requires the adjunction of higher types and, in particular, the transfinite iteration of the powerset operation. Given that, in the example outlined above, we require a formal truth predicate to deduce Con(PA), Feferman suggests that the crux of the issue is the inexpressibility of a formal truth predicate of some language L in L itself. Modulo the one issue mentioned above, I find much of this argument compelling, and so I take Feferman to have successfully undermined the necessity claim in Gödel's doctrine: our understanding of independence/incompleteness phenomena does not *require* the transfinite iteration of the powerset operation.

Feferman then considers a fascinating family of theorems that might cause trouble for this conclusion. These theorems, e.g., Paris-Harrington, Goodstein's, finite Kruskal's, are "finitary independence results," resembling "naturally occurring" combinatorial results relevant to everyday mathematical concerns.[143] For instance, both the Paris-Harrington theorem and Goodstein's theorem are independent of PA and thus independent of ACA_0; more pressing for Feferman, finite Kruskal's theorem is independent of ATR_0. This latter result might offer weak support for Gödel's doctrine insofar as it requires Π^1_1-CA_0 for its proof (or appears to do so). Feferman notes

> Now (the [Gödelian] argument) continues, this system [Π^1_1-CA_0] is justified only if one assumes that the powerset $\mathscr{P}(\mathbb{N})$ exists as a fixed definite totality. Hence, according to this line of argument, it is necessary to make platonistic assumptions concerning the existence of uncountable totalities in order to derive finite

[141] For a brief discussion of this fact, see [Feferman, 1998b].

[142] For instance, consider Maddy's mathematical naturalism.

[143] One might challenge their relevance, of course. Feferman himself does so, citing a "nagging feeling that the statements still have a 'cooked-up' look" ([Feferman, 1987], 201).

> combinatorial truths P_{fin} [finite Kruskal's theorem and an extended version due to Friedman]...In other words, according to this line of thought, at least the first stage of the Cantorian transfinite is necessary for everyday combinatorial mathematics ([Feferman, 1987], 200).

Taking the necessity in question to be "necessary for proving," Feferman has two rejoinders to this reworked Gödelian argument from the finitary independence results. The first is that we require only the 1-consistency[144] of the impredicative comprehension principle, not the principle itself, in order to prove the finitary statements in question. And, as Brouwer stressed to Hilbert, consistency alone does not yield truth. Second, one might appeal to proof-theoretic reductions of subsystems of Z_2 to constructive theories. Informally, we say that a system T_1 is *proof-theoretically reducible* to another system T_2 with respect to some class of formulas Φ when every proof π_1 of T_1 ending in formula $\varphi \in \Phi$ can be effectively transformed into a proof π_2 of T_2 (also ending in φ). Furthermore, this transformation can be formalized entirely in T_2; it should be apparent that a proof-theoretic reduction immediately implies conservativity (over the relevant class of formulas). The basic idea is that, if we can proof-theoretically reduce the impredicative theory $\mathsf{\Pi}^1_1\text{-}\mathsf{CA}_0$ to some weak theory, which is itself justified without appeal to $\mathscr{P}(\mathbb{N})$, then the use of $\mathsf{\Pi}^1_1\text{-}\mathsf{CA}_0$ in the finitary independence results is anodyne.[145] The results contained in [Buchholz et al., 1981] provide examples of such reductions in terms of "iterated inductive definitions." In particular, it is shown that $\mathsf{\Pi}^1_1\text{-}\mathsf{CA}_0$ is (finitarily) reducible to $\mathsf{ID}^{\mathrm{i}}_{<\omega} := \bigcup_{n<\omega} \mathsf{ID}^i_n$ with respect to Π^0_2 formulas. Here $\mathsf{ID}^{\mathrm{i}}_{<\omega}$ is a constructive (i.e., intuitionistic, indicated by the i superscript) system of inductive definitions iterated $< \omega$ times (indicated by the $< \omega$ subscript). To fix intuitions, let me briefly define the simplest theory of iterated inductive definitions, ID_1, first introduced by Kreisel.[146] Note that this theory is not intuitionistic.

Example 2.6.1. ID_1 is an axiomatic theory consisting of the axioms of PA with additional predicates P defined in the following way. Take some arithmetical formula $\varphi(P, x)$ where P has only "positive occurrences" in φ. The positivity condition means that P always occurs without negations when we write everything in negation-normal form. This ensures that we get "new information" only from "previously given information" or "from below." This is quite important as Feferman relies on this kind of intuitive gloss in defending iterated inductive systems as especially perspicuous. Given the positivity constraint, we can then define a monotonic operator:[147]

$$\Gamma_\varphi : \mathscr{P}(\mathbb{N}) \to \mathscr{P}(\mathbb{N}), \quad \Gamma_\varphi(X) = \{x \in \mathbb{N} : \varphi(X, x)\}\,. \tag{2.6.2}$$

The monotonicity of Γ_φ is sufficient to show that there is a smallest fixed point, $I_{\Gamma_\varphi} :=$

[144] We define 1-consistency as: for some theory T and any primitive recursive φ, we cannot have both (i) $\mathsf{T} \vdash \exists x\, \varphi(x)$ and (ii) $\mathsf{T} \vdash \neg\varphi(\underline{0}), \neg\varphi(\underline{1}), \ldots$.

[145] It would then presumably join the good company of $\mathsf{\Delta}^1_1\text{-}\mathsf{CA}$, even though $\mathsf{\Pi}^1_1\text{-}\mathsf{CA}_0$ is indeed fully impredicative.

[146] I have benefitted from sections of the Preface and Chapter 1 of [Buchholz et al., 1981] by Feferman in trying to absorb this information.

[147] We can also formulate iterated inductive systems in terms of very concrete sets of rules; this formulation is in fact equivalent (in a particular sense) to the operator formulation. See [Buchholz et al., 1981], pp. 19-20.

$\bigcap\{X : \Gamma_\varphi(X) \subseteq X\}$. The new predicates P for ID_1 are then just such smallest fixed points associated to each formula φ. We complete the construction of ID_1 by adding the following axioms:

1. $\forall x(\varphi(P, x) \to P(x))$;
2. $\forall x(\varphi(\psi_P, x) \to \psi(x)) \to \forall x(P(x) \to \psi(x))$,

for each ψ and where ψ_P means that we replace atomic formulas P with ψ.

Given the availability of the proof-theoretic reduction of Π^1_1-CA_0 to $\mathsf{ID}^{\mathrm{i}}_{<\omega}$, Feferman believes we can dispense with the problematic impredicative principles. Indeed, in light of this reduction, we are supposed to be in a very epistemically secure place for two reasons. First, we utilize only intuitionistic logic in $\mathsf{ID}^{\mathrm{i}}_{<\omega}$ and thus no appeal to the "completed infinite." Second, $\mathsf{ID}^{\mathrm{i}}_{<\omega}$ employs only countable sets generated from very concrete, previously obtained, information. Thus, we are to conclude that

> [...] the above reductions demonstrate [that] the classical systems in question [viz, Π^1_1-CA_0 and various extensions] have an *alternative constructive justification* which does not require anything like belief in a pre-existing totality of subsets of $\mathbb{N}$ ([Feferman, 1987], 201).

Once more, I find this compelling (and the results incredibly striking). However, Feferman's arguments only retain their force under the assumption that the infinite is required for *provability*. But is infinitary reasoning necessary in a no less important sense for the provision of *perspicuous and explanatory* proofs?[148] My analysis of the ergodic proof of Szemerédi's theorem is intended to support an affirmative answer to this question. Though it is indeed the case that Szemerédi's theorem does not require anything, provability-wise, beyond finitary combinatorial reasoning, restricting ourselves to this alone damages our understanding of the result.

Let us now examine the metamathematics of the ergodic proof(s) of Szemerédi's theorem. The axiomatic strength of these proofs will be determined by the strength of the Furstenberg Structure Theorem (Theorem 2.3.37) and whether the full power of the theorem is required. There is a rather complicated story here. It was originally claimed that the Structure theorem is equivalent over ACA_0 to Π^1_1-CA_0 (Theorem 5.3 of [Avigad, 2009]).[149] However, a few years later, it was noted in [Montalbán, 2011] that Avigad and Towsner were comfortable asserting only the formalizability of the Structure Theorem in Π^1_1-CA_0. The reversal, i.e., that the Structure Theorem $+\,\mathsf{ACA}_0$ implies Π^1_1-CA_0 was—and remains—open. It is not ruled out that this reversal holds, but it would seem to require some delicate work to prove.[150] With these facts in mind, let us assess the status of Furstenberg's proofs. The crucial question is this: how far into the countable ordinals need the tower of factors of measure-preserving X extend? That is, need it extend arbitrarily far and thus require the *full* Furstenberg Structure Theorem? It is well-know that the original proof in [Furstenberg, 1977] did not;

[148] Indeed, one might think that the very intelligibility of a good deal of mathematics requires the infinite.

[149] I first wrote this section under the presumption that this equivalence was solid; however, after discovering [Montalbán, 2011] and corresponding with Jeremy Avigad, I have had to rely on more cautious claims.

[150] See Appendix B for further discussion.

on the other hand, the presentation in [Furstenberg et al., 1982] does appear to use the full Structure Theorem. Nonetheless, it has been shown in [Avigad and Towsner, 2010] that the tower need only extend to the $\omega^{\omega^{\omega}}$th level. The upshot of all this is that, when the Structure Theorem is used to prove Furstenberg Multiple Recurrence for all k, the proof goes slightly beyond the strength of ACA_0, and so is weakly infinitary.

Thus, the ergodic proof(s) in both [Furstenberg, 1977] and [Furstenberg et al., 1982] are then in principle much weaker than they appear because they do not *require* the full Furstenberg Structure Theorem. But this raises an interesting question: why didn't Furstenberg et. al. in [Furstenberg et al., 1982] simply avoid using this much power? As Avigad noted in personal correspondence, "I am sure they knew that it was possible. It would have changed the presentation only slightly: they could throw away the limit argument for the SZ property and the appeal to transfinite induction, and then they only needed to modify one of their calculations slightly." Thus, the answer is, probably, that they did not care to do so: the Structure Theorem is an incredibly interesting result and provides understanding as to why the ergodic analogue of Szemerédi's theorem holds. Why then dispense with it or try to whittle away at its logical strength if it provides a perspicuous proof?

I believe the following morals can be drawn. Even though the ergodic proofs are not as axiomatically strong as they appear, they are still much stronger than Szemerédi's original proof. We should emphasize the *relative distance*, which is nicely precisified by reverse mathematical analysis: Szemerédi's combinatorial proof is axiomatically weak but also incredibly difficult to understand.[151] On the other hand, Furstenberg's proof(s), especially that of [Furstenberg et al., 1982], takes us from RCA_0 (possibly even EFA) to just beyond ACA_0 and thus into the realm of the infinitary. In so doing, we get a perspicuous high-level proof of Szemerédi's theorem that emphasizes crucial structural features of the mathematics. Is this a consequence of the increase in axiomatic strength? I think the answer is yes but with some reservations; a full analysis of this point would require an independent discussion, and so I leave it at that for now.

In any case, the most perspicuous version of the ergodic proof[152] appears to use the full strength of the Structure Theorem, even though this is not strictly necessary for the proof. The clear expression of this structural result, effected by first examining the limiting behavior of measure-preserving systems, is a crucial ingredient of the ergodic proof of Szemerédi's theorem, and, I have claimed, it is the putative "reason why" the theorem holds. This high-powered result is, of course, not the only aspect of the ergodic proof that effects simplification and explanatory dividends over the combinatorial one, but it is the most important. It should, then, be considered necessary in the requisite sense.

Such an "explanationist" argument for the necessity of the transfinite also gains purchase against Feferman's second rejoinder. Even though it is proof-theoretically possible to provide constructive justification for Π^1_1-CA_0 (and other subsystems), in order to counter my argument it would be incumbent on Feferman to show that a constructive reduction of the infinitary elements of the ergodic proof generates an *equally perspicuous and explanatory* proof. I will not claim that this cannot be done, but I find it incredibly unlikely. One might surmise this is what Avigad attempts to do in [Avigad, 2009]. Here, in joint work with

[151] See Appendix B for further discussion.

[152] That of [Furstenberg et al., 1982].

Henry Towsner, Avigad provides a "metamathematical explanation" as to why the Furstenberg Structure Theorem (in particular, the Furstenberg tower constructed in the theorem) can be used to "prove a finitary combinatorial statement with explicit computational content." In particular, they sketch a strategy designed to show that the entire ergodic proof can be carried out in ID_1[153] and that this proof in ID_1 can be interpreted constructively (via a modification of Gödel's Dialectica interpretation). The ultimate goal is "to obtain a perspicuous new proof of Szemerédi's theorem, one that will clarify the combinatorial essence of the Furstenberg approach..." ([Avigad, 2009], 74). However, this is a goal *different* from giving an explanatory proof of Szemerédi's theorem *simpliciter*: it is designed to clarify the surprising infinitary intervention but not to demonstrate the reason why Szemerédi's theorem is true or even show why the ergodic setting is so helpful here.[154] Thus, Avigad and Towsner's aim, while very interesting, is somewhat orthogonal to my own. It would become much less so if a new proof was produced by their strategy that clarified the structural features underlying the theorem in a new way. But, again, I find this unlikely. My pessimism is warranted, in part, by the fact that no real simplifications of Szemerédi's original proof have been produced despite attempts by Tao and others. I would like to say that the ergodic setting is, in some sense, the "right" setting for proving Szemerédi's theorem: it yields the most significant epistemic advantages.[155]

Thus, one might offer an "explanationist" argument for Gödel's doctrine. That is, one might argue for the acceptance of stronger and stronger formal systems because they are necessary for producing explanatory proofs.[156] However, is this acceptance to be dictated solely by mathematical need as, say, a mathematical naturalist would have it? I do not find such a position terribly convincing,[157] so some qualifications are in order. First, if one is convinced by my argument via the ergodic proof, it has, at the very best, only bought us Π^1_1-CA_0. Though this is the strongest of the five most commonly studied subsystems of Z_2 and avowedly impredicative, it is a far cry from anything like the addition of very strong "higher types," e.g., the large cardinal axioms advocated by Gödel. Szemerédi's theorem might offer compelling reason (that is, more than invoking mere mathematical need) to accept the first stage of the transfinite because the theorem is itself eminently comprehensible: it involves assertions about the additive structure of natural numbers. However, it is not clear to me that the addition of something like large cardinal axioms, especially "large" large cardinals, can begin from such an epistemically distinguished starting point. Thus, I should want to say that the most philosophically significant explanationist arguments for Gödel's doctrine should begin, as Gödel himself does, from a restricted class of results we seek to explain, e.g., strictly finitary statements, and these might not take us very far into the infinite (see,

[153] See Example 2.6.1 above.

[154] This is what I have tried to do in Section 2.4.

[155] Furthermore, in personal correspondence, Avigad has noted that his project with Towsner encountered serious obstructions, and so the desired results, i.e., producing a proof in ID_1 and then applying Dialectica, are not forthcoming.

[156] More or less a way of "extrinsically" justifying strong axioms as Maddy seeks to do in [Maddy, 1998] and her later work. However, I find the whole mathematical naturalist project too quietistic. The following restrictions are intended to add some philosophical thrust behind extrinsic justification.

[157] I will not defend this view in detail here. On the whole, naturalism appears unstable as one cannot cleanly distinguish philosophical and mathematical concerns, so it does not make sense to appeal solely to mathematical need.

however, my reflections in the following section).

Interestingly, and somewhat surprisingly, this explanationist argument for infinitary mathematics has roots in Feferman's early work, viz., [Feferman, 1964]. Here Feferman (once more) seeks to undermine the "Cantorian" or "Platonistic" conception of sets in favor of some suitable interpretation of the "predicative" conception. Gödel is (once more) taken to task as the representative *par excellance* of platonism. In so doing, Feferman considers Gödel's argument that a platonistic construal of sets is "as legitimate" as the assumption of physical entities in developing an adequate theory of sense perception. Feferman says of Gödel's argument:

> The actual development of mathematics strongly supports one interpretation of this argument. Abstraction and generalization are constantly pursued as the means to reach *really satisfactory explanations* which account for scattered individual results. In particular, extensive developments in algebra and analysis seem *necessary* to give us real insight into the behavior of natural numbers. Thus we are able to realize certain results, whose instances can be finitistically checked, only by a detour via objects (such as ideals, analytic functions) which are much more "abstract" than those with which we are finally concerned ([Feferman, 1964], 3; emphases my own).

This remark is an excellent distillation of many of my points in this section and has served as an important influence on this chapter.[158] However, despite this initial assessment, Feferman converges on the views found in the later [Feferman, 1987] as he continues:

> The [above] argument is less forceful when it is read as justifying some particular conceptions and assumptions, namely those of impredicative set theory, as *formally necessary* to infer the arithmetical data of mathematics ([Feferman, 1964], 3).

Perhaps, then, whatever differences I have with Feferman are a matter of degree and not of kind.[159] His primary aim in [Feferman, 1964] was to show that much of mathematical practice is predicatively justified, while simultaneously acknowledging the explanatory significance of infinitary mathematics. My own aim is to carefully analyze this explanatory significance, while acknowledging the interest and success of Feferman's predicative program. Indeed, I find it hard to imagine, even though the talk of explanation does not feature in [Feferman, 1987], that Feferman would fail to recognize the explanatory import of the infinitary and ergodic techniques discussed above.

[158] This quote has not escaped Mancosu's attention. See [Mancosu, 2008a], p. 139 and Section 2.6.2 below. A remark that I believe may be interpreted along similar lines (as I mention at the end of Section 2.4.1) is made in [Isaacson, 1996]: "...there can be cases where the higher-order [infinitary] perspective is essential for *actual* conviction as to the truth of arithmetically expressed sentences" and, in a formal, proof-theoretic sense, "[t]he higher-order perspective can be essential, then, for shortening an otherwise unsurveyable proof" (221).

[159] Although when I come to ontological considerations, our positions will diverge significantly.

Another Case to Consider: Fermat's Last Theorem

Is it possible to justify principles even stronger than Π^1_1-CA_0 via the strategy outlined above? Can we find a theorem that, on its own, appears epistemically significant insofar as it is intuitively number-theoretic, making appeal to nothing beyond natural numbers, while requiring strong infinitary resources for an explanatory proof? A natural candidate[160] is Fermat's Last Theorem (FLT): for any positive integer $n > 2$, there are no positive integers x, y, z satisfying $x^n + y^n = z^n$. Famously, this was postulated by Pierre Fermat as a marginal annotation to Diophantus' *Arithmetica* and remained unproven until Wiles's monumental result of 1995.[161] The only proof of FLT thus far remains Wiles's and demands vast prerequisites. Indeed, there is an entire book of over 500 pages designed as prolegomena to Wiles's proof.[162] Thus, FLT is an example of a number-theoretic result requiring significant resources to prove it. Some natural questions for a logician or philosopher of mathematics are: what is the precise nature of these resources? Do we require all of them to prove FLT? Can the proof be reduced in some way? Can we make explicit its foundational assumptions?

A paper by Colin McLarty provides a summary of what is known about answers to these questions ([McLarty, 2010]). In particular, he asks whether the proof goes beyond ZFC or whether it requires, say, only PA.[163] The crux of the issue involves the use of *Grothendieck universes*: we say that an uncountable transitive set U is a Grothendieck universe (or simply universe) if:

1. for all $x \in U$, $\mathscr{P}(x) \in U$;
2. for all $x \in U$ and functions $f : x \to U$, $\bigcup_{i \in x} f(i) \in U$.

That is, for every set x in U, the powerset of x is in U, and, for every function from an element x of U to U, the range of this function is also in U. These amount to a strengthening of the familiar powerset and replacement axioms of ZFC, and so the use of a universe U goes beyond ZFC alone. In fact, it is a theorem due to N.H. Williams[164] that U is a Grothendieck universe iff, for some strongly inaccessible cardinal α, $U = V_\alpha$ (the collection of all sets of rank α).

Thus, our question then becomes: does the proof of FLT necessarily involve universes? In what sense are universes necessary? It is worth quoting McLarty at length here:

> For [Grothendieck] [large cardinals and thus universes] were merely legitimate means to something else. He wanted to organize explicit calculational arithmetic

[160] Indeed, Feferman remarks that various undecidability and incompleteness results of interest to mathematical logicians may not "...have any relevance to the classic unsettled problems that have challenged generations of number theorists. [...] How much different it would be if one showed that Fermat's 'Last Theorem' FLT is not provable in PA, or even more strikingly, in ZFC–and thus demonstrated why it's so difficult to prove FLT (if true)!" ([Feferman, 1987], 196).

[161] Which itself built upon some of the most profound mathematical techniques of the 20th century in number theory and algebraic geometry, e.g., the Taniyama-Shimura-Weil conjecture (restricted to the case of semistable elliptic curves).

[162] See [Cornell et al., 2000].

[163] Or, if Harvey Friedman is correct, even weaker systems like *Elementary Arithmetic* (EA). See [Avigad, 2003] for a nice discussion of Friedman's conjecture.

[164] See [Williams, 1969].

> into a geometric conceptual order. He found ways to do this in *cohomology* and used them to produce calculations which had eluded a decade of top mathematicians pursuing the Weil conjectures. He thereby produced the basis of most current algebraic geometry and not only the parts bearing on arithmetic. *His cohomology rests on universes but weaker foundations also suffice at the loss of some of the desired conceptual order* ([McLarty, 2010], 360-1, emphasis my own).

There are many technical details that McLarty, to his credit, tries to address quite carefully. I will briefly mention some of the main points for the sake of completeness.[165] First, it has long been known by algebraic geometers (McLarty cites Deligne in particular), that universes may always be eliminated for ZFC alone. However, this is *never* done in the "great cohomological proofs" because of the epistemic gains of using universes. I return to this point below. Second, McLarty mentions work by Angus Macintyre[166] attempting to show that the Modularity Theorem[167] is expressible as a Π^0_1 statement of PA and can be proved in PA. Much of this involves replacing various analytical and topological structures (as completions of $\mathbb{Z}$ or $\mathbb{Q}$) with finite approximations in PA. As one might expect, even in its early stages, this work renders the already long and difficult proof of FLT even longer and more difficult. McLarty remarks that

> For the foreseeable future it is likely that any proofs of FLT to be found in weak theories of arithmetic will be discovered in the first place, and will be comprehensible after they are discovered, only by applying metatheorems to some shorter known proof using stronger logic. In this context [Wiles's proof] counts as a short proof ([McLarty, 2010], 364).

Thus, the case of FLT is similar to that of Szemerédi's theorem: in both we have theorems with explicit finitary content whose proofs appeal to infinitary techniques, and these techniques provide some sort of epistemic gain. However, are the philosophical conclusions to be drawn from each case the same? Perhaps not, as there are some important disanalogies. First, there is only one proof of FLT and it uses Grothendieck universes as a matter of fact. Szemerédi's theorem, on the other hand, has multiple proofs; in particular, its combinatorial proof makes no appeal to infinitary techniques. Second, the appeal to universes in the proof of FLT catapults us into the higher infinite (equivalent to the existence of a strongly inaccessible cardinal), whereas the use of Π^1_1-CA_0 in the ergodic proof of Szemerédi's theorem is a mere first step into impredicativity. Can my explanationist argument for Gödel's doctrine buy much stronger set-theoretical principles than I first claimed? And is the case of FLT more convincing because the *only* proof we now possess makes appeal to universes (even though, in principle, this can be weakened)? I must confess that I am not entirely sure of the answers to these questions. McLarty argues, along with many preeminent mathematicians, that Grothendieck universes provide a " 'systematic means' of presenting and proving results" and stresses the "practical value of Grothendieck's high level organization..." (363;

[165] If uninterested, one may skip to the next paragraph.

[166] See [Macintyre, 2011].

[167] This was proved in 2001 by Breuil, Conrad, Diamond, and Taylor. Wiles's proof of FLT uses a restricted version of this theorem (merely a conjecture in 1995), which implies FLT.

372). Obviously, this is all true. But from a philosophical perspective is appeal to the practical value of universes in the proof of FLT[168] sufficient to say that they are necessary in any sense? It seems to me that the organizational power adverted to in the proof of FLT is thinner than the explanatory power afforded by ergodic techniques in Szemerédi's theorem. In particular, it is entirely possible that, though universes might yield a sort of "simplicity" in virtue of their structuring of the proof of FLT, this simplicity need not produce an explanatory proof generative of understanding. I think, then, that one would have to carefully analyze what happens in the proof of FLT when universes are exchanged for something weaker, say, ZFC. It may be that this seriously damages the conceptual flow of the proof; it may not. In any case, pending such an analysis, I would be hesitant to claim that the same conclusions concerning Szemerédi's theorem hold for FLT, viz., that the infinitary resources, at least those going beyond ZFC, are necessary in any sense.

I have proposed an argument for a weakened version of Gödel's doctrine and have suggested that we apply such an argument to an epistemically distinguished class of statements:[169] more or less, those involving assertions about the natural numbers alone. Let me remark upon both (i) the weakening and (ii) the class of statements from which I start.

Regarding (i): the argument is weakened because the necessity in question is not "necessary for proof," but rather "necessary for explanation." I do not take this to be significantly weaker: it would seem that one wants their mathematical theories to be maximally intelligible or generative of understanding almost as much as one wants proofs of theorems. Proofs without intelligibility are not worth terribly much.[170] Regarding (ii): my overall argument looks as though it might advocate for a metaphilosophical view in the vicinity of mathematical naturalism. That is, it may look friendly to Maddy's assertion that, "the grounds on which to criticize and/or justify mathematical methods are to be found in mathematical practice itself; the practice need not answer to, nor can it look for support from, any external standard" ([Maddy, 1998], 136). I find many features of this view troubling.[171] Thus, in order to block wholesale appeal to "whatever mathematics (including set theory) needs," I have started from strictly number-theoretic statements, taken to be epistemically privileged in some way.[172] What do we need to prove these in an explanatory fashion? More than we might have initially thought: Π^1_1-CA_0 and perhaps even the existence of strongly inaccessible cardinals. It would be interesting to see if this argument could be extended even further, but this will have to be done elsewhere.

[168] Of course, in this paper I have only considered individual theorems and their proofs. I have not considered the construction of mathematical theories. It may be (and is probably) the case that the incredible fruitfulness of Grothendieck's cohomological techniques and the attendant use of universes is sufficient to argue for the necessity of universes on a larger scale. However, this is an incredibly complicated affair and cannot be addressed here.

[169] As Gödel himself did, though he did not stress this point.

[170] Indeed, consider the recent issues surrounding Mochizuki's supposed proof of the *abc* Conjecture.

[171] See fn. 157 above.

[172] Here are some well-worn, though to my mind convincing, reasons for thinking so: (1) they have determinate truth values; (2) their axioms are relatively self-evident; (3) we have some geometric intuition about them; (4) statements about the natural numbers are the minimal, non-trivial mathematics we need to get our mathematical theories off the ground.

2.6.2 Explanationist Indispensability Arguments

Let me return to Feferman's starting point: platonism as the "medieval metaphysics" of mathematics. My explanationist argument for Gödel's doctrine has so far been strictly epistemological.[173] If one is convinced by it, then I have provided grounds for something like the rational acceptance of infinitary mathematics up to the strength of Π^1_1-CA_0 and perhaps strongly inaccessible cardinals. Here, unfortunately, the fact that we do not have the equivalence of the Furstenberg Structure Theorem with Π^1_1-CA_0 over ACA_0 becomes crucial. The epistemic claims above are not so affected; however, if we think that mathematical platonism only enters the picture when the full powerset $\mathscr{P}(\mathbb{N})$ is involved, then we must be dealing with Π^1_1-CA_0 for this to be true. Thus, the following discussion must be taken as conditional upon the truth of Avigad's claims in [Avigad, 2009], which have not yet been established. Of course, these worries do not affect the discussion of FLT.

We might frame the situation in terms of Kitcher's scheme from [Kitcher, 1984]: begin with a mathematical practice, (L, M, Q, R, S), a quintuple of our mathematical language, metamathematics, questions, reasonings, and statements of a restricted nature. Then, we can proceed to an expanded practice, (L', M', Q', R', S'), *without epistemic debt* by adjoining the mathematical and metamathematical resources needed to explanatorily prove theorems in the original class of statements S. However, it is then reasonable to ask about the referents of the new terms in L' and M' (from which the new statements in Q', R', and S' are partially constructed). Can our epistemological picture be grounded in a convincing metaphysics? I would like to suggest that, contra Feferman, one can provide an argument for (restricted) mathematical platonism via an "explanationist" indispensability argument.[174]

This form of argument would be analogous to one now familiar in the literature[175] from [Baker, 2005]. He takes as his starting point the classical Quine-Putnam (QP) indispensability argument for platonism. This runs as follows:

1. Mathematical entities are indispensable for our best scientific theories;
2. We ought to believe our best scientific theories;
3. Therefore, we ought to be committed to the entities these theories quantify over, including mathematical entities.

This argument has been the subject of much discussion, but one feature is especially relevant here: its holistic conception of scientific theories. That is, ontological commitment is determined by *all* existentially quantified sentences entailed by the theories. One might find this unsatisfying as holism is neglectful of the different roles that particular posits play internal to the theory. Thus, Baker develops a variety of indispensability argument that does not rely on holism.[176] He proposes that we ought to be committed only to those entities that play an explanatory role in a given theory. Such a move is suggested by the fact that, in order for

[173] And thus my disagreement with Feferman has been relatively minimal. This now changes.

[174] This possibility was, to my knowledge, first raised by Mancosu in [Mancosu, 2008a]. He suggests that the argument quoted above from [Feferman, 1964], p. 3 is "...something in the vicinity of an indispensability argument" ([Mancosu, 2008a], 140).

[175] Suggestions in a similar vein can also be found in the much earlier article by Steiner [Steiner, 1978].

[176] Picking up on a debate between Melia and Colyvan. See the references provided in [Baker, 2005].

any debate about indispensability to get started, one must endorse some version of inference to the best explanation (IBE). Given that Baker is concerned with *external* applications of mathematics, i.e., mathematics applied to natural science, he must establish that there are genuine mathematical explanations of physical facts. Much of [Baker, 2005] is concerned with doing just this via a case study from evolutionary biology. The case study involves the life cycles of the "periodic" cicada, which are invariably prime.[177] Baker claims that a number-theoretic result, i.e., that prime periods minimize intersection relative to non-prime periods, is essential to the explanation that periodic cicadas have life cycles of only thirteen and seventeen years. In short, this mathematical result tells us why it is that prime periods are "evolutionarily advantageous": minimization of intersections helps the periodic cicadas avoid predators and deleterious interbreeding with other insects.

Thus, an "explanationist" indispensability argument takes the following form:

1. There are genuine mathematical explanations of physical facts;
2. We ought to be committed to the existence of explanatory entities;
3. Therefore, we ought to be committed to the existence of the mathematical entities posited in the mathematical explanation of the physical fact in question.

Though such an indispensability argument seems an improvement over the "holistic" QP version, there are many questions to be asked.[178] However, in order not to lose focus, let me move on to the positive proposal. It is important to note that, up until now, all versions of the indispensability argument argue from physical phenomena to mathematical platonism. Broadly speaking, these arguments address those who are realists about scientific or naturalistic entities, but are not necessarily mathematical realists. Then, by dint of either holism or explanatory power, the mathematical entities are shown to be on an ontological par with the scientific entities, thus establishing platonism. One might wonder why this form of argument, in particular, that of the explanationist indispensability argument, has not been applied *internal* to mathematics itself.[179] Perhaps an obvious first answer is that doing so would be question begging: as Leng says, "[...] in the context of an argument for realism about mathematics, it [an explanationist indispensability argument] is question begging. For we also assume here that genuine explanations must have a true explanandum, and when the explanandum is mathematical, its truth will also be in question" ([Leng, 2005], 174).

As Mancosu notes, this worry reflects

> [...] the general use to which indispensability arguments have been put. The main goal is to provide an argument for platonism in mathematics but no attention is truly given to the different kind of mathematical entities we are postulating. From this point of view the existence of the natural numbers is on par with the existence of a Mahlo cardinal or of a differentiable manifold. It is, however, reasonable to ask whether mathematical explanations can be used not as arguments for realism in mathematics *tout court* but rather as specific arguments for realism about certain mathematical entities ([Mancosu, 2008a], 139).

[177] Thirteen or seventeen years depending on geographical location.
[178] See Leng's fictionalist rejoinder in [Leng, 2005]. See also the volume [Leibowitz and Sinclair, 2016].
[179] This point is raised in [Leng, 2005].

The suggestion, then, is that we may apply an explanationist indispensability argument internal to mathematics provided that we are not arguing for mathematical realism *simpliciter*. Rather, the idea would be to argue from a restricted realist position, say, realism about the natural numbers, to a more expansive realism encompassing the mathematical entities needed to prove results about the naturals in an explanatory fashion. This would defeat the circularity complaint. Furthermore, as Mancosu mentions, one would also have to address a realist innocent of particular foundational commitments. This is because some foundational positions (e.g., constructivism, predicativism) might immediately deem the theoretical posits in the argument illicit. I do not think the latter restriction is too severe; however, one might worry about finding realists about $\mathbb{N}$ who are not dyed-in-the-wool realists *simpliciter*. I briefly discuss this below.

Explicitly, then, such an argument would look like:

1. We are committed to the existence of a restricted domain of mathematical entities, say, $\mathbb{N}$;

2. There are genuine mathematical explanations of facts about $\mathbb{N}$;

3. We ought to be committed to the existence of explanatory entities;

4. Therefore, we ought to be committed to the existence of the explanatory entities posited in the mathematical explanations of the facts about $\mathbb{N}$.

Of course, one would now like particular cases to validate this line of argumentation. But I have already provided two examples of number-theoretic facts that require vastly different mathematical entities to prove them in an explanatory fashion: Szemerédi's theorem and (perhaps) Fermat's Last Theorem. Thus, if one is a realist about the naturals, one should also be a realist about the explanatory entities involved in the Furstenberg proof of Szemerédi's theorem via the above indispensability argument.[180]

An immediate unclarity arises: what, exactly, are the explanatory entities? This question arises because we have dropped the commitment to Quinean holism; once this is done, we must more carefully delineate those mathematical posits that are "essential" to the explanation.[181] I have tried to demonstrate in the above sections that ascertaining what constitutes an explanatory proof is no easy matter. In the Szemerédi case, we have the Furstenberg Structure Theorem, which gives the reason why Szemerédi's theorem is true. However, this is not the only mathematical result that contributes to the explanatory power of the ergodic proof: many other considerations come into play. Thus, I think one could adopt either a *conservative* or *permissive* interpretation of the above indispensability argument. The conservative, if pressed, would probably have to say that *only* the entities serving as referents for terms in the putative reason why (e.g., the Furstenberg Structure Theorem) merit ontological commitment. On the other hand, the permissive interpretation would accept the existence of all entities serving as referents in the proof of the explanandum (Szemerédi's

[180] And perhaps a realist about Grothendieck universes. I will discuss only the Szemerédi case in what follows as I have done the requisite work to show that its ergodic proof is indeed explanatory.

[181] This issue arises in Baker's case study as well. Should we be ontologically committed to only the primes 13 and 17? All primes? All resources needed to prove the theorem about primes invoked?

theorem). It is possible that these positions may not come apart all that much. For example, if one is a conservative, the Szemerédi case has bought you the Structure Theorem. This involves quite a bit of ergodic theory (e.g., arbitrary measure-preserving systems, transfinitely many compact factors). It then seems reasonable to iterate the indispensability argument: what further resources does one need to prove facts about measure-preserving systems in an explanatory fashion? About compact factors of measure-preserving systems? This can be extended *ad infinitum*, and it is entirely possible that vast swathes of the contemporary mathematical landscape could be recovered for the platonist. The permissive theorist could also do the same; the difference would be that they would recover more mathematics more quickly than the conservative.

I doubt there is any clean way to adjudicate between these two interpretations (and, again, it may not matter much in the end), but favoring one or the other will depend upon a careful analysis of the putative explanandum. In other cases, the "reason why" may be much less perspicuous than in the ergodic proof of Szemerédi's theorem, and so a conservative interpretation may not be available. On the other hand, a permissive interpretation might feel unnecessary; however, in order to block wholesale commitment to all entities featuring in an explanatory proof of the putative explanandum much unpacking of the proof will be necessary. In the case of FLT, for instance, one would have to exchange the use of universes for merely ZFC (or even PA) to see whether universes are essential for the proof to be explanatory (if it is such). I say all this to indicate that there is unlikely to be any *a priori* desiderata that will tell us how to interpret the indispensability argument.

Let me close by mentioning two further issues. The first involves, once more, a question about content. Let us assume that the Furstenberg Structure Theorem is equivalent to Π^1_1-CA_0 (over ACA_0). Obviously, these statements, though equivalent over a weak subsystem, are very different: the first involves ergodic terms; the latter is simply a set-existence principle. The question is: should we also be ontologically committed to the set-existence principle given that it is equivalent to an explanatorily indispensable theorem? This opens up a Pandora's box of further issues about mathematical content; in particular, is the set-existence principle part of the content of the Structure Theorem? If so, what kind of content? Do answers to these questions depend on translations between sets and "ordinary" mathematics? What guarantees that the translation of the Structure Theorem into the formal language of second-order arithmetic preserves properties we are interested in? I cannot take these up here, but I find them interesting questions.[182] My initial response to the question about commitment to the set-existence principle would be broadly Dedekindean in style: just as Dedekind stressed that we should not *identify* real numbers (or rational numbers) with Dedekind cuts, we should not identify the Structure Theorem with the set-existence principle. Thus, perhaps we should not be ontologically committed to the bare set existence principle. However, the highly uncomputable nature of the tower of extensions involved in the Structure Theorem is actually very important for the explanatory power of the ergodic proof. The fact that we can study measure-preserving systems in the limit eliminates the need to deal with much combinatorial manipulation that muddies our understanding of the pure proof of Szemerédi's theorem. This was one upshot of my discussion of impurity, simplicity, and explanatory power in Section 2.5. Thus, at least for the case at hand, we

[182]See Chapter 3 of [Eastaugh, 2015] for a (somewhat) related discussion.

should be ontologically committed to both the Structure Theorem and the bare set-existence principle as the latter encodes the complexity required to explanatorily prove Szemerédi's theorem. Again, as is true for the conservative-permissive construal of the indispensability argument, judgments like this will depend upon a careful analysis of the results in question.

Second, the above indispensability argument is, by design, not an argument for mathematical realism *simpliciter*. We have addressed "number-theoretic platonists" in order to avoid circularity. Thus, one would like an argument for a realist construal of $\mathbb{N}$. This is certainly no easy task, but seems a more reasonable one than attempting to argue for a platonist construal of all of mathematics at once. For instance, a convincing argument for the reality of $\mathbb{N}$ might be provided via appeal to our geometric or visual apprehension of the natural number structure or perhaps via a linguistic argument from the recursivity of natural language. Or, if one is sanguine about the prospects of a neo-Fregean justification of arithmetic, appeal to *Frege's Theorem*: Hume's Principle[183] + Second-Order Logic + Suitable Definitions $\vdash \mathsf{PA}_2$ (second-order Peano arithmetic). The success of this strategy will, of course, depend on a justification of Hume's Principle and, for those skeptical of the power of full second-order logic, may involve restrictions on its comprehension scheme. In any case, by addressing a number-theoretic platonist, we have a more accessible stepping stone to mathematical realism *simpliciter*.

2.7 Impurity, Unification, and Explanation

2.7.1 Introduction

Finally, I aim to demonstrate that intuitively impure resources generate explanatory proofs via unification. That is, at least in some cases, we have the following chain of inferences: Impurity $\Rightarrow$ Unification $\Rightarrow$ Explanatory Power. I take it that, by now, the impurity present is clear, and so I will focus on making the second inference precise. As mentioned in the introduction to this chapter, the idea that explanation can be analyzed through unification has been pervasive and influential in the literature on scientific explanation. And, despite the failure of formal models of unification proposed by Friedman and Kitcher, it would seem that unification has something to do with explanation in mathematics.[184] In the case study from [Hafner and Mancosu, 2008], where it is shown that Kitcher's model fails to make sense of mathematical practice,[185] Hafner and Mancosu note

> There is a certain irony in this [failure] since Brumfiel [the mathematician in question] champions a kind of unification of real algebraic geometry by insisting on proofs that exhibit a 'natural' explanatory uniformity. Yet, despite its focus

[183] Where Hume's Principle is a second-order abstraction principle of the form: $\forall X, Y(N(X) = N(Y) \leftrightarrow \exists \text{bijection } f : X \to Y)$, where X, Y are second-order entities, and $N(\cdot)$ is the "number of" or "cardinality" operator that lowers the type of second-order entities to first-order entities.

[184] Lange also takes up this in Ch. 8 of [Lange, 2017]. Surprisingly, he makes little use of the literature I cite here (and has little interest, it would seem, in impurity). Thus, I take my work here to be an examination of the same question, but from a different perspective.

[185] In particular, Kitcher's model ranks a decision procedure as most unificatory (and thus most explanatory). Brumfiel unequivocally rejects this proof procedure as explanatory, preferring instead a non-elementary method applicable to all real closed fields.

> on unification Kitcher's account of explanation apparently does not have the resources to provide insight into the controversy over the 'right' proof methods or at least enhance our understanding of Brumfiel's motivations (170).

Despite taking great pains to improve Friedman's more brute unificatory model which focuses on the mere reduction of independent phenomena, Kitcher falls prey to the same, apparently mistaken, intuition that explanatory power and unification can be analyzed by quantitative comparisons alone. Instead, as the case study from real algebraic geometry makes clear, adjudicating between proof methods depends on various qualitative comparisons. Hafner and Mancosu conclude

> [...] even under the assumption that an account of explanation as unification is, in principle, on the right track, Kitcher's model doesn't tell the whole story yet. In general there is more to explanation than unification in Kitcher's sense, a more fine-grained analysis of different types of unification seems to be needed (170-1).

Thus, one of my aims is to understand what unification might mean in practice.[186] To that end, I turn to Margaret Morrison's book, *Unifying Scientific Theories* ([Morrison, 2000]), which is explicitly engaged with this question in the sciences. I believe that various distinctions made there can be carried over to the mathematical case and that these distinctions will help to further the debate concerning unification and mathematical explanation. I ultimately conclude that, although unification (suitably understood) may not be associated with explanation in the sciences (as Morrison argues), the same cannot be said for mathematics. Indeed, this deepens the irony present in Kitcher's work. One of his goals in providing a theory of explanation as unification was to elucidate systematic continuities between the sciences and mathematics; however, one upshot of my discussion is that there may in fact be systematic *discontinuities*.

I also examine Morrison's diagnosis of why explanation and unification often come apart in the sciences (what I have been calling "Morrison's thesis"). One might worry that this issue, like her typology of unification, easily carries over to mathematics. I show that, at least in the case of Szemerédi's theorem (as well as in the case of theorems included in the genus of results dependent upon the structure-randomness dichotomy), this does not hold. In particular, my account of structural content is serves as a preventative against the mathematical analogue of Morrison's thesis. This is likely too strong a condition to impose, i.e., requiring shared structural content may block genuine cases in which unification yields explanatory power. Nonetheless, this condition serves as a helpful starting point in analyzing how impurity, unification, and explanation may be associated.

2.7.2 Morrison's Thesis

In *Unifying Scientific Theories*, Morrison shows quite convincingly that explanation and unification come apart in the natural sciences. Through a careful consideration of detailed case studies from both biology and physics, Morrison concludes that the mathematical techniques that facilitate unification in the sciences are not those techniques that actually explain

[186] This is analogous to my discussion of simplicity.

why a particular phenomenon occurs. Morrison argues that the latter inevitably involve the "machinery" or "mechanisms" of a system, i.e., the system's causal behavior, and the mathematical apparatus abstracts from these details (Morrison's thesis).

This understanding of the relationship (or lack thereof) between explanation and unification proceeds from an analysis of the unity of theories in the sciences. The point is to "provide an analysis of how [unity] is achieved and how it functions" (1). We find very quickly that unification/unity is said in many ways and is a much more complicated notion than has been appreciated in the philosophy of science literature. One especially salient aspect of theoretical unity is the following:

> [...] truly unified theories display a particular feature in virtue of which the phenomena are joined together, enabling diverse phenomena to be combined into a single theoretical framework. It is this combining of phenomena through a particular parameter in the theoretical structure that constitutes an important part of the unifying process, a process that is represented in the mathematical framework of the theory [...] I want to argue that in true cases of unification we have a mechanism or parameter represented in the theory that fulfills the role of a necessary condition required for seeing the connection among phenomena (32).

For example, in the third chapter, Morrison argues that a theoretical parameter called the displacement current is a necessary condition for Maxwell's field equations. This is because it enables us to make sense of the notion of a quantum of electricity crossing some boundary without which we could not formulate the notion of a field. These field equations in turn show that both electromagnetic and optical processes are the results of waves travelling through space, and thus Maxwell's theory is an instance of a unificatory theory. However, genuine "theoretical understanding" was absent from this formulation:

> [Maxwell's theory] provides very accurate descriptions of the behavior of optical and electromagnetic processes using the field equations, yet there was no explanation of just what the field consisted in or how these waves could be propagated through space (107).

I am in general agreement with Morrison's account. Indeed, precisely because the mathematics present in scientific theories often serves to *represent* physical phenomena, it is quite natural that these mathematical structures should fail to capture all the information sufficient for explanation.[187] An important question for us is: does the same sort of representational gap occur internal to mathematics? This is certainly a very difficult question and relates to my concerns about ontological commitment in indispensability arguments.[188] For instance, when translating an ergodic result into the formal language of second-order arithmetic some information is levelled out. However, for the purposes of this discussion, we need only consider a narrower question: do the techniques that enable unification internal to mathematics cause the loss of explanatory data?

That is, I would like to examine whether Morrison's thesis about unification and explanation holds in the mathematical setting. My claim is that it does not (or at least not

[187] I don't mean to be glib here. Surely, this is a very complex question, but I find myself relatively on board with most of Morrison's account.

[188] See the first issue raised at the end of Section 2.6.2.

uniformly) and that the ergodic proof of Szemerédi's theorem will help us to see why. First, a few preliminary remarks: given that we are concerned solely with mathematics, we will not be able to avail ourselves of the same conceptual explications. Certainly, we cannot claim that explanation in mathematics has anything to do with causal mechanisms. It is still worth asking, though, whether the "lower-level" content (i.e., the explicit, combinatorial data involved in the pure proof of Szemerédi's theorem), which we abstract away when passing to the ergodic setting, plays a role analogous to these physical processes. One might put this in the following somewhat hand-wavy fashion: how much does the "lower-level" content of the proof of a theorem contribute to the reason why the theorem is true? For example, in Szemerédi's theorem, is it essential to know, explicitly *for each* $A \subset \mathbb{Z}$ under consideration (resp. system in Furstenberg Multiple Recurrence), that it has a particular classification (structured, random, somewhere in between) in order to see why the theorem holds? Or can we pass to a more abstract context in which we lose this information but retain an explanatory proof? Finally, what does any of this have to do with unification? Hopefully, I can provide some answers to these questions, at least in a restricted setting.

2.7.3 What Do We Mean By Unification and What Kinds Generate Explanatory Proofs?

First, since I propose that unification can be analyzed as a special case of explanation in mathematics, let me briefly classify what is meant by unification. After doing so, I will demonstrate why the unification I have in mind serves to provide an explanation of Szemerédi's theorem. As Morrison does in the scientific case, it is possible to understand unity as a property belonging to theories, where theories are taken in a "naïve" or intuitive sense. For instance, just as she takes Newtonian mechanics and Maxwellian electromagnetism to be theories, I take combinatorics and ergodic theory in the same way.[189] However, in the mathematical case, we might also consider unity to be a property of particular theorems: a theorem may serve to unify because it collects many cases into a single classificatory result. Both unity as a property of theories or, more loosely, frameworks, and as a property of theorems will feature in our discussion.

How, exactly, are we to understand the properties in question? Morrison offers a helpful distinction:

> I distinguish two different types of unity: reductive unity, where two phenomena are identified as being of the same kind (electromagnetic and optical processes), and synthetic unity, which involves the integration of two separate processes under one theory (the unification of electromagnetism and the weak force) (5).

The reductive type of unity is the one most familiar from the literature.[190] Indeed, the idea that explanation and unification go together in the sciences proceeds from the following in-

[189]There are many questions concerning this intuitive reading of what a theory is: most basically, how does one circumscribe the content of such theories, especially given the inter-theoretic penetration with which I am concerned in this paper? I have addressed this to some extent in Section 2.4: I believe that we should avail ourselves of an "intuitive" reading of the content of theories (as with theorems), but be aware of the fact that there are high-level relational properties shared by intuitively distinct theories.

[190]See [Friedman, 1974].

sight: unification serves to reduce the number of brute facts we must countenance. The fewer the brute facts, the more comprehensible the world becomes, thereby increasing our understanding of the world. Finally, if we think there is no explanation without understanding, the identification of reductive unification and explanation seems quite natural. For example, Friedman claims that Newtonian mechanics served to effect this kind of unification by showing us that the laws of both celestial and terrestrial phenomena could be derived from Newton's laws.[191] We move from two distinct sets of phenomena to one, and thus increase our understanding of the world.

Synthetic unity, on the other hand, is evident in Maxwell's development of the electromagnetic field via his field equations.[192] The electromagnetic field does not reduce electricity and magnetism to one force, but rather shows the relationships between the two: "where a varying electric field exists, there is also a varying magnetic field induced at right angles, and vice versa" (107). This different understanding of unification is particularly interesting because it "calls into question the traditional relationship between unified theories and theoretical reduction," which is taken to hold between *reductive* unity and explanation (as in [Friedman, 1974]). I find this very helpful as I do not think that unification construed as mere quantitative reduction can get us very far. Let me then try to give a more nuanced account of unification, inspired by Morrison, in the mathematical context utilizing the ergodic proof of Szemerédi's theorem.

Here is a high-level overview of what we have seen: Szemerédi's theorem (Theorem 2.3.4) concerns the existence of arithmetic progressions in sufficiently dense subsets of the integers. This theorem can be proved by pure means, i.e., using techniques that intuitively belong to combinatorics and do not appeal to infinitary methods. This proof is quite subtle and ultimately turns on a Structure Theorem (Szemerédi Regularity Lemma; Lemma 2.5.3, Lemma A.1), i.e., a result that allows us to classify the behavior of the subsets and deduce the existence of arithmetic progressions. On the other hand, we can show that Szemerédi's theorem is equivalent to a result in ergodic theory, Furstenberg Multiple Recurrence (Theorem 2.3.12). Our strategy is then to prove Multiple Recurrence: this proof also turns on a Structure Theorem (Furstenberg Structure Theorem; Theorem 2.3.37), which plays a role similar to the Regularity Lemma in an ergodic context. In brief, it allows us to decompose any measure-preserving system into structured and random parts and then deduce the presence of recurrence properties. Thus, via our detour through ergodic theory and infinitary techniques, we get another proof of Szemerédi's theorem.

How, then, do we understand the unification, if it is such, taking place via this detour? We have two theorems T_C (combinatorial; Szemerédi's theorem) and T_E (ergodic; Furstenberg Multiple Recurrence), which we can now say both belong to a class of theorems $\mathscr{T}$. This $\mathscr{T}$ contains all "Recurrence Theorems," i.e., those theorems about the existence of various recurrence properties of various objects (sets of integers, topological spaces, measure preserving systems) in different contexts.[193] Another way of classifying $\mathscr{T}$ is to say that it is the class of theorems that all essentially involve a Structure Theorem, i.e., a theorem that

[191] Though this analysis is somewhat questionable. I believe that the distinction between reductive and synthetic unity is not always so clear-cut.

[192] Morrison also considers how Maxwell's field equations might be construed as reductive insofar as optical phenomena were reduced to electromagnetic phenomena.

[193] Van der Waerden's theorem (Theorem 2.3.7) and its topological analogue would also be members of $\mathscr{T}$.

shows us how to decompose the various objects in question into a structured and random part (because the recurrence property asserted by each theorem in $\mathscr{T}$ can only be proven by appeal to a Structure Theorem). One way to understand this situation $T_C, T_E \in \mathscr{T}$ is *reductively*: we have shown two *prima facie* very different theorems to be *equivalent*, and thus they belong to $\mathscr{T}$ trivially. I believe this explicit equivalence renders the reductive type of unity especially evident in our case. Nonetheless, it is not the mere reduction of two phenomena to one that generates understanding here but rather the comparative potential of the two contexts (combinatorial and ergodic) subsumed under one framework. If we simply reduce T_C to T_E we have not really gained any insight into why T_C is true in the combinatorial setting; the reduction *ignores* that setting. Perhaps if we had no other means of proving T_C, then we could claim the reduction is what drives the explanation, but that is not what occurs.

We can also consider this situation as an instance of synthetic unification: we have two different theorems, each of which may be independently proved via its own Structure Theorem. However, these theorems also belong to one mathematical genus, that delineated by their shared structural content, i.e., the entities intuitively involved in both theorems instantiate the dichotomy between structure and randomness (see Section 2.4.2). If we consider T_C and T_E under a single framework (a framework of structure theorems), then we get illuminating inter-theoretic comparisons. Though it did not happen this way, we might imagine that Szemerédi was not able to provide a pure, combinatorial proof of this theorem, in particular because he did not see the exact role the Regularity Lemma was to play. Instead, imagine Furstenberg first showed that Szemerédi's theorem was equivalent to Furstenberg Multiple Recurrence (which one may do without having an explicit proof of either). Then, because the Furstenberg Structure Theorem exhibits the crucial dichotomy between structure and randomness in a particularly explicit way, one might import this insight to the Regularity Lemma and achieve a better understanding as to why Szemerédi's theorem is true.[194] The foregoing case study was, in a way, an exercise in this fictional scenario: I have tried to show that the ergodic proof generates understanding in a way that the combinatorial proof does not and proceeded from the impure, ergodic proof to the pure, combinatorial one.[195]

Thus, the unification here is Janus-faced. It seems entirely plausible to think of it as both reductive and synthetic. We have that Szemerédi's theorem and Furstenberg Multiple Recurrence are equivalent, and thus we collapse two theorems to be proved into one, which is then proved via the Furstenberg Structure Theorem.[196] However, the synthetic variety of unification is also present: precisely because we have two theorems which depend on Structure Theorems, we are able to avail ourselves of inter-theoretic relationships. And now we come to my ultimate claim: these inter-theoretic connections are *explanatory* in the case

[194]However, the inter-theoretic comparison may not provide a clear way forward in actually proving Szemerédi's theorem.

[195]As in Section 2.4, the Prime Number Theorem (PNT) supports my case here. I briefly described how the pure proof of the PNT, though undoubtedly ingenious, took as its starting point the analytic properties crucial to the impure proof and deduced arithmetical analogues. This might also be thought of as an instance of synthetic unification.

[196]Of course, one could also proceed from Szemerédi Regularity to Szemerédi's theorem to Furstenberg Multiple Recurrence, *but not in an explanatory fashion.*

at hand. Thus, the synthetic unification achieved, i.e., seeing that T_C and T_E are intuitively independent, and yet explicable by one overarching framework, drives the explanation of T_C.

There are further distinctions to be made. I have been considering explanatory power in a "local" context, i.e., explanatory power is a property of particular proofs. One might also consider explanatory power to be a property of larger theoretical frameworks. Similarly, one can consider unification as a property either of proofs or of wider theoretical contexts. This distinction, obviously, does not arise for Morrison: she does not deal with local, circumscribed proof contexts. I find it likely that there are complicated relationships between local-global distinctions: for instance even though one theory may be explanatory on the whole, particular proofs using its theoretical apparatus may fail to be explanatory. It may also be the case that, though one has a "local" unity, one fails to get an explanatory proof. The story of the dynamics of these different distinctions is undoubtedly a long one to tell. Let me see what can be done with our case.

Internal to both the combinatorial and ergodic proofs we have a "local" sort of unity. This is precisely what the Structure Theorems, in their respective contexts, do. For instance, the Furstenberg Structure Theorem allows us to take *any* measure-preserving system and, in a flash, as it were, decompose it into structured (compact) and random (weak-mixing) components. Similarly, the Szemerédi Regularity Lemma allows us to take any $A \subset \mathbb{Z}$ and break it up into structured (periodic) and random subsets.[197] Furthermore, because we know that the relevant patterns are present in structured and random components (of both sets of integers and measure-preserving systems), we get proofs of Szemerédi's theorem and Furstenberg Multiple Recurrence. Thus, "local" unity in this case, i.e., unity internal to a particular proof, is precisely Morrison's reductive unity. Instead of having to consider, say, *all* the different subsets of a particular A, we need only consider a representative sufficiently dense A. Thus, we need not dig into the "lower-level" content of each particular entity to understand its behavior. We do, however, need to understand why the extreme cases (structured and random) exhibit the patterns in question; however, this is a very modest contribution of "lower-level" content, of a different order of magnitude than what Morrison believes is essential for any explanation in the sciences, viz., the causal behavior of entire physical systems. In any event, we have a two-fold reductive unity present in the Szemerédi case, at both a local and global level. Reductive unity can thus occur at different levels of theoretical inquiry in mathematics.

The following kinds of unity are then present in our case study:

1. Local Reductive Unity (internal to particular proofs; many cases collected under one classification);

2. Global Reductive Unity (equivalence of different theorems; two results reduced to one);

3. Global Synthetic Unity (both Szemerédi's theorem and Furstenberg Multiple Recurrence, irrespective of their equivalence, may be understood as belonging to the framework of Structure Theorems).

The question now remains: which of these or which combination results in an explanatory proof? I believe that, in our case, both Local Reductive Unity and Global Reductive Unity

[197] See Section 2.3.3.

serve as necessary, but not sufficient, conditions on an explanatory proof. It is the Local Reductive Unity that allows us to pass from what would be a horrifying proof by cases (e.g., consider each sufficiently dense A, look at its behavior, ascertain the presence of arithmetic progressions) to a genuinely comprehensible proof. But this alone does not guarantee that the proof is explanatory. I have been claiming that the pure proof of Szemerédi's theorem *fails* to be explanatory despite the presence of a combinatorial Structure Theorem. And this is because the delicate computations necessary for this proof obfuscate the role of the Structure Theorem as the putative "reason why." This is where the impurity becomes essential. The equivalence of Szemerédi's theorem and Furstenberg Multiple Recurrence facilitates the explicit intervention of ergodic and infinitary techniques to prove Szemerédi's theorem. However, as noted above, this equivalence on its own offers no real insight into why Szemerédi's theorem is true. Thus, the unification produced by impure techniques that generates an explanatory proof is Global Synthetic Unity. Once we understand that both Szemerédi's theorem and Furstenberg Multiple Recurrence depend on Structure Theorems, the ergodic context sheds light on Szemerédi's theorem.[198] And this is because the decomposition of measure-preserving systems is particularly explicit and clean (in virtue of the infinitary nature of the ergodic context), so the role of the Furstenberg Structure Theorem as the "reason why" is easily recognizable. Thus, the Global Synthetic Unity is what primarily drives the explanatory gains.

This analysis is particularly interesting because the interaction of different kinds of unity is a marriage of the unificatory techniques of Friedman and Kitcher (along with Morrison's precisification). As Morrison notes, Friedman emphasizes reductive unity: we simply need to minimize the numbers of facts we take as brute. Kitcher, on the other hand, focuses on explanation as a global phenomenon: we wish to minimize the number of *argument patterns* that generate the most conclusions. Indeed, he says, "Understanding the phenomena is not simply a matter of reducing the 'fundamental incomprehensibilities' but of seeing connections, common patterns, in what initially appeared to be different situations" ([Kitcher, 1989], 432). This recognition of common patterns has been absolutely crucial to my discussion, viz., the presence of the dichotomy between structure and randomness in both combinatorics and ergodic theory. However, unlike Kitcher, I have not tried to force this insight into the Procrustean bed of quantitative comparisons. Rather, it is more fruitful to consider a stratified picture:

1. Local Reductive Unity brings us from a mere conjunction of results (possibly transfinitely many) to a comprehensible classification;

2. This may not generate an explanatory proof if, for various reasons, the context is not suitable, e.g., combinatorial;

3. Global Reductive Unity serves as a bridge between the original context and a new, explanatorily appropriate context;

[198] Indeed, it seems possible to me that, even in the absence of an explicit equivalence, we could still say that the shared framework of Structure Theorems could generate explanatory dividends. This would, however, require that we move to a more global setting: the impurity in question could not be the presence of impure techniques in a particular proof, but rather something like a "hybrid" theory.

4. Global Synthetic Unity allows for the comparison of contexts that might shed further light on the original theorem, where the choice of appropriate contexts might be determined by the presence of structural content.

5. It is this Global Synthetic Unity that ultimately leads to an explanatory proof.

Indeed, Tao himself emphasizes the importance of this last kind of unity, saying, in a rather tantalizing fashion:

> These theorems [structure theorems] have remarkably varied contexts—measure theory, ergodic theory, graph theory, hypergraph theory, probability theory, information theory, and Fourier analysis—and can be either qualitative (infinitary) or quantitative (finitary) in nature. However, their proofs tend to share a number of common features, and thus serve as a kind of "Rosetta Stone" connecting these various fields ([Tao, 2006], 584).

Thus, it seems there is much more interesting work, both philosophical and mathematical, to be done.

To close, let me address an important disanalogy between my account and Morrison's. As we saw above, she claims that examples of true unification in the sciences involve a theoretical parameter representing the unification. Is there such a parameter in the case I am describing? Is such a parameter ever possible in mathematics? I am somewhat doubtful. For example, in Maxwell's theory of electromagnetic and optical processes, we are able to characterize such processes using eight field equations.[199] These field equations are then constitutive of the theory of classical electromagnetism (and so the classical limit of quantum electrodynamics). As Morrison makes clear, the theoretical parameter that serves as a necessary condition for these equations (and thus for unification) is that of the displacement current. Perhaps somewhat surprisingly, we rarely, if ever, have sets of equations that constitute what might be called a mathematical theory.[200] There are no such constitutive equations for combinatorics, nor for ergodic theory. Certainly there are central concepts, definitions, and results, but these play a more ambiguous role. Of course, we do have axioms of both a local and global sort, viz., those that characterize particular concepts and those that serve a foundational role. But in passing to formalized axioms we lose the intuitive content of the concepts we were originally interested in. And so it does not seem that we could ever come to identify a theoretical parameter of the sort Morrison means. Nonetheless, we can still identify the conceptual boundaries that engender unification in mathematics. In the above case, we saw that we can understand both T_C and T_E as members of $\mathscr{T}$ because all theorems in this class depend upon Structure theorems classifying the behavior of the mathematical entities in question as either structured or random. This dependence property might be considered a sort of "theoretical parameter," but we must take this in a looser sense than Morrison does. At the very least, what we do have is a classificatory property, i.e., the property that

[199] As Morrison writes them at least. See [Morrison, 2000], p. 87.

[200] For example, one might suggest the Cauchy-Riemann differential equations. These provide necessary and sufficient conditions for characterizing a complex-valued function as holomorphic (complex-differentiable). But this is merely the characterization of a concept central to a particular domain of mathematics (complex analysis), but a far cry from characterizing the domain itself.

allows us to understand that some variety of unification is occurring, which also serves as an explanatory property, i.e., the reason why the theorems we survey hold. And this seems to me a rather surprising result.

2.8 Concluding Remarks

Let me now bring together some of the conceptual strands. The general philosophical moral of this study is that impure methods often play a distinctively explanatory role in proving theorems, in particular, by facilitating greater simplicity and unification. Of course, each of these terms (impurity, simplicity, unification) must be suitably understood.

The impurity I have considered is *radical* insofar as it is both elemental and topical: ergodic techniques are both infinitary (and thereby more complex) and intuitively different in kind from the combinatorial techniques utilized in the pure proof of Szemerédi's theorem. Both the elemental and topical impurity play a role in generating explanatory power. However, the topical impurity must be sufficiently constrained. Indeed, though ergodic theory and combinatorics should be granted their intuitive differences, they may not be *so* different in light of their shared content. I have argued that Szemerédi's theorem and Furstenberg Multiple Recurrence should be understood as having shared *structural content* because the entities involved in both theorems instantiate the dichotomy between structure and randomness (registered by their dependence upon Structure Theorems). It is this shared content that undergirds the move from impurity to simplicity and from simplicity to explanatory power. The presence of this content allows for the highly infinitary Furstenberg Structure Theorem to abstract from the interlocking epsilon computations present in the Szemerédi Regularity Lemma, thus simplifying the proof of Szemerédi's theorem, *without losing the reason why the theorem holds*. To summarize the dialectic:

> Impurity [constrained topical + elemental] ⇒
> Simplicity [conceptual speed-up/overall proof structure] ⇒
> Local Explanatory Power [proof-based].

Similarly, we are licensed in moving from impurity to unification to explanatory power. In virtue of the fact that essential explanatory data is not lost (the shared structural content), the unificatory mechanisms at work are able to serve an explanatory function, i.e., both Szemerédi's theorem and Furstenberg Multiple Recurrence can be understood as part of a framework of Structure Theorems. Again, to summarize:

> Impurity [constrained topical + elemental] ⇒
> Unification [Global Synthetic] ⇒
> Local Explanatory Power [proof based].

Finally, I have noted that there may be other forms of unification that serve to make the Global Synthetic Unity possible. These should not, however, be thought of as playing a distinctively explanatory role.

I find it instructive to think of all this in terms of conceptual "distance." Obviously, the

conceptual distance in question will depend upon our analysis of the content of a theorem (thus motivating my decision to place the discussion of content first). If we cleave to an intuitive notion of content, then, in the case at hand, it seems we travel very far to prove the theorem: ergodic theory is indeed conceptually far from operations on natural numbers. However, perhaps this distance is only "surface level": once we have excavated, e.g., structural content, the theorems may not be so "far away" after all. Thus, the mathematical universe may be clustered in interesting ways, and ostensibly unrelated domains might have deep conceptual connections only discoverable by, to quote Bourbaki, "higher and frequently difficult stage[s] of abstraction." And thus, explanation, while not being local in the sense of requiring a purity constraint, as Aristotle and Bolzano would have it, cannot be radically global; there must be *some* shared content between domains in order for one to explain the other.[201]

Finally, I have deliberately avoided many—now canonical—questions about the status of explanation; in particular, whether it is an "ontic" or "epistemic" notion. I have treated it as lying on the side of knowledge: having an explanatory proof is a distinguished form of mathematical knowledge that generates understanding of why the proof holds. However, despite this avowedly epistemic language (e.g., understanding), this does not necessarily commit me to any sort of internalism or "irrealism" about explanation. Namely, I need not say that the relation that obtains between the statement of a theorem and its explanatory fact fails to be supported by some other objective, mind-independent relation between mathematical entities. Let me offer a few reflections on this.

I take it that much support for an ontic or realist construal of explanation proceeds from the following line of thought. Some proposition C serves as the *explanans* of *explanandum* proposition E in virtue of the fact that there is some determinate objective reality in which facts c and e stand in some appropriate relation r. The fact that c and e (represented by C and E) stand in relation r provides the content of explanatory realism. The question now, of course, is: what could the objective relation r be? The explanatory realist quite plausibly suggests causation. Thus, one might think that any flavor of explanatory realism requires a causal realism to support it[202] (lacking any other plausible relation r). Unfortunately, in the mathematical case, we cannot specify r to be causation.[203] Does this require that we be committed to a notion of explanation in mathematics that is wholly internal, wholly about the way in which concepts formulated by inquirers hook up to one another?

It is not at all apparent that the answer is "yes." One might attempt a heavy-duty metaphysical defense of grounding, thus providing another candidate relation. On the other hand, we might begin, as I have done, from a careful examination of particular mathematical phenomena and see what we can glean from this. In working through the details of Sze-

[201] [Lange, 2017] makes a point that might be construed along these lines: he argues that an explanatory proof must involve what is "salient" in a particular theorem. Unfortunately, he analyzes salience in a phenomenological fashion (what "jumps out" at us from a theorem). This risks trivializing explanation as something ephemeral and *entirely* subjective. Furthermore, as I have argued, the content of a theorem relevant for an explanatory proof may not be available up front, but rather requires a good deal of careful work to discover.

[202] This is one of the main points of [Kim, 1988].

[203] At least as we now understand it. Aristotle's notion of "formal cause" could serve us well here. Indeed, *Posterior Analytics*, from which I have drawn much inspiration, considers, it would seem, only formal causes.

merédi's theorem, I am struck by, for lack of a better phrase, the "givenness" of it all. We are presented with two *prima facie* completely different domains which, in the final analysis, share robust structural connections. In attempting an ergodic proof of Szemerédi's theorem, we find a "naturally occurring" result in ergodic theory that allows us to write a proof that generates a much clearer sense of why the theorem is true. We have not artificially constructed a different mathematical apparatus to solve this problem. We have not projected any sort of desire to find structure in randomness; this simply falls out of the analysis (recall Tao's remarks about the presence of Structure Theorems as a mysterious "Rosetta Stone"). Of course, one might claim that the structural confluences we do find were constructed long ago, and we are merely reconnecting conceptual relationships laid down by other inquirers. This *may* be the case; however, I find it to be entirely contrary to the phenomena and thus quite implausible. It may also be the case that platonism strikes us, along with Weyl and Feferman, as a "medieval metaphysics of mathematics" and elicits discomfort. But it may be correct, and if mathematics does not yield to more palatable philosophical reductions, so much the worse for these reductions.

3 Cosmic Topology, Conventionality, and the Constitutive *A Priori*

3.1 Introduction

A great deal of philosophical attention has been paid to the epistemology of physical geometry. This attention is certainly warranted as problems in physical geometry provide an especially illuminating case study of the relationship between evidence and theory, science and mathematics, and the philosophical presuppositions underlying all these.[1] Nonetheless, I believe that the mathematical and scientific developments of the past half century indicate that it is not the *geometrical* features of space and time that are most philosophically interesting but rather the *topological* features. Here I am using "topology" somewhat loosely as is commonly done in the cosmological literature.[2] That is, I am concerned with qualitative properties, e.g., shape, connectedness, causal structure, of an antecedently defined topological manifold.

Geometry is, ultimately, concerned with *local* and *quantitative* notions of distance or measurement. This suggests that an intelligible—if somewhat complex—story can be told as to how spatial geometrical facts relate to observers via the mediation of a properly constructed physical theory. On the other hand, topology is concerned with *global* and *qualitative* properties.[3] The role of topological properties in our physical theories and the relationship between these properties and scientific enquirers is, consequently, much more puzzling than in the case of geometrical properties. How is it possible for us to make claims about the structure of space *as a whole*, especially if claims about local properties of distance and measurement

[0]Content from Sections 1-8 of this chapter first appeared in the *European Journal for Philosophy of Science*. ©, the author, 2024. Please cite the published version: "Cosmic topology, underdetermination, and spatial infinity." *European Journal for Philosophy of Science*. 14:17 1-28 (2024). https://doi.org/10.1007/s13194-024-00576-7.

[1]For classic philosophical treatments of these topics see, e.g., [Sklar, 1974] and [Torretti, 1978]. See also [Gray and Ferreirós, 2021].

[2]See, e.g., [Geroch, 1967], [Ellis, 1971], [Geroch and Horowitz, 1979], [Luminet and Lachièze-Rey, 1995].

[3]The distinction between global/topological and local/geometrical features of space was, to my knowledge, first made in Riemann's, "Über die Hypothesen, welche der Geometrie zu Grunde liegen" (Concerning the Hypotheses on which Geometry is Based). See [Riemann, 2017] for the German text and [Riemann, 2004] for an English translation. Thus, this distinction is a relatively recent one that may not have been accepted by Riemann's predecessors. In this brilliant (though somewhat cryptic) address, Riemann attempts to show that we should consider Euclidean space to be a special case of the more general class of "n-fold extended quantities," i.e., the modern concept of a Riemannian manifold. Once we do this, we see that there is an explosion of mathematical options that may be used to characterize space. In particular, there are unbounded, finite spaces, e.g., the Riemannian hypersphere. This was Einstein's preferred model for the spacetime manifold when he first developed the foundations of general relativity.

are already fraught with difficulties and confusions (as witnessed by the long dialectic concerning these between mathematicians, physicists, and philosophers)? What, then, is the epistemic status of global topological properties?[4]

Let us briefly consider a few examples of topological properties. First, we can classify a space[5] as having a boundary or not, i.e., having an "edge" or not. For instance, neither infinite Euclidean space nor the surface of a finite sphere has a boundary. Second, a space may be *simply* or *multiply connected*: informally, in a simply connected space,[6] any closed loop can be contracted to a point. On the other hand, in a multiply connected space, the presence of a "hole" (or many "holes") causes such a contraction to fail. For instance, Euclidean space is simply connected while a torus is multiply connected.

Topological properties, unlike geometrical ones, do not suggest any obvious connection to sensory experience or to our cognitive faculties; in particular, it would seem that they cannot be measured or detected in any direct way. Indeed, think for a moment about what it would be like to live in a space with a "hole." You cannot directly perceive the "hole" because this would require living in a higher dimension and viewing the space "from the outside." But certainly your experience on this space will be different from, say, that of someone living in infinite Euclidean space? Perhaps, but determining *how* your experience might differ is no easy matter.

Furthermore, topological properties are not mere mathematical abstractions to be skimmed from the surface of respectable empirical theories; our best theory of gravity, viz., general relativity, requires, for its cogency, that space and spacetime have determinate topological properties. If one accepts general relativity, then one must accept that spacetime is (at the very least) well-modeled by the mathematical structure of a topological manifold.[7] One then asks: does this manifold have a boundary or not? Is it simply connected or multiply connected? Thus, given that these topological features are fundamental components of our best scientific theory of spacetime, the question of their epistemic status, especially in light of their "global" nature, is philosophically pressing.

This chapter addresses two questions involving the topology of space. The first concerns underdetermination and the role of epistemic virtues in relativistic cosmology. It is well-known that cosmology faces an underdetermination problem: there are many cosmological models compatible with our best observational data.[8] At first blush, this may be quite unsurprising given that cosmology deals with physics at extremely large scales. A rather more surprising fact is that, even under strong hypotheses about the global structure of space (the Cosmological Principle), this underdetermination persists. In particular, we are still unable to ascertain the global topology of space.

Is there any way to break this topological underdetermination? After discussing the

[4]See [Manchak, 2009], [Manchak, 2011], [Manchak, 2013] for related discussions. See also the masterful survey [Ellis, 2007].

[5]When I say "space" here, I mean an abstract topological space. We are not yet in the realm of physics.

[6]More formally: a topological space X is said to be simply connected if it is path connected and the fundamental group $\pi_1(X, x_0)$ (relative to base point x_0) is trivial, i.e., $\pi_1(X, x_0)$ reduces to the identity element. Note that the fundamental group does not depend on the choice of base point. My informal gloss should suffice for most of the discussion.

[7]It will be made clear below why I am considering the topology of space and not spacetime.

[8]For discussion of underdetermination in cosmology and related issues see [Beisbart, 2009], [Manchak, 2009], [Smeenk, 2013], [Butterfield, 2014].

mathematical background for relativistic cosmology and the general problem of underdetermination (Section(s) 3.2, 3.3, 3.4), I survey recent work in observational cosmology that has aimed to provide definitive answers on this front and conclude that the prospects for empirically determining spatial topology are not promising (Section 3.5). However, a familiar point in the philosophical literature is that underdetermination by data may not be so worrisome.[9] This is because one may be able to find significant epistemic reasons for preferring one theory (or model) over another.

As such, I argue that we can muster epistemic reasons to prefer various topologies over others. In particular, I argue that we should prefer cosmological models with multiply connected topologies on grounds of simplicity, Machian considerations, and explanatory power (Sections 3.6, 3.7, 3.8). We are able to ascribe such features to multiply connected models because they generate spatially finite universe models, which in turn avoid extremely thorny issues concerning the postulation of an actually infinite universe. Thus, though a purely observational underdetermination remains, we can avoid a more robust underdetermination, viz., one in which all epistemic reasons underdetermine the choice of topology.

The second question, already mentioned above, concerns the epistemic status of spatial topology. What are we to say about the topology of space in light of the topological underdetermination? A natural suggestion, especially given the long and lively dialectic concerning the conventionality of spatial *geometry*, is that the topology is "conventional." Using the mathematically simpler and better known geometrical conventionalism of Poincaré as a springboard, I provide a few arguments for the conventionality of spatial topology and assess how this conventionality is to be understood. However, conventionality of any stripe does not *fully* capture the epistemic status of the topology of space. This is because of its foundational role in our cosmological theorizing. It would appear that there is a sense in which spatial topology makes possible the application of fundamental physical concepts and subsidiary physical laws. Thus, I turn to Michael Friedman's work on the "relativized" or "constitutive" *a priori* to make sense of this predicament. Nonetheless, Friedman's account cannot do this in a straightforward fashion; rather, we require a new epistemic category, involving a fusion of conventionalism and the relativized *a priori*, to capture the epistemic status of spatial topology.

3.2 Fundamentals of Spacetime Structure

I begin by rehearsing some of the basic details of relativistic cosmology. Speaking circumspectly, we can say that cosmology is the study of the large scale structure of the universe. By "universe," we might mean either everything that exists in the physical sense or that which comprises everything that exists physically. Both of these notions are useful and can be understood rigorously. Namely, we can think of the universe as the *spacetime* in which everything is contained together with the *distribution* of matter and energy in this spacetime. I am primarily interested in the former, though details about the latter will become relevant later.

Given that gravitation is the dominant force at large scales, we must consider our best extant theory of gravity: Einstein's theory of general relativity. Thus, ultimately, cosmol-

[9] For nice discussions, see [Laudan, 1990],[Earman, 1993]

ogy is concerned with finding models of general relativity that are consistent with our data concerning the spacetime structure and energy distribution of the universe at large scales.[10] In terms of the standard formalism, we say that a *model of general relativity*[11] is a triple $(\mathcal{M}, g_{ab}, T_{ab})$, where $\mathcal{M}$ is a connected four-dimensional real smooth manifold without boundary of variable curvature,[12] g_{ab} is a metric tensor (field) of type (0,2),[13] and T_{ab} is the energy-momentum tensor (field). The metric g_{ab} characterizes the geometric properties, e.g., curvature and geodesics, of $\mathcal{M}$ at a given *point*, $p \in \mathcal{M}$. Finally, T_{ab} characterizes the energy distribution of $\mathcal{M}$ and is described by suitable equations of state relating its components, again at a particular point, $p \in \mathcal{M}$.

We must now understand how these elements of models of general relativity interact. In particular, we seek a field equation relating the metric g_{ab}, characterizing geometry, and the energy-momentum tensor T_{ab}, characterizing energy distribution. This relationship is expressed as:

$$R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = \kappa_0 T_{ab}, \tag{3.2.1}$$

which is now known as *Einstein's Equation*.[14] More precisely, the left-hand side characterizes the curvature of $\mathcal{M}$ at a point p given the specification of g_{ab}.

Technically, Equation 3.2.1 expresses ten non-linear partial differential equations of immense mathematical complexity. In order to obtain "exact solutions" that can be studied both mathematically and physically, one must lay down plausible simplifying assumptions that accord with observational data. I turn to these assumptions in a moment; however, before complicating matters, we can already express a general sort of cosmological underdetermination.

3.3 Underdetermination and the Cosmological Principle

In providing a model of general relativity, we provide a particular kind of ambient manifold structure and a metric and energy distribution solving Einstein's Equation. How do we go about doing so? Certainly, we wish such a model to match our observations at a given point in spacetime. The hope is that our observational data can narrow down a unique model (or unique class of models).

It should be noted that by "unique" we really mean "unique up to isometry." That is, we

[10] This distinction between model and theory is slippery and usage varies, but my meaning should be reasonably clear in what follows. See [Butterfield, 2014], 58-9.

[11] I will drop T_{ab} later, but it is helpful here in describing the Einstein Equation.

[12] See [Ellis and Hawking, 1973] and [Wald, 1984] for details.

[13] More precisely, g_{ab} is a smooth, non-degenerate, pseudo-Riemannian metric of spacetime/Lorentz signature (–, +,+,+).

[14] Here R_{ab} is the Ricci tensor, g_{ab} is the metric tensor, R is the Ricci curvature scalar, Λ is the Cosmological Constant, $\kappa_0 = 8\pi G/c^4$ is the Einstein gravitational constant, and T_{ab} is the energy-momentum tensor. Λ was originally included in the field equations by Einstein to achieve a static cosmological model (among other things). Today it is invoked as a dark energy candidate to explain the observed acceleration of the expansion of the universe. See [Earman, 2001] for a nice discussion of Λ.

say that two models $(\mathcal{M}, g_{ab}, T_{ab})$, $(\mathcal{M}', g'_{ab}, T'_{ab})$ are isometric if there is a diffeomorphism[15] $\varphi : \mathcal{M} \to \mathcal{M}'$ such that $\varphi_*(g_{ab}) = g'_{ab}$.[16] This easily descends to the local case of open sets on the manifolds. The crucial point is that two isometric manifolds (resp. open sets of manifolds) do not constitute distinct physical possibilities because they cannot be distinguished using observations. This is so because the isometry preserves the metric structure across manifolds and thus preserves solutions to Einstein's Equation.

We can now state the conditions required for isolating a unique model/class of models for general relativity. According to general relativity, anything we observe at a spacetime point p (for some $p \in \mathcal{M}$ in some model) must be causally related to p; however, signals cannot propagate faster than the speed of light. Thus, the events with which we can have causal contact sit either on or within a particular region of spacetime bounded by the paths of light that arrive at p. We denote this region by $J^-(p)$ and call it the *past lightcone at* p or simply the *observable universe at* p.[17] For reasons of mathematical convenience, we follow [Manchak, 2009] and work primarily with the interior of $J^-(p)$, denoted by $I^-(p)$.[18] Let us write $I^-(p_0)$ for our observable universe.

Thus, if we are to pick out a unique class of models compatible with our observations at p_0,[19] we require:

Condition 3.3.1. Up to isometries, there is a unique model $(\mathcal{M}, g_{ab}, T_{ab})$ that has a point $q \in \mathcal{M}$ such that $I^-(p_0)$ and $I^-(q)$ are isometric.[20]

Unfortunately, it is well known that the uniqueness condition cannot be satisfied. There are various ways to see this, but perhaps the slickest is by appeal to recent results by Manchak.[21] In particular, Manchak shows that virtually any model $(\mathcal{M}, g_{ab})$ (subject to a few reasonable constraints[22]) will be observationally indistinguishable[23] from another model $(\mathcal{M}', g'_{ab})$ that is *not isometric* to $(\mathcal{M}, g_{ab})$. Consequently, an ideal observer at $p \in \mathcal{M}$ who knows all metrical data about $I^-(p)$ can know very little about the global structure of their spacetime, since there will be many spacetimes possessing markedly different global properties that contain regions isometric to $I^-(p)$.

Thus, it would appear we are in very bad shape when we try to provide a unique model of general relativity that matches our observational data. And thus we are confronted with a severe underdetermination of models by data. What's worse, by the above results of Manchak,

[15] A smooth, bijective map with smooth inverse.

[16] See [Wald, 1984], [Manchak, 2009].

[17] Also, technically, $J^-(p)$ must sit to the future of the time of decoupling.

[18] The $I^-(p)$s are mathematically simpler because they are open sets, as opposed to the $J^-(p)$s which are closed. See [Cinti and Fano, 2021] for a brief discussion of the physical significance of this restriction.

[19] When I write p_0 and $I^-(p_0)$ in Condition 3.3.1, I am not quantifying over points in different models. These notions simply serve as shorthand for *our* observable universe from an arbitrarily selected spacetime point p_0.

[20] Here I simply follow the requirement given in [Beisbart, 2009]. It is a natural and widely acknowledged one. See also [Butterfield, 2014].

[21] See [Manchak, 2009]. His results make rigorous ideas found in [Malament, 1977].

[22] In particular, well-behaved causal structure.

[23] Manchak defines two models of general relativity $(\mathcal{M}, g_{ab})$, $(\mathcal{M}', g'_{ab})$ to be *observationally indistinguishable* iff for all $p \in \mathcal{M}$, there is some $p' \in \mathcal{M}'$ such that $I^-(p)$ and $I^-(p')$ are isometric. See [Cinti and Fano, 2021] for alternative notions of observational indistinguishability.

this almost always appears to be the case. However, the severity of this underdetermination can be greatly reduced by restricting the models of general relativity considered. Appeal is usually made to the following:

Principle 3.3.2. (Cosmological Principle) On average, at large scales, the universe is spatially homogeneous and isotropic around every point.[24]

Thus expressed, the Cosmological Principle is essentially an *a priori* prescription imposed on all possible models of general relativity. Once imposed, it has the effect of restricting our attention to a particularly well-behaved class of models, the *Friedmann-Lemaître-Robertson-Walker* (FLRW) models. There are many intricate arguments, drawing on a wide variety of considerations (some empirical), for the Cosmological Principle.[25] These arguments are of great philosophical interest, for the Cosmological Principle, if acceptable, would significantly reduce the underdetermination. I will, however, simply assume the Cosmological Principle here. The reason for this is that, even under the strong hypothesis of the Cosmological Principle, the model underdetermination persists when we consider topological properties of our spacetime manifold $\mathcal{M}$. Indeed, somewhat astonishingly, for each metric solution of Einstein's Equation internal to the class of FLRW models, there may be *infinitely many* compatible topologies. I will now discuss these models and the relevant topological properties in greater detail.

3.4 FLRW Models and Topology

The Cosmological Principle amounts to the imposition of spatial[26] symmetry constraints. In particular, spatial homogeneity means, roughly, that every point in space at a given time "looks the same," and spatial isotropy means that there are no preferred spatial directions. We represent the spacetime manifold, $\mathcal{M}$, as the product of a three dimensional spatial manifold and a temporal continuum, i.e., $\mathcal{M} := \mathcal{M}_3 \times \mathbb{R}$. The spatial manifold can then be thought of as a "stack" of surfaces, each indexed by a particular cosmic time. The metrical structure of these FLRW models is particularly tractable, and, crucially for our discussion, the spatial sections have *constant curvature* with values $k = \pm 1, 0$, respectively.

Once more, the essential point is that, even with all these simplifications, we have said nothing about the topology of $\mathcal{M}_3$. Until quite recently, it has been assumed in the cosmo-

[24] As expressed in [Wald, 1984], 92-3. Also, before imposing the Cosmological Principle, one must assume that space and time can be "split," i.e., the entire spacetime manifold, $\mathcal{M}$, can be written as $\mathcal{M}_3 \times \mathbb{R}$, otherwise we could not make sense of imposing only spatial symmetry constraints. A strong—but common—assumption that guarantees this is called *global hyperbolicity*. This condition amounts to claiming we can determine the evolution of spacetime from our information about a spatial hypersurface, Σ, at a given time. More precisely, Geroch showed that a spacetime is globally hyperbolic iff it admits a Cauchy surface (see [Geroch, 1970]). This means that a globally hyperbolic spacetime admits a hypersurface, Σ, whose domain of dependence, $\mathcal{D}(\Sigma)$, is the entire spacetime manifold. That is, from the physical "information" given by Σ, one can deduce the evolution of the entire spacetime manifold.

[25] See [Ellis, 2007], Section 4, [Beisbart, 2009], [Smeenk, 2013], [Butterfield, 2014].

[26] This is crucial. We do not have *spatiotemporal* symmetries. Metrical structure is only preserved on spatial hypersurfaces of $\mathcal{M}$ but not throughout $\mathcal{M}$. The only exception among FLRW models is the de Sitter universe, which neglects ordinary matter. The de Sitter universe satisfies the "perfect" Cosmological Principle that imposes homogeneity and isotropy throughout space and time.

logical literature that the topology of $\mathcal{M}_3$ is *simply connected*. Once more, informally, in a simply connected space, any loop through a point x_0 can be continuously deformed into any other closed loop through x_0.[27] However, neither observational data nor the FLRW models themselves dictate such a choice. It is entirely possible that the spatial sections are *multiply connected*, i.e., there is a "hole" (or many "holes") that renders such a continuous deformation impossible.[28] For instance, a hypertorus[29] is multiply connected, while Euclidean space is simply connected.[30]

Let us consider the possible simply connected models (SCMs). There will then be three candidates for the spatial section $\mathcal{M}_3$: the 3-sphere ($\mathbb{S}^3$), Euclidean 3-space ($\mathbb{R}^3$), and the 3-hyperboloid ($\mathbb{H}^3$). These correspond, respectively, to the three possibilities for constant curvature and will be equipped with their respective classical geometries, viz., spherical, Euclidean, and hyperbolic. The possible SCMs along with their central mathematical properties are summarized in Table 1 below.

Table 1: Mathematical Properties of SCMs

Spatial Section	Geometry	Curvature	Topology	Extent of Universe
$\mathbb{S}^3$	Spherical	$k > 0$	SC	Finite
$\mathbb{R}^3$	Euclidean	$k = 0$	SC	Infinite
$\mathbb{H}^3$	Hyperbolic	$k < 0$	SC	Infinite

When we deal with SCMs, note that the determinant of the spatial extent of the universe is the *curvature* of $\mathcal{M}_3$ alone.

Let us now turn to multiply connected models (MCMs). The effect of a multiply connected topology for $\mathcal{M}_3$ is equivalent to considering a particular simply connected space (the universal covering space, denoted by $\widetilde{\mathcal{M}}_3$) tiled with particular polyhedra (fundamental polyhedra).[31] This tiling of the covering space is achieved by the action of a group Γ on the covering space. Since we only deal with constant curvature models, we need only consider three universal covering spaces $\mathbb{S}^3, \mathbb{R}^3, \mathbb{H}^3$ under the action of such a Γ. In order to get a multiply connected topology, we form a quotient manifold $\widetilde{\mathcal{M}}_3/\Gamma$, which is gotten by identifying points equivalent under the action of Γ on the covering space $\widetilde{\mathcal{M}}_3$, where $\widetilde{\mathcal{M}}_3$ which is one of the three constant curvature SCMs. For example, $\mathbb{R}^3/\Gamma \cong T^3$, the hypertorus, where Γ consists of discrete translations identifying faces of the fundamental polyhedra (parallelepipeds) tiling $\mathbb{R}^3$.

To summarize, we can re-write Table 1 above with the choice of multiply connected topology. See Table 2 below.

[27] More formally: a topological space X is said to be simply connected if it is path connected and the fundamental group $\pi_1(X, x_0)$ reduces to the identity element.

[28] More formally: X is multiply connected if it has a non-trivial fundamental group.

[29] $T^3 = S^1 \times S^1 \times S^1$. See below.

[30] For foundational texts on alternative topologies for space see [Ellis, 1971], [Luminet and Lachièze-Rey, 1995], [Luminet, 2015].

[31] See [Wolf, 1967], [Ellis, 1971], [Luminet and Lachièze-Rey, 1995], [McCabe, 2004] for further mathematical details.

Table 2: Mathematical Properties of MCMs				
Spatial Section	Geometry	Curvature	Topology	Extent of Universe
$\mathcal{M}_3 = \mathbb{S}^3/\Gamma$	Spherical	$k > 0$	MC	Finite
$\mathcal{M}_3 = \mathbb{R}^3/\Gamma$	Euclidean	$k = 0$	MC	Infinite or Finite
$\mathcal{M}_3 = \mathbb{H}^3/\Gamma$	Hyperbolic	$k < 0$	MC	Infinite or Finite

This will not affect the geometry of each case, so, e.g., geometrically $\mathbb{R}^3$ and the hypertorus T^3 are the same, and so will be observationally indistinguishable, provided the topology cannot be empirically determined. However, the topology change will affect the possible size of the universe, e.g., T^3 is finite, while $\mathbb{R}^3$ is infinite.

Thus, we see that the effect of considering MCMs (in addition to SCMs) produces an explosion of new FLRW models of general relativity consistent with our best data.[32] Once more, this is the case even under the very strong assumption of the Cosmological Principle. Is there, then, any means of breaking the underdetermination of models? In recent years, there has been active research in the field of *cosmic topology* whereby cosmologists have attempted to empirically ascertain the global topology of space. I will now briefly review the most promising aspects of this research.

3.5 Recent Cosmological Research on Spatial Topology

Now that we have motivated the study of different spatial topologies and indicated the various options available, let us examine some recent cosmological research concerning the possibility of "empirically" verifying MCMs. I merely provide a summary in this section and invite my reader to consult the references provided.

First, we need to get clearer on what is meant by "empirical verification" in the context of large-scale cosmological research. It seems to me that when most cosmologists talk of "empirical verification" they are using the phrase in a rather broad sense that includes a pretty unrestricted appeal to inference to the best explanation. The particular point at issue is this: certainly, we cannot *observe* the topology of the universe in any direct way; however, we can help ourselves to downstream observational consequences of particular spatial topologies. For instance, one of the preferred methods for detecting spatial topology relies on examining particular configurations in Cosmic Microwave Background (CMB). Particular configurations are consistent with particular topologies; thus, the observation of CMB is taken as justification for the claim that the universe has a particular spatial topology. My methodology in this paper is to take as many hypotheses commonly employed by working cosmologists on board, e.g., the Cosmological Principle; thus, I do not object to this usage of "empirical verification" nor to the use of inference to the best explanation, though, of course, there are many philosophical questions about this sort of reasoning.[33] I simply wished to say that it is a rather generous usage, and we should should be aware of it. Indeed, there are even further issues which arise given the reliance of such observations on massive (but still too small) data sets and complex methods of data collection. I will flag some of these issues below.

[32] In particular, we see that spatial extent is no longer exclusively determined by the curvature of space as in SCMs.

[33] See [Lipton, 2004] for an authoritative discussion.

3.5.1 The Observable Universe

Before engaging in any empirical investigation, we must ask: what is the putative domain from which our data might be drawn? How is this domain delineated? It turns out that in asking these questions yet another important distinction arises, namely, that between the the observable and non-observable universe. And, what's worse, this distinction does not map perfectly onto the finite versus infinite universe models.

There are two observational limits of particular import. First, the observational data of cosmology is electromagnetic radiation from various times and places in the evolution of the universe. However, this data is only available from the *time of decoupling*, i.e., when the universe became transparent to radiation (before decoupling, matter and radiation could not be distinguished). Any information about the universe present before the time of decoupling is thereby inaccessible to us.[34] Second, a finite time has elapsed between the present and the time of decoupling, and light has traveled a finite distance in that time. Since no signal can travel to us faster than the speed of light, it is the case that data about entities beyond this finite distance is inaccessible. A natural consequence of these observational barriers is that knowledge of the universe as a whole is beyond our grasp. Certainly, we can (and do) theorize about it, but we must be careful to acknowledge that we do not have observational or causal access to it.

There are then three possible combinations of infinite versus finite universe size with observable versus non-observable data (since "infinite + observable" makes little sense):

1. The whole universe is infinite as in, e.g., the SCMs with $k = 0$ or $k = -1$ (see Table 1 above). In this case, the observable universe is an infinitesimal patch of the whole universe. Here *neither* topological features of space nor the size of the universe will be empirically testable.

2. The whole universe is finite but exceeds the observable universe. If the difference between the whole and observable universe is "too large," then, as in (1), both topology and extent are not empirically determinable properties. However, there might be cases where the actual universe is not too large, and, if very favorable circumstances obtain, this could be empirically determinable.[35]

3. The whole universe is a small universe, i.e., the observable exhausts the actual. Both topology and size would be, in principle, empirically determinable.

In this section, we are concerned with (3), i.e., particular universe models in which it makes sense to talk about observationally determining both topology and size of the universe.

3.5.2 Three Detection Techniques

The basic idea underlying all recent attempts at determining the spatial topology of the universe is the following. If we live in a universe that enjoys a multiply connected topology,

[34] And this data would be marvelous to have. For instance, many physicists postulate a time of quantum inflation (before decoupling) in order to account for the large-scale isotropy of the universe. This is, however, merely a postulate and cannot be verified by canons of scientific experiment.

[35] See [Fabre et al., 2013].

then space can be represented via a universal covering space tiled by a fundamental domain. That is, an MCM is topologically equivalent to an SCM subject to particular periodic boundary conditions. The immediate physical effect of this periodicity is that sources of radiation will produce multiple images (because there will be multiple shortest paths along which light travels) occurring at particular points in a lattice, which is in turn consistent with a particular multiply-connected topology. All recent work has attempted to exploit this fact in some way. For instance, Figure 3.1 below represents the universal covering space of the two-torus, T^2, i.e., a two-dimensional MCM:[36]

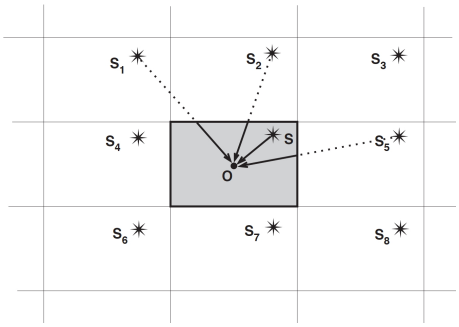


Figure 3.1: Universal Covering Space of T^2

The shaded region is the fundamental polygon, subject to periodic boundary conditions, S is a source of radiation, O is an observer. S propagates light to O along the "intuitive" geodesic (ray SO). However, this light would also scatter in infinitely many directions, "wrap around" the universe, and generate the appearance of infinitely many "apparent" or "ghost" images $\{S_1, \ldots, S_8, \ldots\}$. All recent work has attempted to exploit a higher-dimensional version of this representation in some way. Figure 3.2 below represents the case of T^3 viewed from a "corner" of real space with Earth closest to us:[37]

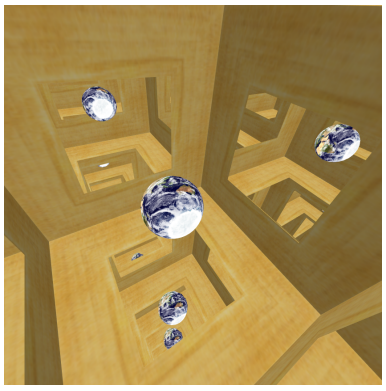


Figure 3.2: Universal Covering Space of T^3

[36]Image from [Luminet, 2015].
[37]Image generated using the Curved Spaces package by Jeff Weeks.

Let us now turn to the possible methods for empirically determining spatial topology.

Straightforward Observation. An initial attempt to ascertain whether the topology of the universe is trivial or not consists in identifying multiple images of particular celestial objects, e.g., galaxies, at regular intervals. The basic idea is to see if such objects lie in a regular lattice consistent with a particular topology. There are, however, immediate difficulties with this method: (i) images are viewed from different directions, at different angles; (ii) quality of images may obscure similarities and differences; (iii) even if one can identify two images as those of the same object, these images may correspond to different periods in the object's evolution, and so the images may not be "genuine" multiple images of the same (i.e., spatially and temporally) object. The upshot of all these problems is that direct observational data of distant celestial objects is not sufficiently reliable to determine the topology. (Also, note that I say this approach is "direct", but, of course, the observations are mediated by very complicated technology, viz., extremely powerful telescopes, which introduces yet another level of complexity.)

Cosmic Crystallography. This technique utilizes statistical methods that avoid, at least in part, the difficulties of the straightforward observational approach.[38] In particular, there is no direct dependence on visually recognizing the morphology of sources of radiation. The basic idea behind the technique of cosmic crystallography is to collect as many galaxy images as possible and attempt to discern particular statistical properties over the distribution of these images. The main obstacle here is that the available data sets are far too meager to provide any convincing evidence of multiply-connected topologies. There are also some rather complicated dependencies that indicate this method may only be applicable to manifolds with a particular geometry. In short, as of now, there are serious uncertainties regarding this technique. Despite these difficulties, statistical analyses of cosmic images (cosmic crystallography) is one of the better techniques available for determining cosmic topology.

Circles in the Sky. Let us conclude with the most promising technique, the so-called "circles in the sky" method. It is also the most complicated method on offer.[39] Here is an extremely coarse sketch of what is involved.

According to the standard Big Bang theory, the universe is generated from an extremely hot, dense energetic plasma. This plasma is entirely opaque to light because photons will scatter off of hot charged particles. As the universe expands, the plasma cools sufficiently to permit radiation to pass through it (experimental evidence postulates that this point of cooling or "decoupling" occurs about 300,000 years into the universe's life). Note that this cooling happens very quickly, but it is not instantaneous and introduces some difficulties into the following method.

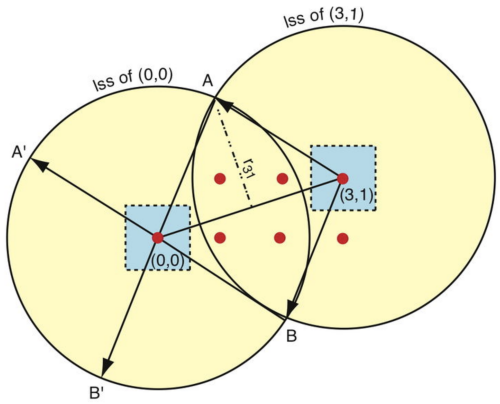
In any case, the ancient scattered radiation has by this time cooled to be observable in the microwave spectrum and is called Cosmic Microwave Background (CMB). CMB carries

[38] See [Lehoucq et al., 1996] for an early survey of these techniques and [Rebouças and Gomero, 2004] for a more recent assessment.

[39] See [Cornish et al., 1998] for the original paper on this technique. [Levin, 2002], [Rebouças and Gomero, 2004], [Cornish et al., 2004], [Luminet, 2015] also provide further details and discussion.

coveted data about the very beginning of the universe and, as it turns out, might be useful in detecting non-trivial cosmic topologies. If we imagine the initial state of the universe as a "point," then CMB would have scattered in every direction from this point to reach us now, forming a "sphere" of radiation processing from the center of the sphere. This sphere of radiation is called the "surface of last scattering" (SLS).

The question that concerns us here is: how can the SLS be used to detect cosmic topology? If the universe has a non-trivial topology, then, as we have seen, it can be represented as its covering space tiled by fundamental polyhedra. Each "copy" of the observer (that is, each analogous point in each fundamental polyhedron) in each polyhedron will come associated with its own SLS; provided that the diameter of the fundamental polyhedron does not exceed that of the diameter of the SLS, then the SLS spheres will intersect, generating a "circle" of CMB radiation. Since there will be an observer and their "copy," these intersections will come in pairs viewed in different directions. The presence of pairs of circles in CMB radiation will then be a sign of non-trivial topology. See the image below for the circles method applied to the 2-torus with fundamental polygon a square.[40] Note that this method will help to deal with issues of "genuine" multiple image identification, because a genuine multiple image in this case will have an SLS of identical temperature fluctuations and the same radius. Thus, this appears to be a very powerful method for detecting cosmic topology.



One very important advantage of "circles in the sky" over cosmic crystallography is that it will apply to all non-trivial topology models, and the model can be directly reconstructed from the radius, number, and distribution of the circles. This circumvents the issues in cosmic crystallography where various models could not be disambiguated. However, there are still observational issues involved in the circles method. In particular, the velocity and density of the SLS can become obscured, thus affecting the accurate detection of circles (e.g., the gravitational pull of coalescing galaxies could be a potential obstruction).

3.5.3 Evaluation of Empirical Techniques

Have traces of "small" MCMs been found by the above method? Unfortunately, the results thus far are not promising, though they have been fiercely debated. It seems that many

[40] Image from [Luminet, 2015].

of the favored "small" MCMs have been ruled out: no matching circles have been found for hypertori nor for other important models. However, the results employed to rule out these models do not apply to all MCMs; in particular [Aurich et al., 2004] has claimed to have found some evidence for a multiply connected hyperbolic model called the *Picard Horn*. The Picard Horn is particularly interesting because, though one direction of its fundamental polyhedron is infinite, the space as a whole has finite volume.

Thus far, we have considered the "best case scenario" for empirically determining cosmic topology, i.e., Small Universes. It is worth asking whether we can lift this assumption and consider cases in which the universe is finite and exceeds the observational horizon, but only by a "negligible" amount. Astonishingly, there has been some recent work that has shown it would be possible to distinguish an infinite universe from a finite, though technically non-observable, universe for particular multiply connected topologies. This means that, even if we did not live in a small universe, but rather a "relatively" small universe, both the topology and size of space could be empirically detectable ([Fabre et al., 2013]).

Despite these developments, there has been no especially compelling evidence for a multiply connected spatial topology. As I have tried to indicate in this section, this does not, of course, rule out the possibility. It does seem, though, that the set of models both (i) consistent with our best evidence and (ii) either small or "relatively" small is shrinking.

3.5.4 The Topological Underdetermination Thesis

Before turning to further complications, let us summarize our findings and make explicit our underdetermination thesis. We have taken the Cosmological Principle on board as an assumption about the global structure of space. An immediate consequence of this assumption is that we must restrict our attention to the FLRW models of general relativity. We then saw that, internal to this highly symmetric class of models, we might distinguish between simply connected models (SCMs) and multiply connected models (MCMs). The existence of a tractable (because spatially finite and particularly small) subset of MCMs, the so-called Small Universes, suggested that we might empirically investigate spatial topology. However, given the lack of empirical evidence that we live in a Small Universe, combined with the many sensitivities and difficulties of the empirical techniques used, we postulate:[41]

Thesis 3.5.1 (Topological Underdetermination Thesis)**.** For any simply connected FLRW model $(\mathcal{M}, g_{ab})$, there exists a multiply connected FLRW model $(\mathcal{M}', g'_{ab})$ that is not isometric to $(\mathcal{M}, g_{ab})$ such that $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ are observationally indistinguishable.[42]

For example, let the spatial section $\mathcal{M}_3$ of $\mathcal{M}$ be $\mathbb{R}^3$. This is simply connected and infinite. Let the spatial section $\mathcal{M}'_3$ of $\mathcal{M}'$ be T^3, the hypertorus. This is multiply connected and finite. These two models share the exact same kinematics and dynamics and so, given

[41] Note that, in general, my underdetermination thesis does not follow from Manchak's result in [Manchak, 2009]. This is because I impose further conditions on $(\mathcal{M}', g'_{ab})$; in particular, that it be an FLRW model.

[42] Again, following [Malament, 1977], we say that two models $(\mathcal{M}, g_{ab})$, $(\mathcal{M}', g'_{ab})$ are *observationally indistinguishable* if for all $p \in \mathcal{M}$, there is some $p' \in \mathcal{M}'$ such that $I^-(p)$ and $I^-(p')$ are isometric.

the observational difficulties noted above, are observationally indistinguishable.[43]

In sum, despite some impressive advances in the study of cosmic topology, I believe that we are still faced with a severe underdetermination of spatial topology by observational data. And this is so even under the assumption of the Cosmological Principle.

3.5.5 MCMs, Manchak's Theorem, and Isotropy

Finally, let us consider whether our topological underdetermination would vanish if strong evidence for a particular MCM emerged, e.g., matching circles of CMB indicating a particular multiply connected topology. The first issue to examine is the nature of the "empirical evidence" that could be adduced for MCMs. Obviously, the meaning of "empirical evidence" here must be quite generous in that sense that we infer a particular topology from its "signature" in the CMB. Second, there are a number of worrisome restrictions on popular techniques for determining topology. For example, it has emerged that some techniques have complicated dependencies upon the geometry (viz., group of isometries) of a given manifold; thus, a "negative" result may not in fact be such given that some possible models cannot be detected by the technique. Finally, even the preferred circles method is susceptible to many observational difficulties.

But let's proceed in the spirit of optimism. If it becomes clear that a particular MCM is the best explanation for a given set of observational data, have we dispensed with the topological underdetermination? Initially, it would appear that the answer is no. This is because the models we are considering are causally well-behaved, and so Manchak's theorem on observationally indistinguishable spacetimes still applies. It is not worth rehearsing the exact details of his proof here, but the basic point is that we can form spacetimes by cutting and pasting together subsets of other spacetimes ([Manchak, 2009], 55). Assume that we have been convinced that we live in a Small Universe, i.e., such that $I^-(p_0)$, our observable past from point p_0, contains all of space. It can then be shown that $I^-(p_0)$ can be attached to various other regions in the *future* of p_0, thereby generating a spacetime that is globally different from the assumed Small Universe model.

The significance of this result for Small Universes is somewhat unclear and ultimately depends upon a complicated set of considerations. For instance, one might already be concerned about the body of assumptions laid down internal to which the above observational data would be considered evidence. The fact that, even after all these assumptions have been stated, there is *still* the possibility that strong empirical evidence for an MCM would not break our underdetermination could incline one to a skepticism about knowledge of spatial topology. On the other hand, one might be relatively unconcerned about the assumptions in question and further contend that Manchak's result, in the face of evidence for an MCM, lacks bite. Namely, the cutting and pasting technique generates a spacetime that has a decidedly "cooked-up" look, and this is a much less reasonable explanation for the evidence in question.

Following the discussion in [Cinti and Fano, 2021], we can make these concerns about Manchak's result more precise. Their primary contention is that Manchak's construction

[43] More precisely, these models are governed by the exact same FLRW solutions to Einstein's Equation. See [Luminet and Lachièze-Rey, 1995] and [Luminet, 2015] for details.

is not physically reasonable and thus the implications of his result for physical cosmology are overstated. That is, Manchak has only shown that "[...] the mathematics of General Relativity allows for certain structures, not that these structures are physically relevant, and thus relevant to our actual attempts at modelling the universe" ([Cinti and Fano, 2021], 104). In particular, they argue that the spacetime $(\mathcal{M}', g'_{ab})$ produced by Manchak is "pathological" because it possesses features that lack a physical explanation in terms of some physical process. This is because $(\mathcal{M}', g'_{ab})$, in virtue of its production by the cutting and pasting construction, has singularities in the form of deleted boundaries of 3-spheres.[44] However, no physical explanation is provided for the existence of these singularities, e.g., gravitational collapse of a body. In fact, by Manchak's own lights, no such physical explanation could ever be produced because such a physical process would violate another property of $(\mathcal{M}', g'_{ab})$ stated in his theorem, i.e., that $(\mathcal{M}', g'_{ab})$ and $(\mathcal{M}, g_{ab})$ are locally isometric.[45] Clearly, if there were such a physical process responsible for the singularities in $(\mathcal{M}', g'_{ab})$, there would be no corresponding process in $(\mathcal{M}, g_{ab})$, thereby violating the local isometry of $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$. The authors continue by isolating a particular formal condition, *local b-boundary inextendibility*, which they argue characterizes a physically reasonable spacetime, and show that Manchak's $(\mathcal{M}', g'_{ab})$ fails to satisfy this condition.[46]

I find the discussion of [Cinti and Fano, 2021] quite compelling. As such, we should conclude that, *if* strong evidence supporting a particular MCM emerges, e.g., matching circles in the CMB, then we should accept this evidence on its face rather than appeal to Manchak's skeptical scenario. However, as I have noted, there are many reasons to doubt that such evidence is forthcoming. Furthermore, [Cinti and Fano, 2021] bolsters the aims of this chapter insofar as it indicates the importance of topological underdetermination.[47] This is because I have not merely provided "possible geometric objects which might be studied in General Relativity," but rather have shown that there are spacetimes that are both physically reasonable and observationally indistinguishable. Thus, I have produced a case of "genuine" underdetermination in physical cosmology.

I would like to close this section by making our statement of the Cosmological Principle more precise. This is because all MCMs[48] violate the *global* isotropy of space, i.e., there will be "principal" or "preferred" directions in space reflecting the identification of faces of the fundamental polyhedron under the action of group Γ.[49] However, these preferred directions will only be present beyond a particular length scale, and so "local isotropy everywhere" still holds in MCMs.[50] This local isotropy, along with spatial homogeneity, ensures consistency with our best data. Thus, we have identified precisely the assumption that would rule out

[44] My reader is encouraged to consult [Manchak, 2009] and [Cinti and Fano, 2021] for full technical details.

[45] Two spacetimes $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ are said to be locally isometric if, for each $p \in \mathcal{M}$, there is an isometry from a neighborhood U of p to a neighborhood U' of $p' \in \mathcal{M}'$. Local isometry guarantees that the local physics for each observer in $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ will be the same.

[46] See [Cinti and Fano, 2021], p. 109 for technical details.

[47] Indeed, in their conclusion, they suggest that results about topological underdetermination would be "extremely valuable."

[48] With one exception, real projective space, $\mathbb{RP}^3$. See [Luminet and Lachièze-Rey, 1995].

[49] Indeed, the violation of global isotropy by MCMs is a consequence of the mathematical fact that the group of isometries of most MCMs is smaller than that of its universal covering space. See [Luminet and Lachièze-Rey, 1995], Section 9.3 and [McCabe, 2004], pp. 551-6.

[50] See [McCabe, 2004] for a definition of local isotropy everywhere.

MCMs, viz., the imposition of global isotropy. Providing an argument for this claim would be difficult indeed, and typically something weaker is assumed ([Ellis, 2007], [Beisbart, 2009]). In any case, given the difficulty in justifying global isotropy and the potential benefits imparted by MCMs (see below), our Cosmological Principle should take the form of "spatial homogeneity + local isotropy everywhere."

I will now argue that, though an observational underdetermination of spatial topology seems unavoidable, there are nonetheless reasons to prefer MCMs thereby avoiding a robust underdetermination. I consider reasons that exploit the relationship between multiply connected topologies and *finite* universe models compatible with our best data.

3.6 Issues Concerning the Infinite

It is a truism that many questions arise when we countenance an infinite universe. Can the actual infinite be instantiated in the physical world? What are some philosophical implications of an actually infinite universe? I cannot, of course, do justice to these questions here; rather, I wish to illuminate how they relate to our topological underdetermination. The logical flow of the argument is as follows: by Thesis 3.5.1, the topology of space is observationally underdetermined. However, in virtue of the relationship between the infinitude-finitude of the spatial universe and spatial topology, MCMs enjoy theoretical virtues that SCMs do not. In particular, here I will argue that MCMs possess a particular kind of simplicity because there are spatially *finite* MCMs consistent with our best data.[51,52] In virtue of this finitude (and hence simplicity), particular MCMs avoid difficulties latent in cosmological models that postulate the existence of an actually infinite universe.[53] Thus, we have good reasons for preferring MCMs, thereby providing a means for dispensing with a robust topological underdetermination, viz., an underdetermination in which all epistemic reasons underdetermine the choice of topology.

Before turning to my main discussion concerning the relationship between spatial topology and the size of the universe, I must clear some ground. In recent years, various authors have claimed that it is a direct consequence of inflationary theory that the universe is actually infinite, and this claim has been used to support substantive philosophical conclusions.[54] However, the claim that inflationary theory implies an actually infinite universe relies upon a feature of the idealized mathematical model of inflation that is difficult to support on either physical or philosophical grounds.[55] In virtue of this fact, along with the various ontological and physical scruples one might have about inflationary theory, we should resist arguments from inflationary theory for the infinitude of the universe.

[51] In particular, consistent with the near flatness of the universe thus far observed, i.e., the value of spatial curvature is $k \approx 0$.

[52] Another very important consideration that could be discussed under the auspices of "simplicity" is the fact that a universe with compact spatial sections would exclude various cases of the classical multiverse. See [Ellis and Larena, 2020]. This is deserving of an independent discussion, so I set it aside for future work.

[53] In particular, as discussed above, if our topology is simply connected, the only cosmological models consistent with our best data about the curvature of space are spatially *infinite*.

[54] For a discussion of inflationary theory, see [Guth, 1981]. For a discussion of possible philosophical consequences, see [Knobe et al., 2006].

[55] See my discussion in Appendix D following [Ellis and Stoeger, 2009]

Consequently, it would appear that the topology of the spatial sections will be the primary determinant of the size of the spatial universe (constrained by the value of curvature k). As such, our choice of topology is immediately related to questions of the cogency of an actually infinite universe. Thus, if there are serious philosophical and physical problems with an infinite universe, this should motivate the choice of a topology that avoids ascribing an actually infinite size to space.

3.6.1 Actual Infinities in Cosmology

I would like to begin by considering [Brundit and Ellis, 1979] in which a particularly strange consequence of an actually infinite universe is dramatized. Although the conclusion drawn by the paper is not my primary objection to positing actually infinite universe models, the discussion serves as a useful starting point for a few reasons. First, it is one of the few articles that emphasizes the difficulties latent in the use of infinite universe models.[56] Second, the authors realize the possible theoretical dividends of positing a multiply connected topology, though they dispense with this option much too quickly.

Brundrit and Ellis argue that, provided we live in an FLRW universe that is nearly flat ($k \approx 0$), it is highly probable that[57]

> [...] there exist infinitely many worlds on which there are 'duplicate' populations (i.e., populations identical in number and genetic structure) to that on our own world ([Brundit and Ellis, 1979], 37).

Furthermore,

> It soon becomes clear that it is difficult to provide a *precise* argument against there existing elsewhere in the Universe an identical person reading the identical article on 'Life in the Infinite Universe' [...] for, with an *infinite* family of histories to look at, it is difficult to provide an incontrovertible argument as to why a particular history should occur only once. [...] There is no need to postulate some hypothetical statistical ensemble—it exists in the infinite universe! ([Brundit and Ellis, 1979], 38)

This is a remarkable and somewhat dizzying conclusion. Nonetheless, though the reduplication is odd and unsettling, its strangeness need not incline us to posit one model over another. However, I will show that there are a number of theoretical disadvantages that come with the postulation of an actually infinite spatial universe and the attendant possibility of infinite populations and infinitely much matter.

[56] G.F.R. Ellis makes similar remarks in [Ellis, 2007] and [Ellis, 2014]. For a more recent discussion, see [Ellis et al., 2018]. This paper distinguishes "placeholder" and "essential" uses of infinity in mathematized physical theories. The authors suggest that our best physical theories, even at the most fundamental level of analysis, should not involve essential uses of infinity, i.e., the use of actual infinities. This is because actual/essential infinity satisfies relations that "cannot occur in physical reality; in essence, it fails to obey conservation laws" ([Ellis et al., 2018], 770).

[57] For the details of the argument, see [Brundit and Ellis, 1979], pp. 37-8.

Before detailing these disadvantages, I would like to examine the options available to us for dispensing with them. First, further developments in quantum theory or in our understanding of how "very small" variations of parameters may affect population formation might be of use. A survey of these possibilities would, however, merit an extensive independent discussion. Brundrit and Ellis themselves propose that one could: (i) deny the application of probability theory to scientific models; (ii) deny the Cosmological Principle; (iii) assume that the spatial sections have multiply connected topologies. Option (i) seems incredibly restrictive and should be set aside. Option (ii), though a licit possibility, is a standing assumption of this chapter and is employed in much cosmological research. Finally, option (iii), as we have seen, provides particular universe models in which the spatial sections are *finite*, thereby disrupting the reduplication argument and the postulation of infinite quantities. It is set aside in [Brundit and Ellis, 1979] because

> [T]here appears to be no philosophical reason—based on the uniformity principle, or any other principle—why the space-sections should not have their 'natural' [simply connected] topologies (40).

I would like to press on this reasoning. Indeed, it seems that this is a case in which appeal to the classical theoretical virtue of simplicity may be of use. Of course, the notion of "simplicity" is extremely slippery and multi-faceted (though no more so than that of "naturality"). As such, let us try to be a little more precise.[58] Two distinct sorts of simplicity are commonly acknowledged: ontological and syntactic. Ontological simplicity is given by the number and complexity of entities postulated by a theory, while syntactic simplicity is given by the number and complexity of the laws of the theory.[59] Following Quine, these two aspects are often thought to be inversely related: an expansion in ontology usually results in a contraction in laws, while an expansion of hypotheses/laws usually results in a contraction of entities. However, as we shall see, this relationship is highly non-obvious and difficult to evaluate.

Another distinction is also required: when considering ontological simplicity, one can have either a *qualitative* species or a *quantitative* species. In the former, the number of *kinds* (however one construes a "kind") is minimized, while in the latter the number of entities *simpliciter* is minimized. I am here concerned with the quantitative ontological simplicity of MCMs. Let me now elucidate the theoretical advantages afforded by this simplicity.[60]

First, one might think that the postulation of *actual* infinities is unscientific in the sense that an actually infinite quantity is untestable.[61] So, if one thinks that science should deal with statements that are at least *in principle* testable, actual infinities, e.g., actually infinite spatial sections, actually infinite populations of organisms, should be avoided. Put differently: our best science should not postulate entities that are, by their nature, beyond

[58]The literature on what, exactly, "simplicity" amounts to is vast. My remarks on the notion of simplicity employed here are, of necessity, skeletal. See [Baker, 2016] for the following distinctions and discussion. See also [Sober, 2015] for an extended discussion of simplicity.

[59]In our case, the models of a theory.

[60]For some recent discussions that support conclusions about space and time friendly to my own (though from more metaphysical perspectives) see: [Nolan, 2022], [Sorensen, 2014], [Tallant, 2013]. For a more general consideration of quantitative simplicity (or parsimony), see [Nolan, 1997].

[61]See [Ellis et al., 2018], p. 772.

the scope of science itself. Furthermore, it would seem that, since actual infinities are unobservable and untestable, any scientific theory that makes use of actual infinities (in an essential way) is committed to some form of mathematical platonism.[62] Platonism, as is well-known, involves a host of philosophically questionable theses, most notably that abstract entities are assumed to exist independently of any means of human definition, construction, or observation. Furthermore, on its face, platonism is at odds with scientific realism, especially its epistemic component. I think it preferable to avoid these issues whenever possible.

Second, if one finds this response too hard-nosed, there are a few specific consequences of an infinite universe that would require explanation: most obviously, the generation of infinitely much space at the time of the Big Bang. Furthermore, since we are working under the auspices of FLRW models, we will also have to explain the generation of infinitely much matter (due to the homogeneity of these models).[63] Of course, one might then ask why the existence of infinitely much matter and space is more in need of explanation than the existence of finitely much. I would then be inclined to fall back on the first consideration: the existence of finitely much matter and space is *in principle* testable, while this is not true in the infinite case.

Finally, considered in the context of extant theories of quantum gravity (and more broadly grand unification programmes in physics), the existence of actually infinite space is problematic. In particular, if one subscribes to string theory, then one deals with additional "small" compactified dimensions. Without any reason for thinking that our three "ordinary" spatial dimensions are distinguished in some way, it seems much more reasonable that our "ordinary" dimensions are also compact, as given by a finite MCM.[64]

Objection: Syntactic Complexity

However, following the Quinean insight above, the use of an MCM will require the postulation of additional laws and hypotheses, thus increasing the syntactic complexity of our cosmological model. Starting from a strictly mathematical perspective, an MCM is rather more complicated than an SCM: we must pass to the universal covering space, consider which polyhedra tile the space, etc. This requires the use of various theorems linking SCMs, MCMs, and their geometries. Even from a physical perspective, an MCM could generate a further complexity: the need to reinterpret particular observations of radiation as "ghost images." Namely, in an SCM, there is (generally) a 1-1 correspondence between an object in space and an event in spacetime, e.g., a star produces observable radiation.[65] In an MCM this correspondence fails; there are, rather, multiple spacetime events associated with a single entity in space. However, unlike the complexities generated by positing an infinite universe, e.g., why infinitely much space and matter, we have a good handle on the mathematics

[62] This remark applies to many contexts, especially in physics, since the real numbers $\mathbb{R}$ are constantly used. For a nice discussion of related issues see [Feferman, 1998a]. It goes far beyond the scope of this paper to address the relationship between the postulation of infinite space and the use of continua; I set this aside for future work.

[63] In essence, matter does not occur in "distinguished" regions of the universe. Thus, since matter is uniformly distributed throughout an infinite universe, there is infinitely much matter.

[64] For an excellent survey on quantum gravity, see [Rovelli, 2008].

[65] There is one case of a simply connected space in which this 1-1 correspondence fails.

and physics underlying the added syntactic complexity. Thus, I am inclined to think the ontological simplicity gained outweighs the additional syntactic complexity of MCMs.

Finally, adopting an MCM would raise the question: why *that* particular multiply connected topology (with its particular fundamental polyhedron of particular size)? However, if we no longer expect spatial topology to be empirically determinable, I do not see why we should expect our best cosmology to answer this question. Indeed, no explanation is provided for the preference of a simply connected topology, and so this issue is a wash.[66]

In sum, I propose that we have good reason to prefer MCMs over SCMs in virtue of their simplicity (of the kind indicated) and the advantages this simplicity affords our physical theorizing. Thus, though we have a topological underdetermination by observational data, we can still avoid a robust underdetermination by attending to theoretical virtues. More explicitly, consider Thesis 3.5.1 once more. Let $(\mathcal{M}, g_{ab})$ be an FLRW model with simply connected topology and let $(\mathcal{M}', g'_{ab})$ be an FLRW model with multiply connected topology such that $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ are non-isometric and observationally indistinguishable. I have argued that we should select $(\mathcal{M}', g'_{ab})$ as our preferred model in virtue of the fact that such a model could have finite spatial sections consistent with our best data. In virtue of this spatial finitude, $(\mathcal{M}', g'_{ab})$ enjoys the sort of simplicity discussed above and avoids the difficulties presented by models that employ actually infinite spatial sections. Thus, we have good theoretical reasons to prefer MCMs, and thus we need not resign ourselves to a robust topological underdetermination.

3.7 Einstein and Mach's Principle

Another consideration that might dispense with a robust topological underdetermination involves *Mach's Principle*. In its original formulation, Mach's Principle was developed as a challenge to a substantivalist conception of space.[67] Mach's central idea was that the inertia of a given body derives from its relation to the "fixed stars" and matter throughout the universe rather than its relation to the "absolute space" postulated by Newton. In more modern terminology, we might say that Mach's Principle is the claim that *all inertial properties of an object are entirely determined by the distribution of mass-energy throughout space*.[68] I now turn to an argument originating from Einstein that relates Mach's Principle to the extent of space (and thus to its topology).

In *The Meaning of Relativity*, Einstein provides three arguments "against the conception of a space-infinite" ([Einstein, 1950], 107). The first argument is of the same sort I have offered above: namely, from the standpoint of general relativity, the postulation of a finite universe is "very much simpler" (though he does not give any indication of what this is supposed to mean) than the infinite case. He does not mention topological considerations, but, as we have seen, multiply connected topologies appear to be the only possible way that the universe could turn out finite.[69] In any case, I take this argument to be further grist for

[66] Again, pending a workable theory of quantum gravity, which would hopefully provide insight into the topological structure of the early universe.

[67] See the classic account given by [Sklar, 1974].

[68] See the various formulations given in [Wheeler, 1964].

[69] And, again, consistent with our current value of k.

my mill.

Let us now turn to his second argument from Mach's Principle:

> But in the second place the theory of relativity makes it appear probable that Mach was on the right road in his thought that inertia depends upon the mutual interaction of matter. For we will show in the following that, according to our equations, inert masses do act upon each other in the sense of the relativity of inertia, even if only very feebly. What is to be expected along the lines of Mach's thought? (*ibid.*, 100)

Einstein then proceeds to list three consequences of Mach's Principle and demonstrates that, internal to the formalism of general relativity, these consequences are borne out.[70] Given his derivation of these consequences, he concludes that our interpretation of general relativity should validate Mach's Principle. He continues,

> The idea that Mach expressed, that inertia depends upon the mutual action of bodies, is contained, to a first approximation, in the equations of the theory of relativity; [...] But this idea of Mach's corresponds only to a finite universe (*ibid.*, 107-108).

The point here is one that we have already seen. Einstein's Equation does not specify *solutions* that can be studied from a physical standpoint. Further assumptions, sometimes expressed as "boundary conditions at infinity," are required. However, these boundary conditions simply replace the role played by Newtonian absolute space. This is because, when boundary conditions are specified, it is no longer simply the presence of mass-energy that determines the geometry of spacetime, and hence mass-energy no longer determines the inertial properties of any given object. For instance, general relativity appears to validate the idea that, in a universe without matter, there is nonetheless an "absolute" flat (Minkowskian) spacetime structure, as dictated by the imposition of natural Minkowskian boundary conditions.

In response to these concerns, Einstein suggested a radical way to deal with the problem: simply postulate a finite universe model, thereby obviating the need for boundary conditions at infinity that contradict Mach's Principle. Concisely, we might say, "[I]t is likely that the requirement of Mach's Principle is identical with the requirement of a finite universe" ([Wheeler, 1964], 306). Thus, since our best experimental evidence indicates that the universe is not sufficiently dense to force finiteness, the presence of a multiply connected spatial topology would be the only way to produce a finite universe model. And so, if one is convinced by (some version of) Mach's principle, one should be compelled to select a multiply connected spatial topology.[71]

The reasons for preferring a multiply connected topology that proceed from Machian considerations are not obviously of the same sort adduced above. It is, however, possible to think in these terms, since we get both an ontological and syntactic simplicity. The former is clear; the latter occurs because, instead of having to specify boundary conditions for each

[70] As Einstein notes, these effects are so minuscule as to rule out the experimental confirmation.

[71] For extremely brief mention of "Machian considerations," see [Fagundes, 1983], [Ellis and Schreiber, 1986], and [Ellis, 2007].

possible solution of Einstein's Equation, we add a single stipulation (multiply connected topology) to the mathematics of our physical theory that eliminates the need for boundary conditions. Nonetheless, as in both Mach's case and Einstein's case, the reason for eliminating boundary conditions is not a desire for simplicity, but rather a preference for a highly plausible metaphysics of space and characterization of inertial properties (as well as the purely formal consequences derived by Einstein).

3.8 The Explanatory Power of MCMs

To conclude my discussion of topological underdetermination, I would like to consider the explanatory power of positing a finite MCM. I examine two sorts of recent cosmological data for which a finite MCM would appear to be a good (or even the best available) explanation.

Following the common practice of cosmologists, we have assumed the Cosmological Principle and thereby have considered FLRW models. However, as we have noted, the assumption of the Cosmological Principle is an *a priori* prescription imposed on all possible models of general relativity. To what extent is the spatial isotropy and homogeneity[72] assumed in the Principle observationally justified? It turns out that spatial isotropy about our location is strongly supported by observational evidence, consisting of both the observation of luminous sources and the CMB radiation discussed in Section 3.5. As George Ellis has noted,

> Considered on a large enough angular scale, astronomical observations are very nearly isotropic about us, both as regards source observations and background radiation; indeed the latter is spectacularly isotropic, better than one part in 10^4 after a dipole anisotropy, understood as resulting from our motion relative to the rest frame of the universe, has been removed. Because this applies to all observations (in particular, there are not major observed matter concentrations in some other universe region), this establishes that in the observable region of the universe, to high accuracy *both the space-time structure and the matter distribution are isotropic about us* ([Ellis, 2007], 1225).

Clearly, this isotropy (and homogeneity) cannot be explained by the commonly used FLRW models, as such models assume these properties. It would be desirable, however, to have some sort of explanation for them. It is commonly postulated that spatial homogeneity results from an inflationary period in the early universe[73]; however, it has been shown that inflation would only produce the observed homogeneity if the pre-inflationary universe were *already* sufficiently homogeneous ([Luminet and Lachièze-Rey, 1995]). Thus, inflationary theory only pushes the problem back. As usual, a fully developed theory of quantum cosmology and quantum universe formation would deal with this issue, but no such theory is yet operative.

However, the postulation of a finite MCM seems a reasonable and readily available explanation of the phenomenon of homogeneity.[74] More precisely, a finite MCM with especially small volume would produce the special initial conditions necessary for a "chaotic" (in the

[72]Spatial isotropy along with some version of the so-called *Copernican Principle*, i.e., we are not distinguished observers, implies spatial homogeneity. See [Ellis, 2007], Section 4.2.2 for discussion.

[73]See, especially, [Guth, 1981] and [Gibbons et al., 1983]

[74]This possibility is hinted at in [Ellis and Schreiber, 1986].

sense of inhomogeneous) early universe to transition to a homogeneous one. The details here are extremely complex and depend upon the dynamics of the precise models considered, but the basic point is that, at a sufficiently young age, a small MCM is causally connected. In virtue of causal connectedness at an early age, the universe can homogenize before the the scattering of the CMB. Thus, the isotropy of the CMB reflects the homogeneity of the early universe, which is itself explained and made possible by a causal process occurring in a sufficiently small MCM.[75] It is worth noting that more calculations have to be done before MCMs can confidently discharge this explanatory role. Nonetheless, they serve as one of the best available explanations of the homogeneity and isotropy of space.

I would like to close with what I take to be the strongest explanatory function of finite MCMs. As we have noted, the key data for observational cosmologists are surveys of the CMB.[76] It was hoped that CMB data could decisively reveal the spatial topology of the universe via such methods as circles in the sky; however, we have adopted a position of justified pessimism towards this endeavor. Nonetheless, particular measurements of the CMB may provide reasons to posit a finite MCM. I will now show why this is so.

Of particular interest are temperature fluctuations (anisotropies) in the CMB as these can yield information about the physical conditions of the early universe. (It is worth noting that these anisotropies are, on the whole, minuscule, so they in no way impugn the large scale isotropy of the CMB; see Ellis's quote above.) It is believed that the early universe was crossed by acoustic waves soon after the Big Bang and, in turn, these waves left imprints on the universe ($\approx$ 380,000 years after the Big Bang) as density fluctuations in the primordial plasma. The anisotropies detected in the CMB reflect these density fluctuations, which can be mathematically constructed from vibrational modes of space.[77]

Suppressing the mathematical details of the spherical harmonics, the crucial quantity for measuring anisotropies is the *full-sky two-point correlation function* of temperature fluctuation $\delta T(\hat{\mathbf{n}})$, observed for our sky in the direction of unit vector $\hat{\mathbf{n}}$. This function is written as

$$C^{\text{obs}}(\theta) := \langle \delta T(\hat{\mathbf{n}}), \delta T(\hat{\mathbf{n}}') \rangle \tag{3.8.1}$$

where $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}' = \cos\theta$. The brackets denote averaging over directions $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}'$ separated by angle θ. Using CMB datasets, values of $C^{\text{obs}}(\theta)$ have been computed for $0 \leq \theta \leq 180$ (degrees). However, there are a number of "intriguing discrepancies" between the observational values of $C^{\text{obs}}(\theta)$ and predictions of the "standard" cosmological model, which is flat, infinite, and Euclidean with simply connected topology (this model is often written as ΛCDM). In particular, for angular scales over 60 degrees, there is very little correlation between the CMB observations and ΛCDM simulations. As has been noted in very recent studies, especially [Aurich et al., 2021], this discrepancy finds "a natural explanation in cosmic topology." That is, a finite MCM would make good sense of this discrepancy because the spatial sections are finite and so space is not large enough to support the longer wavelengths produced by larger

[75] For full details see [J. Richard Gott, 1980]. For an investigation of a different possible model see [Hayward and Twamley, 1990].

[76] The most recent space probe missions are WMAP (Wilkinson microwave anisotropy probe; active life-time 2001-2010) and Planck (Planck probe; active life-time 2009-2013).

[77] For exhaustive details see [Levin, 2002].

angles. Indeed, *as of yet*, there appears to be no other explanation of this phenomenon of "angular power spectrum suppression."[78]

It is important to note, however, that this does not impugn our underdetermination thesis (Thesis 3.5.1). The observational constraints and extreme sensitivities of even our best methods (circles in the sky, statistical techniques for anisotropies) warrant a healthy does of skepticism. Nonetheless, especially given the discussion of anisotropies, finite MCMs remain a live *possibility* and, furthermore, provide the best explanation for our current data.

3.9 Conventionalism: Geometrical and Topological

In light of the underdetermination of spatial topology by observational data, what then are we to say of it? Provided that we accept the last century of work in relativity theory, we must accept that spacetime has (or is well-modeled by) a manifold structure, and this manifold structure must be equipped with particular topological properties. Indeed, these properties are required for the very cogency of relativistic cosmology.[79] A natural suggestion, especially in light of the long and lively dialectic concerning the conventionality of spatial geometry, is that topological properties of space are "conventions."[80] Namely, because no empirical evidence can compel us to choose one topology over another, we simply make a choice, motivated by the concerns mentioned above: simplicity, explanatory power, etc. It is, however, not entirely clear what a topological conventionalism amounts to or what the precise kinds of arguments for favoring it are. Thus, I will first chart out a few possible arguments, taking the (mathematically) simpler and better known geometrical conventionalism of Poincaré as a guiding example. Indeed, this will also be of use as my ultimate position concerning the epistemic status of spatial topology will share many features with Poincaré's own complex conception of the role of "conventions" in a physical theory.

3.9.1 Poincaré's Geometrical Conventionalism

Geometrical conventionalism is usually taken to originate in Poincaré's famous epistemological critique of geometry in *Science and Hypothesis*[81] and is principally motivated by the following, now well-known, thought experiment.

Imagine a finite Euclidean sphere of constant temperature T_c at its center and absolute zero at the surface. The temperature T_p of any given point p between center and surface varies according to the gradient $R^2 - r^2$, where R is the radius of the sphere and r is the distance of the given point p from the center. Now imagine that the sphere's inhabitants are trying to ascertain empirically the geometry of their space using rigid rods. Importantly, the length of these rods is susceptible to variations in temperature; in particular, as the

[78] For a very convincing and thorough discussion see [Aurich et al., 2021]. See also the earlier paper [Aurich et al., 2008].

[79] As already noted, it has simply been assumed by more or less all cosmologists over the course of the 20th century (modulo those who took an interest in cosmic topology, especially G.F.R. Ellis, J.P. Luminet, and M. Lachièze-Rey) that spatial topology is simply connected. That is to say, physicists have not remained agnostic concerning spatial topology.

[80] Topological conventionalism originates in [Reichenbach, 1957]. See below.

[81] See [Poincaré, 2015] and [Poincaré, 2017].

temperature decreases, the rods proportionally contract in length. This is unknown to the inhabitants who proceed to measure distances between points using these rods. Say that one inhabitant initially uses a rod of length $\ell = 1$ meter and attempts to measure the distance from center c to point p. But observe what happens: as they continue laying down their rod, $r \to R$, $T \to 0$, and $\ell \to 0$. Thus, they will think that they live in an infinite,[82] hyperbolic space, when, in fact, they live in a finite, Euclidean one! Thus, the inhabitants of this world could assert, consistent with all empirical evidence, either that they live in a finite Euclidean space equipped with a special temperature field or that they live in an infinite hyperbolic space without such a field. Furthermore, Poincaré allows that the inhabitants have access to light rays, suitable for triangulating their space. Analogous to the shrinking of the rods, these light rays are refracted through their medium of transmission proportional to $1/(R^2 - r^2)$. The inhabitants assume that, as usual, light travels along geodesics (paths of shortest distance) and their index of refraction in a vacuum is constant. Thus, they could again describe the geodesics followed by the light beams as either straight lines in hyperbolic geometry or as circular arcs in Euclidean geometry.

From this thought experiment, Poincaré concludes that experience is compatible with various geometries, and so we must choose which geometry we will use to describe physical phenomena. This choice is, however, far from arbitrary. First, he argues that our choice is restricted to the three geometries of constant curvature: Euclidean, elliptic, hyperbolic.[83] (Note that, given our own physical assumptions above, we are also limited to these three geometries.) Second, he argues that we have good reason to select Euclidean geometry, and this choice is motivated by concerns of simplicity and convenience:

> *One geometry cannot be more true than another; it can only be more convenient.* Now, Euclidean geometry is, and will remain, the most convenient: (1) Because it is the simplest, and it is so not only because of our mental habits or [because] of I know not what direct intuition we may have of Euclidean space; it is the simplest in itself, just as a polynomial of the first degree is simpler than a polynomial of the second degree [...]; (2) Because it sufficiently agrees with the properties of natural solids, those bodies which our hands and our eyes compare and with which we make our measuring instruments ([Poincaré, 2017], p. 80; emphasis my own).

We might summarize by saying that we have no epistemic (in the sense of truth-tracking) reasons to prefer Euclidean geometry, but we do have practical reasons that are sufficient to determine our choice. Note that there are definitely some unclarities here.[84] How, exactly, are we to make sense of simplicity "in itself"? Why think the *physical theory* employing Euclidean geometry is simpler than the *physical theory* employing hyperbolic geometry? I discuss this further below where I attempt to provide a sharpening of Poincaré's argument in the case of spatial topology.

[82] Poincaré asserts that the measurer will think they live in an infinite space; however, as we have seen, one could live in a finite space, without any "distortions", and still think space is infinite. Namely, space could be finite but still "too large" to measure.

[83] Thus, he makes essential use of the Helmholtz-Lie theorem. See footnote 89.

[84] As has long been recognized. See, for example, [Sklar, 1974], p. 93.

A more immediate concern is how we are to interpret the above thought experiment in order to establish Poincaré's desired conclusion. Here are a few possible readings.

Reading I: From Observational Equivalence. Since the sphere interpreted with a particular set of physical laws and hyperbolic geometry $(\mathcal{L}_1, \mathcal{H})$ is observationally equivalent to the sphere interpreted with a different set of physical laws and Euclidean geometry $(\mathcal{L}_2, \mathcal{E})$, no empirical facts can determine whether the geometry is hyperbolic or Euclidean. Thus, only conventional choice (and not matters of fact) can determine spatial geometry. Such an argument from observational equivalence was favored by the logical positivists in their own adoption of Poincaré's geometrical conventionalism. However, as Friedman points out in Chapter 4 of [Friedman, 1999], this argument has nothing to do with geometry in particular and can be applied to any part of a physical theory. Indeed, the crucial ingredient to this argument is Duhemian holism: our physical theories (physical laws and geometry) confront experience as a whole. Thus, the only fact about geometry that renders it conventional is that it is a part of our complete physical theory. This is certainly not a reasonable interpretation of Poincaré (as Friedman argues) and is generally unpalatable. For indeed every part of a physical theory could be rendered conventional, and *geometrical* conventionalism would seem to be of little independent interest.

Reading II: From Observational Equivalence and Further Geometrical Facts. One way to block the move to a thoroughgoing conventionalism would be to appeal to specific geometrical facts that produce the observational equivalence. I have in mind something like the following. It is well known that one can construct models of hyperbolic geometry within Euclidean geometry,[85] and thus we have a precise and systematic way to move between these different systems. This can then, in turn, suggest to us the changes to our physical laws that would be needed in order to get an observational equivalence between $(\mathcal{L}_1, \mathcal{H})$ and $(\mathcal{L}_2, \mathcal{E})$, i.e., the temperature gradient and contraction of objects. Indeed, it seems likely that this is precisely how Poincaré proceeded.[86] Thus, there is something distinguished about the geometrical parts of our physical theories that drives the observational equivalence and thus generates the conventionality of geometry. I will return to this kind of argument below when I discuss topological conventionalism.

Poincaré himself constructs a "dictionary" between hyperbolic (Lobachevskian) geometry and Euclidean geometry with the aim of obtaining theorems of hyperbolic geometry from those of Euclidean geometry.[87] However, this argument does not on its own establish the conventionalist thesis. All it does is show that we can produce a model of hyperbolic geometry within Euclidean geometry, which is merely a feature of formal geometrical systems and not physical theories as a whole.[88] Thus, it seems that the most convincing way to interpret

[85] Thereby establishing the relative consistency of hyperbolic geometry, i.e., the question of the consistency of the system of hyperbolic geometry reduces to the question of the consistency of the system of Euclidean geometry.

[86] See [Torretti, 1978] on Poincaré's study of Fuchsian functions and hyperbolic geometry.

[87] See Chapter 3, [Poincaré, 2017].

[88] See [Ben-Menahem, 2001] where this is noted and used to clarify the relationship between Poincaré's various arguments for conventionality. See especially pp. 486-488 where she suggests, entirely in line with my point above, that "Geometric equivalence does not entail physical equivalence, but provides direction on

Poincaré's thought experiment is to appeal to both the formal features of geometrical systems and how such formal features interact with physical laws and observations.

Finally, a sort of conventionalism that is always available and need not be explicitly tied to the above thought experiment is:

Reading III: Semantic Conventionalism. Under this reading, we note that the truth of any geometrical statement will be dependent upon how we have defined "distance." Since the meaning we assign to this word is a matter of convention, so the argument goes, the truth of any geometrical statement will be conventional. Thus, the body of truths that comprise a geometrical system will be conventional. As with **Reading I**, this appeals to nothing in particular about geometry, and thereby also leads to a thoroughgoing conventionalism.

It is not my intent to judge whether any of these readings of geometrical conventionalism adequately describes Poincaré's position. Indeed, it is almost certain that none does. For Friedman has provided a convincing interpretation of Poincarean conventionalism in terms of Poincaré's hierarchical conception of the sciences and the group-theoretic basis for metrical properties of space.[89] What these readings do provide, however, is a few reasonable starting points for how we might construe conventionalism with respect to a mathematical component of a physical theory.

When we consider topological properties of space, I argue that some version of Reading II could describe our situation, provided we elaborate further on how we are to understand the role of "conventions" in a physical theory. This elaboration will, in fact, lead us back to Poincaré, who interprets "conventions" in a rather strong sense.[90] I find Readings I and III less interesting, in this context at least, for the same reasons noted above: they do not indicate why a *topological* (or geometrical) conventionalism is particularly compelling.[91] Furthermore, I will provide reasons for thinking Reading III in the topological case requires assumptions that are difficult to accept. With these points in mind, let me now turn to a discussion of topological conventionalism.

3.9.2 Topological Conventionalism

Reichenbach's *The Philosophy of Space and Time* ([Reichenbach, 1928], [Reichenbach, 1957]) provides the crucial stepping stone between Poincaré's pre-relativistic views on the conventionality of geometry and my discussion of spatial topology. Reichenbach first provides a modified version of Poincaré's thought experiment in order to obtain the observational equivalence of Euclidean and hyperbolic geometries in the context of general relativity. The crucial change is the following. Poincaré's argument relies on non-uniform, distorting, forces,

how to generate it."

[89] As codified by the Helmholtz-Lie theorem, which provides a classification of the geometries of all (3-dimensional) manifolds of constant curvature (i.e., "free mobility"). For Friedman's argument, see [Friedman, 1999], Chapter 4. [Ivanova, 2015] also contains a related discussion as to how we might square Poincaré's many philosophical and mathematical commitments.

[90] That is, though there is a degree of freedom in our choice of geometry, and this choice is determined by non-epistemic reasons, the conventions we choose play a distinguished role in our physical theories.

[91] I do not deny, however, that they might have some bite in more general discussions of conventionalism.

which, according to Reichenbach, are in principle detectable. Thus, Poincaré's thought experiment does not necessarily establish observational equivalence. Reichenbach's insight was that, in space-time theories, one can introduce uniform or universal forces that are always undetectable, thereby establishing the desired equivalence.[92]

This is all extremely interesting and has received a great deal of attention. However, Reichenbach did not stop there. He goes on to consider whether a conventionalist thesis might also hold for topological properties.[93] The discussion remains extremely good, despite some technical unclarities. Ultimately, though, it requires that we countenance what Reichenbach calls "causal anomalies," viz., we allow causal loops. I have been operating under the assumption (as do many cosmologists) that, though such anomalies are possible, they represent rather exotic spacetime models and are to be ruled out. Interestingly, Reichenbach concludes his discussion of topological conventionalism by saying,

> Topology is an empirical matter as soon as we introduce the requirement that no causal relations must be violated ([Reichenbach, 1957], 80).

This is precisely the question with which we have been concerned. That is, if we accept the usual well-behaved spacetime models, equivalent to some rather strong physical and mathematical hypotheses (e.g., the Cosmological Principle), does spatial topology become empirically determinable, as Reichenbach claims? The answer, thus far, appears to be *no*.[94] Even Reichenbach, who seemed to think that conventionalism could be extended very far indeed, did not envisage that it could be extended to spatial topology (provided we disallow causal anomalies). But this is precisely what we will now consider.

I should also mention an important precursor of my investigation, that of Glymour in [Glymour, 1972]. In this paper, Glymour too takes up the Reichenbachean suggestion and applies it to relativistic cosmology in a more precise manner. Some of my points below will overlap with Glymour's early work; I note these when applicable. There are, however, a few ways in which Glymour's account must be supplemented. First, serious research into cosmic topology was not yet under way at the time of his writing, and this brings with it further interesting possibilities left untouched by [Glymour, 1972]. Second, and relatedly, this recent cosmological work makes clearer the relationship between finitude-infinitude and topology, a connection I will exploit a great deal below in my philosophical discussion of conventionality. Third, the various senses in which Glymour employs "convention" and "conventionality" are not perspicuous (an issue not endemic to his discussion; compare to geometrical conventionalism above). Nonetheless, my conclusions[95] concerning the empirical determination of the universe's spatial topology agree with Glymour's main thesis that

> [F]or each of a class of fashionable cosmological models there is another (unfashionable) model different from the first in the topology it ascribes to space-time, and there are good reasons to think that any two such cosmological models are,

[92] This is a traditional reading of the historical development found in, e.g., [Sklar, 1974]. For more recent evaluations of Reichenbach's geometrical conventionalism (or whether it was such) see [Friedman, 1999] (Chapter 3) and [Gimbel, 2004].

[93] See [Reichenbach, 1957], pp. 58-81.

[94] Again, see the discussion above in Section 3.5.

[95] Again, see Section 3.5.

> both in fact and in principle, experimentally indistinguishable. Any bit of evidence which we can account for with one model, we can account for with another, and conversely ([Glymour, 1972], 196).

In any case, following Glymour, my first step in an argument for the conventionality of spatial topology is to produce two models of spacetime that have different topologies (simply or multiply connected) and yet are observationally indistinguishable. That is, we wish to produce the topological analogue to Poincaré's thought experiment (or Reichenbach's generalization thereof).

More precisely, we produce two FLRW models of spacetime, one with a simply connected topology (SCM), the other with a multiply connected topology (MCM), that are nonetheless observationally indistinguishable (in the sense of Section 3.3). But this has already been done! Consider again the Topological Underdetermination Thesis (Thesis 3.5.1) in which we postulated that for any simply connected FLRW model $(\mathcal{M}, g_{ab})$, we can find a multiply connected FLRW model $(\mathcal{M}', g'_{ab})$ such that these models are observationally indistinguishable and yet non-isometric. Since we are concerned with the topology of space, we select $\mathcal{M}_3 = \mathbb{R}^3$ and $\mathcal{M}'_3 = T^3$. These models are locally the same insofar as for any $p \in \mathcal{M}_3$, there is a $p' \in \mathcal{M}'_3$ such that $I^-(p)$ and $I^-(p')$ are isometric. However, they differ significantly in the large: $\mathcal{M}_3$ is infinite and simply connected, whereas $\mathcal{M}'_3$ is finite and multiply connected. If the Topological Underdetermination Thesis holds, then these models will always be observationally indistinguishable. Furthermore, unlike Poincaré's geometrical thought experiment (or Reichenbach's generalization) we have not needed to resort to imposing further distortions (temperature fields, universal forces) to produce the equivalence. Thus, if one finds those postulates suspect insofar as they produce different ontological commitments, one need not worry about them here.

The question of observational indistinguishability here becomes especially interesting. It seems that there are two possible options. First, $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ could be observationally indistinguishable due to limitations in our capacities as scientific enquirers. This would be a case in which our two models are *in principle* observationally distinguishable, but we fail to produce, e.g., the correct techniques for doing so.[96] That is, the multiply connected topology of $(\mathcal{M}', g'_{ab})$ is in principle detectable, say, by cosmic crystallography or circles-in-the-sky (see Section 3.5), but we always fail to distinguish statistical signal from noise or fail to identify matching circles. Such failure might occur for technological reasons (our instruments are not sufficiently discriminating) or because of insufficient data sets.

This case is, I think, rather close to Poincaré's original thought experiment under **Reading II**. For here, just as with his spherical world, we can use our knowledge of the mathematical part of the theory to reinterpret physical parts of our theory. Namely, just as Poincaré

[96] By "in principle detectable," I mean that the fundamental polyhedron of the multiply connected model is sufficiently small, so as to generate multiple images of a single source of radiation. Recall the UC space with Euclidean geometry in Section 3.5, which corresponds to a two-dimensional version of $(\mathcal{M}', g'_{ab})$. Note that *only* star S, and not each $\{S_1, S_2, \ldots\}$, is a genuine object in spacetime. However, S produces, for observer O, multiple spatiotemporal events registered by $\{S_1, S_2, \ldots\}$, which is an immediate consequence of the multiply connected topology. Radiation from star S scatters in infinitely many directions, wraps around the hypertoric universe, and produces multiple images of S. On the other hand, were we in the other model we could safely assume a 1-1 correspondence between objects in spacetime and observed radiation, i.e., there would be distinct objects (stars, quasars, or what have you) corresponding to each $\{S_1, S_2, \ldots\}$.

exploited the "dictionary" between Euclidean and hyperbolic geometries to discover the contraction law needed to render the two versions of the spherical world equivalent, here we can exploit the mathematics behind the differing topological models to suggest requisite physical changes. That is, presented with some data set, e.g., of luminous sources, we could either assume that we are in $(\mathcal{M}, g_{ab})$ or $(\mathcal{M}', g'_{ab})$. In the first model, we would interpret any observation to be 1-1 correlated to a source, whereas in the second we would reinterpret by not assuming such a correlation. That is, in $(\mathcal{M}', g'_{ab})$ we would reinterpret some observations as indicating the presence of mere "ghost" images generated by the wrapping of light due to the multiply connected topology of the model. In short, under these different topological models, we need to accommodate for a different relationship between observations and objects in our model, whereas in the geometrical case we needed to account for different metrics. In the former case, we do so by giving different referents (genuine object or ghost image) in each model, whereas in the latter we do so by imposing the distorting temperature field. Crucially, both of these procedures exploit the mathematics underlying the physical situation.

The second sort of observational indistinguishability is rather different from that of Poincaré's spherical world. Here we can say that $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$ are observationally indistinguishable not solely because of our epistemic limitations, but rather because of the way the universe is constituted. This would be a case in which the "signature" of the multiply connected topology far exceeds the in principle observable universe. Namely, no wraparound effect/multiple images are even there for detection. In this case we get observational equivalence without having to reinterpret anything in the models. Rather, the equivalence proceeds from the fact that, in either $(\mathcal{M}, g_{ab})$ or $(\mathcal{M}', g'_{ab})$, we are living in an inappreciably (or even infinitesimally) small patch of the universe.

This second sort of observational indistinguishability is Glymour's primary focus. He does not make entirely clear whether the observational indistinguishability in question is epistemic or metaphysical, but the latter seems closer to the mark. For he concludes

> The significant cases of conventional topologies are those in which the world is one way or the other, not both, but things are so arranged that we cannot discover which way the world is ([Glymour, 1972], 216).

In any case, these two different sorts of indistinguishability produce two different arguments for the conventional nature of spatial topology. The first proceeds from an indistinguishability induced by epistemic limitations and by our understanding of the mathematical components of the theory:

1. We have a systematic technique to move between SCMs and MCMs that maintains a local equivalence between $(\mathcal{M}, g_{ab})$ and $(\mathcal{M}', g'_{ab})$;

2. This technique suggests what physical hypotheses have to be reinterpreted in the newly formed MCM;[97]

3. Therefore, spatial topology will always be underdetermined by empirical evidence;

[97] It is an interesting and difficult question as to whether there might be techniques available to distinguish these *prima facie* equivalent models. For instance, an MCM and an SCM could, respectively, contain finitely much and infinitely much matter. Could this difference induce local and observable differences?

4. Therefore, it is *precisely* the topological properties of an FLRW model that are conventional.

This is analogous, following [Ben-Menahem, 2001], to a reasonable Poincarean argument for the conventionality of geometry:

1. We have a "dictionary" systematically relating results in Euclidean and hyperbolic geometries;
2. This dictionary suggests what physical hypotheses have to be reinterpreted in each geometrical model of space;
3. Therefore, spatial geometry will always be underdetermined by empirical evidence;
4. Therefore, it is *precisely* the geometrical features of a physical theory that are conventional.

The second argument for the conventionality of spatial topology is much more straightforward and follows simply from the way in which the universe is constituted:

1. We live on an appreciably small (possibly infinitesimally small) patch of the universe;
2. Spatial topology is thus always beyond even in principle detectability;
3. Therefore, the topology of space, along with much else, must be conventional.

This is closer to **Reading I** of geometrical conventionalism, though here it would seem that it is not Duhemian holism driving the observational equivalence. Rather, it is the global nature of topological properties. So, in particular, there is nothing special about spatial topology that renders it conventional. There are many other global spatial (or spatiotemporal) properties beyond empirical determination, and these could be conventional in the same way.

3.9.3 Assessment of Topological Conventionalism

I think that, in either argument, we find that connectivity properties are conventional in a rather distinguished (or perhaps strange) sense. For it would appear that these properties are closely intertwined with extremely fundamental properties of space and the matter occupying it. This is one reason for thinking that a semantic conventionalism for spatial topology (**Reading III** above) is not promising. In the geometrical case, it does not seem totally unreasonable that two physical theories might have the same observational consequences and yet be employing the word "distance" in different ways. However, in the topological case, we would have to say that there is a systematic ambiguity in ascriptions of identity. This is because an MCM is generated by means of a new identification procedure, e.g., in order to form a hypertoric model, points judged to be distinct on a flat Euclidean model are identified under the action of Γ.[98] Thus, if we wanted to cleave to a semantic conventionalism with respect to spatial topology, we would also have to allow that the models disagree

[98] See [Glymour, 1972], pp. 201-202 for the same remark. This is yet another way in which the models "disagree as to the basic individuals in the universe."

about the identification of points. This seems quite troubling. Furthermore, the fact that an SCM and an MCM might disagree on the finitude of the universe, despite being observationally equivalent, is also quite strange. Finally, if we assume the uniform distribution of matter on average throughout the universe, an assumption of the FLRW models, we could have two observationally indistinguishable models that radically disagree over the amount of matter in the universe. This is because a finite, multiply connected spacetime will have finitely much matter, while an infinite, simply connected spacetime will have infinitely much matter.[99] Thus, a difference in topology makes an enormous difference to the models under consideration, even if they are observationally indistinguishable. This fact indicates that the conventionality at work here is in no way anodyne; these models cannot be the same theory in disguise, as it were, given the close relationship between connectivity properties, ascriptions of identity and finitude, and the amount of matter present in the universe.

3.9.4 A Stronger Sort of Convention

Recent scholarship on Poincaré's geometrical conventionalism provides us with an extremely useful way of understanding topological conventionalism.[100] I will briefly sketch the main points of this work and then apply it to connectivity properties of space. In short, I wish to say that these properties, given what we know at this juncture, are conventions in the sense that no empirical evidence can compel us to choose a simply or multiply connected topology. The best way to understand this fact is that, at a rather abstract and general level of our cosmological theory, we have a particular degree of freedom, represented by the possibility of choosing different spatial topologies consistent with our data. However, these conventions, these degrees of freedom, do not amount to mere arbitrary elements of a theory. Rather, a choice of spatial topology *constitutes* the basic entities of our cosmological theory in extremely different ways, and this constitutive role must be attended to when we do ultimately select a particular topology.

Many recent accounts of Poincaré's work on geometrical conventionalism stress his desire to square a generally Kantian scientific and epistemic framework with the emergence of non-Euclidean geometries. Roughly, his invention of conventions as a new sort of "epistemic category"[101] proceeds from the following line of thought. The axioms of Euclidean geometry must be either analytic *a priori*, synthetic *a priori*, or synthetic *a posteriori* propositions (given Kant's classification of propositions). Poincaré then argues that each class of proposition fails to appropriately capture the nature of the Euclidean axioms.

Thus, Poincaré cannot comfortably accommodate geometrical propositions in the Kantian framework and so calls them "conventions." In light of Poincaré's commitment to a hierarchy of sciences, these conventions take on a particularly interesting valence. As Friedman says,

> [T]he determination of particular physical forces presupposes the laws of motion, and the laws of motion in turn presuppose geometry itself: one must first set up a geometry before on can establish a particular theory of physical forces. We have no other choice, therefore, but to select one or another geometry on conventional

[99] See my discussion of [Brundit and Ellis, 1979] above.
[100] I draw from, in particular, [Friedman, 1999], Chapter 4, [Ben-Menahem, 2001], and [Ivanova, 2015].
[101] This very helpful way of putting the point is found in [Ben-Menahem, 2001].

> grounds, which we can then use, so to speak, as a standard measure or scale for the testing and verification of properly empirical or physical theories of force ([Friedman, 1999], 78).

That is, for Poincaré, the choice of geometry in the construction of a physical theory plays a constitutive role.

Now, of course, Poincaré's conception of geometry as conventional and constitutive is inconsistent with contemporary relativistic cosmology. By virtue of the fact that the curvature of spacetime is determined by the presence of energy-momentum (given by particular exact solutions of the Einstein Equations), geometry becomes an empirical matter. However, as Friedman has noted in his work on the *relativized a priori*,[102] a version of Poincaré's view can be salvaged. That is, though geometry might not play the epistemic role that Poincaré thought, other principles do in fact make possible or constitute the empirical content of our scientific theories. We must, however, weaken the claim that there is a unique, fixed set of such constitutive *a priori* principles, as Kant and Poincaré believed, and understand these constitutive principles as relative to a given scientific theory at a given time.

Let me now turn to this conception of the constitutive or relativized *a priori*. Before doing so, we may summarize our findings in the following way: the topology of space is an excellent candidate for a principle of relativistic cosmology that is both conventional and constitutive. Conventional because there are many topologies compatible with our observational data; this induces a particular degree of freedom at a very abstract level of our scientific theory and requires that a choice be made.[103] Constitutive because the ascription of particular topological properties are required for the cogency of subsidiary physical laws and for the application of fundamental physical concepts.

3.10 The Constitutive or Relativized A Priori

3.10.1 The Basics of Friedman's Account

In an influential collection of writings,[104] Michael Friedman has articulated a neo-Kantian picture of scientific knowledge designed both to accommodate Kuhn's theory of scientific revolutions and to resist the pull of Quinean epistemological holism. In particular, Friedman envisages a dynamical and hierarchical conception of knowledge consisting of three levels. At the lowest level sit empirical principles of natural science, namely, those principles susceptible to empirical confirmation or disconfirmation, e.g., Newton's law of gravitation or exact solutions of Einstein's Field Equations. At the next highest level sit the so-called "constitutive" or "relativized" *a priori* principles which define the fundamental framework internal to which empirical principles are tested. As Friedman says, these are

> [...] relatively stable sets of rules of the game, as it were, that define or make possible the problem solving activities of normal science ([Friedman, 2001], 45).

[102] See, in particular, [Friedman, 2001].

[103] See the discussion in Sections 3.6, 3.7, and 3.8 for how such a choice might be made.

[104] See especially [Friedman, 1999] (Ch. 3), [Friedman, 2001], [Friedman, 2007], and [Friedman, 2009]. I follow most closely the presentation in [Friedman, 2001].

Furthermore, though these constitutive principles are not fixed for all time, since it is precisely these principles that are changed in periods of "deep conceptual revolution," they are not straightforwardly empirical. This is because, in the absence of an agreed upon set of constitutive *a priori* principles, there is no agreed upon process of empirical verification. It is only internal to a particular constitutive *a priori* framework that various observations could be said to count as confirmation (or disconfirmation) of a particular empirical law. For instance, one must first lay down the mathematical framework of pseudo-Riemannian manifolds (along with auxiliary hypotheses[105]) before one can claim that solutions of Einstein's Equation are either *empirically* true or false. Only then, for example, does Einstein's calculation of Mercury's perihelion "count" as empirical evidence for the theory of general relativity as codified by the Einstein Equation.[106]

Finally, the third level of Friedman's hierarchy consists of "philosophical meta-paradigms or meta-frameworks," which serve to effect the transition from one constitutive *a priori* framework to another in periods of conceptual revolution. It is by identifying this role for philosophical investigation that Friedman avoids Quinean holism in which both principles commonly construed as *a priori* and philosophy *tout court* are absorbed into empirical science.

Let us articulate further what is meant by the constitutive *a priori*. As Friedman notes, there is something rather strange about calling any principles *a priori* when they are susceptible to revision. Indeed, the full, Kantian notion of the *a priori* means (i) necessary and unrevisable (in light of epistemic independence from experience) and (ii) constitutive of the concept of the object of knowledge. However, Friedman argues (following the logical positivists) that these two aspects of the Kantian *a priori* can be distinguished and that particular principles of a scientific theory may be *a priori* in the sense of (ii) alone. That is, particular principles of a scientific theory ensure the meaningfulness of empirical laws by making possible their confirmation or disconfirmation by data.

Friedman identifies two components of the constitutive *a priori* part of a physical theory: (i) strictly mathematical principles and (ii) fundamental "coordinative" principles. The latter are "coordinative" in the sense that they mediate between abstract mathematical entities and the concrete empirical phenomena these mathematical entities are intended to represent.[107] This need for coordination between mathematics and empirical phenomena became especially pressing with the advent of general relativity as this theory employs mathematical concepts that bear no obvious relation to human sense perception ([Friedman, 2007]).

[105]Like Einstein's Principle of Equivalence. I will discuss the difference between purely mathematical components of the constitutive a priori and "coordinative principles" (like the Principle of Equivalence) below.

[106]A more contemporary example would be the recent imaging of supermassive black holes at the center of galaxy M87 and at the center of the Milky Way. These discoveries have provided further empirical support for Einstein's theory of general relativity because the respective masses of the black holes can be computed to a high degree of accuracy. This then dictates how "large" they should be, viz., how much they curve spacetime. Thus far, the measurement of the images matches the masses of the black holes, thus confirming Einstein's theory. However, this is only possible once one has laid down the framework of pseudo-Riemannian manifolds.

[107]The need for such principles was already noted by Reichenbach in [Reichenbach, 1920] and [Reichenbach, 1957].

Much of Friedman's discussion[108] focuses on these coordinative principles because he (correctly) finds this problem of coordination a philosophically interesting feature of contemporary physics. On the other hand, he seems to find the strictly mathematical part of the spacetime theories under consideration rather less interesting. However, as we have seen above, the choice of different topological properties of space can have profound effects upon our scientific theory, indicating that the mathematical part of the constitutive *a priori* is perhaps more interesting than Friedman allows.

Before investigating particular mathematical features of the constitutive *a priori*, let us discuss Friedman's conception of general relativity in more detail. We begin with the following strata of a physical (spacetime) theory:

1. Philosophical Meta-paradigms.
2. Constitutive *A Priori* Principles:
   (a) Mathematical Principles;
   (b) Coordinative Principles.
3. Empirical Principles.

We can then think of general relativity as consisting of three components: (2a) the theory of pseudo-Riemannian manifolds of variable curvature; (2b) Einstein's Equivalence Principle;[109] (3) exact solutions to Einstein's Equation. This is because, if we accept general relativity as the best extant scientific description of gravity and spacetime, then we accept that spacetime is (or is well-modeled by) a smooth manifold $\mathcal{M}$ satisfying further constraints.[110] And so we should think of this general topological structure as a "constitutive" component of general relativity insofar as it makes possible the extraction of particular empirical laws like those of local metrical structure.[111] Note, however, that the strictly mathematical part of a physical theory is a merely necessary condition on the extraction of empirical laws: sufficiency is achieved only after coordinative principles have been set up. What the mathematical part alone does is determine the space of *logical possibilities*:

> Einstein's field equations are thus logically possible as soon as we have Riemannian manifolds available within pure mathematics, but they are only [possible as an actual description of some empirical phenomena] when these abstract mathematical structures have been successfully coordinated with some or another empirical reality ([Friedman, 2001], 84).

This is all well and good, and I am in complete agreement with this conclusion; indeed, Friedman's conception of the relativized *a priori* provides us with an extremely helpful

[108] See Part Two, Lectures 1 and 2 of [Friedman, 2001] as well as [Friedman, 2009].

[109] This is the most fundamental example of a coordinative principle in general relativity. There are many other such, e.g., the Light Principle and whatever assumptions are required to produced exact solutions to Einstein Equation.

[110] See the definitions in Section 3.2.

[111] Friedman also suggests that the differentiable structure and the pseudo-Riemannian *form* of the metric are also constitutive. See fn. 4, [Friedman, 1999], p. 82.

way of analyzing the epistemic status of connectivity properties. However, as I will argue below, Friedman's "three levels" cannot fully make sense of our situation. It will turn out that the connectivity properties of space, though intelligibly constitutive, are rather more conventional and contingent than other components of the constitutive *a priori* structure of a spacetime theory.

3.10.2 Articulating the Problem

The crux of the problem is this. We have seen that connectivity properties can be construed as conventional in the sense that empirical evidence need not compel us to choose one or the other. However, they are also constitutive in precisely Friedman's sense because they are included among the mathematical principles of general relativity. In particular, these properties make possible the application of fundamental physical concepts as well as the determination of subsidiary physical laws. On the other hand, there is no unique choice for connectivity: each property (simply or multiply connected) seems to be a licit option internal to the mathematics of general relativity, rendering them unlike, say, the manifold structure *simpliciter* without which one could not even *state* general relativity as such.[112] Furthermore, as my discussion of topological underdetermination has shown, observational data will not help us to decide. This places the connectivity properties in a strange "intermediate" position in Friedman's hierarchy. This is because Friedman seems to think that our choice of constitutive *a priori* components of a scientific theory is guided by "philosophical meta-paradigms," and, furthermore, these paradigms, once one has carefully considered "integrated intellectual history," issue in a *practically unique* such choice.[113] This is not, however, what happens with connectivity properties.

Once more, a comparison with the evolution of the epistemic status of spatial geometry will help us to see why the case of spatial topology is interesting. Consider first Euclidean geometry. For Kant, the axioms of Euclidean geometry were *a priori* constraints on any scientific theory of space. This was similarly the case for Poincaré, though he had to muster new arguments for this fact given his knowledge of non-Euclidean geometries. However, once we entered the era of general relativity, it became clear that geometry could not be *a priori* even in the weaker "relativized" sense. Rather, considerations of metrical structure were "demoted" to merely empirical principles. This is because of Einstein's postulation of the (weak) "Equivalence Principle."[114] Once Einstein had developed special relativity, he realized that this theory was incompatible with classical Newtonian gravitation, since Newtonian gravitation postulates instantaneous "action at a distance" and thus absolute simultaneity ([Friedman, 2001], pp. 83-92). Einstein therefore sought a new theory of gravitation compatible with relativistic spacetime. He did this by appealing to a well-known *empirical* fact: that gravitational and inertial mass are equal, so all bodies fall with the same acceleration in the gravitational field. In short, he appealed to the strangely "universal" nature of gravity.

[112] Connectivity properties also seem different from, say, the assumption that the spacetime manifold is without boundary, for a manifold with boundary would introduce regions in which Einstein's Field Equations does not apply. I discuss this in other work.

[113] I explicate this claim below in Section 3.10.3. See also [Friedman, 2009].

[114] Einstein himself did not carefully distinguish a strong from a weak Principle of Equivalence, resulting in much confusion. See [Ryckman, 2007].

As Friedman notes, Einstein then "elevated" this mere empirical fact to a postulate that gravitation and inertia are the same phenomenon, i.e., the so-called Equivalence Principle. This principle then served as one of the constitutively *a priori* principles of general relativity (a coordinative one). Finally, on the basis of this Equivalence Principle, Einstein took the spacetime metric, registering the curvature of spacetime, as the mathematical representation of the gravitational field. Einstein's Field Equations relate this curvature to the presence of matter, and thus spacetime *geometry* became an empirical matter.

Hopefully, it is clear that both the development of general relativity from special relativity and the preceding development of special relativity from pre-relativistic Newtonian physics constitute rather deep conceptual revolutions in scientific practice. As such, according to Friedman, such transitions required guidance from philosophical meta-paradigms. For instance, he claims that Einstein was able to communicate with practitioners of classical physics by situating his constitutively *a priori* principles in a long historical dialectic concerning absolute versus relative motion ([Friedman, 2001], pp. 105-117). This historical dialectic guided scientific practice from one constitutive framework to another, and, in particular, guided the evolution of spacetime geometry from *a priori* to empirical.

In selecting spatial topology we cannot, however, appeal to the guidance of meta-paradigms because such a choice must be made *internal to our current physical theory*, viz., general relativity with its attendant postulates. This is because we would need to select particular topological properties *after* we have already set up much of our mathematical and coordinative framework for general relativity. Furthermore, in light of our topological underdetermination, we cannot make our choice on the basis of empirical data. Thus, the connectivity properties must be selected on the basis of epistemic (e.g., simplicity, explanatory power, coherence with other areas of physics) or even pragmatic considerations, and yet are also constitutively *a priori*. Consequently, the epistemic status of these fundamental features of space is not adequately captured by the categories on offer. In terms of Friedman's hierarchy, we have the following:

1. Philosophical meta-paradigms/integrated intellectual history (e.g., dialectic concerning absolute versus relative motion from 17th-20th century).
2. Constitutive *A Priori*:
   (a) Mathematical Principles (e.g., theory of pseudo-Riemannian manifolds):
      i. Spatial Topology (How is this to be determined?).
   (b) Coordinative Principles (e.g., Equivalence Principle).
3. Empirical Principles (e.g., exact solutions of Einstein's Field Equations dictating spacetime geometry in general relativity).

The connectivity properties, as mathematical properties of the spacetime manifold, are part of the constitutive *a priori* level. However, which property shall we choose: simply or multiply connected? As we shall see below, the historico-philosophical record cannot decide in any way, as Friedman argues it does for other constitutively *a priori* principles. But neither can empirical findings internal to general relativity decide. It would seem, then, that only epistemic or pragmatic considerations can guide such a choice, and this choice must be made internal to the constitutive *a priori* level of Friedman's hierarchy.

3.10.3 The Contingency of Spatial Topology

Here I will show why particular mathematical properties of the constitutive *a priori* "level" cannot be comfortably integrated into Friedman's account. In particular, though connectivity properties are part of the constitutive *a priori* apparatus of general relativity, these properties are also interestingly different from other examples of constitutive *a priori* principles. Let us first consider how they differ from coordinative principles. (Beyond the obvious fact that connectivity properties do not "coordinate" mathematical *abstracta* with empirical phenomena.) Since Friedman claims that these general mathematical structures and coordinative principles jointly exhaust the constitutively *a priori* level of general relativity, my findings then indicate that particular mathematical properties occupy an odd position that does not quite fit into Friedman's schema of a spacetime theory or of scientific knowledge more generally.

Consider [Friedman, 2009] where Friedman acknowledges a new problem for the constitutive *a priori*. The problem is this: once we give up the ambition of delineating the *a priori* structure of *all* possible scientific theories up front (as Kant sought to do), it would seem that any argument for the constitutively *a priori* status of particular principles will depend entirely upon "the concrete details of the historical process in question" ([Friedman, 2009], 254). Consequently, one of Friedman's primary aims in this paper is to show that both Kant's project and his own contemporary extension of it do not fall prey to a radical contingency. The argument is quite complicated, but it is worth spelling out as my inclusion of connectivity properties among the constitutive *a priori* principles affects it.

Our first concern is: how did Kant seek to explain that the axioms of Euclidean geometry and laws of Newtonian mechanics were (fully) *a priori*, i.e. both necessary and constitutive? This was done by appealing to the structure of our cognitive faculties of sensibility and understanding. Immediately a problem arises,

> [...] How can such proposed transcendental explanations inherit the (assumed) *a priori* necessity of the sciences whose possibility they purport to explain unless we can also somehow establish that they are the *unique* such explanations? (*ibid.*, 254).

Friedman's argument then reduces to defending Kant's transcendental method as "practically unique," given the philosophical and scientific resources (especially those of Leibniz and Newton) at his disposal. Finally, he proposes that this particular conception of the transcendental method can be extended to post-Kantian developments in the dialectic between philosophy, mathematics, and science:

> That each of these successive new intellectual situations has its own "inner logic" implies that the enterprise does not collapse into total contingency [...] therefore *integrated intellectual history* of both the exact sciences and scientific philosophy takes over the role of Kant's original transcendental faculty psychology (*ibid.*, 256).

Friedman seems to suggest that aspects of evolving scientific theories that should be considered properly "transcendental" are those that arise as "practically unique" solutions to problems generated by specific intellectual circumstances and these circumstances themselves

have evolved against the background of Kant's original work. The constitutively *a priori* principles of a given scientific theory are those that arise in such a fashion, and thus these principles are not merely *a posteriori* matters of fact.

Returning to the case of spatial topology, it is very unclear why a particular choice of connectivity property, though constitutively *a priori* in some sense, should present itself as a "practically unique solution" to a scientific problem. According to Friedman's proposal, in order for an aspect of a scientific theory to count as constitutively *a priori* it must: (i) not be subject to empirical confirmation or disconfirmation; (ii) make possible the application of subsidiary physical laws; (iii) be a "practically unique" choice given the intellectual context; (iv) come to light against an intelligibly Kantian background of intellectual development. Given the current state of affairs in classical general relativity and observational cosmology, connectivity properties satisfy (i). These properties also make possible the application of empirical laws in virtue of being properties of the topological manifold, thus satisfying (ii). However, what are we to say of (iii) and (iv)? If we cannot show that the properties in question satisfy these final two conditions, then Friedman's account does appear to fall prey to an historical contingency and, independent of his project, we are left with fundamental properties of space that require a new category to classify them.

Thus, in order to avoid this situation, is there an argument that shows the choice of a particular connectivity property is "practically unique"? Let us turn to Friedman's discussion of how coordinative principles are selected to see whether this reasoning can be applied to the choice of connectivity properties.

This is done by "elevating" mere "empirical laws" to the status of constitutive conditions, which in turn make possible the construction of a rigorous mathematico-scientific theory. For instance, Friedman claims that Poincaré began with the supposedly empirical fact that Euclidean geometry (approximately) governs our perceptual experience of bodily displacements. In virtue of this fact, along with the purported "simplicity" of Euclidean geometry, Poincaré takes Euclidean geometry as an *a priori* constitutive condition on "a precise mathematical framework within which alone our properly physical theories can be subsequently formulated" (*ibid.*, 262). Similarly, Einstein began with the empirical fact that the gravitational and inertial mass of bodies are equal such that all bodies fall with the same acceleration in the gravitational field. Einstein elevated this empirical fact to a constitutive condition of general relativity, the (weak) Equivalence Principle, which gave rise to the entirely new inertial-kinematical structure of general relativity.[115] Friedman does not make clear how it is that more general mathematical features of the constitutive *a priori* are selected, though presumably the process would look similar: we search internal to pure mathematics for an abstract structure that can represent and systematize new and surprising experimental evidence. Is this, however, what we have seen in the case of connectivity properties? Do these instantiate the process Friedman envisages for how constitutively *a priori* principles are selected?

In particular, in the case of topological properties of space, do we find empirical data that are then elevated to constitutive principles of general relativity? Not at all. There is no compelling *empirical* evidence internal to classical general relativity to prefer either a simply connected or a multiply connected topology. (Developments in theories of quantum gravity

[115] One requires also the Light Principle.

could very well change this situation, but such theories are not yet operative.) Rather, the reasons one might provide for preferring, say, a multiply connected topology are either of an epistemic character, viz., the simplification and explanatory power we gain in our physical theories, or consist of a complicated mixture of physical and metaphysical reasons, viz., the plausibility of Machian considerations. As we have seen, the common thread present in both lines of reasoning is the complications that arise with the postulation of an infinite universe. Thus, the process by which we "elevate" a connectivity property to a constitutively *a priori* condition of general relativity is quite different from the one proposed by Friedman. Consequently, we obtain a constitutively *a priori* principle of general relativity whose selection does not seem possible to anticipate using the "inner logic" described by Friedman.

Perhaps one could try to understand my (briefly sketched) arguments for multiply connected topologies in a way more salutary for Friedman. One might first try to find an empirical principle concerning either simply or multiply connected topology which could then be elevated to the status of a constitutive condition. Finding such a principle is, however, quite difficult precisely because these topological properties are global and thus do not obviously relate to our perceptual experience (or even empirical knowledge, broadly construed). This makes them quite different from, say, assessing local metrical structure. Thus, it does not seem that a claim analogous to "Euclidean geometry approximately governs our perceptual experience of bodily displacement" is available for spatial topology. Without this, we are forced to turn to Poincaré's more pragmatic considerations for preferring Euclidean geometry, transposed to the topological case. But, of course, Friedman does not emphasize these pragmatic considerations in his discussion.

A more nuanced attempt to find an empirical principle related to our perceptual experience might concern our perception (or lack thereof) of infinite quantities. For instance, since we only ever perceive finite regions of space,[116] we might elevate this empirical principle to the status of a constitutive condition on our scientific theorizing, resulting in our construction of finite universe models only.[117] And, if we wish to remain consistent with out best current data about the curvature of space,[118] finite universe models can only be produced by adopting a multiply connected topology. The crucial idea here would be that, in order to produce genuine scientific *knowledge*, we would have to constrain our physical theories so as to accommodate basic facts about our cognitive capacities. I am rather skeptical of such a proposal, however. This is because our perceptual capacities are so severely limited that elevating what can be perceived to the level of a constitutive condition is tantamount to undercutting most of contemporary physics. Indeed, it is certain that there is an $n \in \mathbb{N}$ such that even the most perceptually acute human being could not perceive anything outside a region of radius n nor inside a region of radius $1/n$. To say that this empirical fact should be elevated to a constitutive constraint on our spacetime theories would be absurd. Thus, I

[116] Setting aside thorny issues of infinite divisibility.

[117] There is surely an interesting connection to be made here with Kant's arguments in the Metaphysical Exposition of Space in the *Critique of Pure Reason*. In particular, Kant argues that space is given to us as "infinite" and "boundless" and no conceptual representation can account for this. Thus, our representation of space is intuitive in character. However, one might think that here Kant oversteps his own prescribed boundaries of "possible experience," returning us to the suggestion that we only ever experience finite spatial regions. See Essay 4 in [Parsons, 1983].

[118] These data indicate that space is approximately flat, i.e., its curvature $k \approx 0$. See [Spergel et al., 2007].

doubt that Friedman's "inner logic" that extends the Kantian transcendental project can be intelligibly applied to the topological properties under consideration. And thus, though they are reasonably constitutive of general relativity, they are also rather more contingent and *conventional* than Friedman would like. Therefore, these fundamental, global properties of space seem to escape adequate classification by an otherwise attractive theory of scientific knowledge.

3.11 Conclusion

In this chapter, I have argued first, that even assuming the Cosmological Principle, the topology of space is underdetermined by observational evidence. Indeed, even if we had strong evidence for a particular spatial topology, it would still be a live option that the underdetermination persists (recall, however, the concerns about Manchak's theorem). Nonetheless, I believe that we have good reasons to prefer multiply connected topologies. In particular, I argued that we should prefer MCMs on grounds of simplicity, Machian considerations, and explanatory power, where many of these grounds follow from troubling consequences of a spatially infinite universe. Thus, we have good reasons to think that a robust underdetermination is avoidable.

Second, I have argued that our understanding of spatial topology requires a fusion of various aspects of conventionalism and the relativized *a priori*. The connectivity properties in question are conventional in the sense that no empirical evidence can compel us to choose one over the other. However, following the work of Poincaré and Friedman, I have also argued that these connectivity properties are not "mere" conventions insofar as they play an intelligibly constitutive role in our cosmological theories. They are, however, neither coordinative principles nor do they function in the same way as more general mathematical features of relativistic cosmology.

Even after all this work has been done, we must not think that these conclusions are unrevisable: the development of a workable theory of quantum gravity would, once again, change the rules of the game. This would, in fact, cohere very nicely with my analysis. Surely, the transition from classical general relativity and cosmology to a theory of quantum gravity would constitute a paradigm shift requiring the guidance of Friedman's highest level of "philosophical meta-paradigms." That is, such a meta-paradigm would effect the transition from one set of constitutive *a priori* principles (those of classical general relativity) to another, as of yet unknown, set. And, under the auspices of these new principles, spatial topology could be demoted to the level of empirical principles, just as occurred with spatial geometry in the transition from pre-relativistic to relativistic theories of spacetime.[119] It is also possible that our questions about spatial topology will be rendered moot: some theories of quantum gravity postulate that the topology of spacetime fluctuates at a quantum level (the so-called "space-time foam"), and it is not at all apparent what it would even mean to ask about the topology of space in this context. My hope is that such further scientific developments will help to shed light on and to enrich the philosophical issues here discussed.

[119]The results of work by [Almheiri et al., 2020] suggest this possibility in the context of string theory and AdS-CFT (anti de Sitter-conformal field theory) duality.

4 Intuitions of the Infinite and Probability

4.1 Introduction

As discussed in the introduction to the dissertation, the notion of the infinite, even in strictly mathematical contexts, can be understood in many ways. This chapter considers various techniques for "measuring" the infinite and shows how these techniques might be used to shed light on a family of difficult and misunderstood philosophical puzzles.[1]

In contemporary mathematics, there are at least three intuitive criteria (with attendant formalizations) for measuring infinite collections/sets. Let us call the first criterion **Cantor's Principle (CP)**: two infinite sets A and B have the same "size" if and only if their elements can be put into 1-1 correspondence. A rich mathematical theory flows from formalizing this intuition via Cantorian cardinalities; however, it is well known that the use of cardinalities produces counterintuitive results. For instance, the set of even numbers has the same cardinality as the set of all natural numbers, despite the fact that the evens are a proper subset of the naturals. This suggests an alternative criterion, **Part-Whole (PW)**, for determining the size of a set: if A is a proper subset of B, then the size of A should be strictly less than the size of B. Famously, Bernard Bolzano argued that **PW** should serve as the criterion for measuring infinite collections; however, he was not able to develop the mathematics needed to formalize this intuition.[2] Indeed, **PW** has been adequately formalized only very recently by Benci and Di Nasso's *theory of numerosities*.[3] Importantly, **PW** and **CP** yield the same verdicts on finite collections but are incompatible when applied to infinite collections.

The third and final criterion for measuring infinite sets involves a **Frequency (FR)** intuition: if infinite sets A and B occur "equally often" in an ambient set C, then A and B have the same size. **FR** finds expression in the number-theoretic notion of *density*.[4] For example, the even and odd numbers have the same (natural) density because they occur equally often in the natural numbers. This intuition is somewhat under-explored, primarily because it does not offer a way to generalize counting from the finite to the infinite case.[5]

In the past decade or so, a significant philosophical literature has sprung up around the

[1] In this chapter, "measuring" should be taken to mean "measuring the 'size' of an infinite set/collection" no matter how "size" is construed in the various approaches here considered.

[2] See [Bolzano, 1972] and [Bolzano, 1975]. Note also that Galileo and Leibniz were drawn by the part-whole intuition, though, ultimately, they concluded that one should not attempt to construct a theory of measuring the infinite.

[3] See, for instance, [Benci and Nasso, 2003b], [Benci et al., 2006], [Benci and Nasso, 2019]. See also [Katz, 1981] for an early precursor.

[4] See, e.g., [Tenenbaum, 1995]. Note that there are various types of densities employed in number theory.

[5] I wish to pursue questions about the conceptual significance of **FR** in future work.

theory of numerosities.[6] The motivation for much of this work is that numerosities allow us to generalize counting from the finite to the infinite case such that **PW**, rather than **CP**, is preserved. Many rich philosophical questions immediately follow, e.g. how does this new theory of counting compare with Cantorian cardinalities in terms of epistemic and mathematical usefulness? Beyond such questions about what constitutes a "good" theory of infinite counting, it is plausible that, given the fundamental nature of these mathematical developments, there should be wide-ranging consequences in areas of study employing infinitary considerations. Here I consider the effect of applying **PW** and numerosities to infinitary probability theory. There has already been some very nice mathematical and philosophical work done here, but my aims are somewhat different.[7]

In particular, after examining three central paradoxes of infinitary probability theory, the Label Invariance Paradox, God's Lottery, and Bertrand's Paradox, I came to the conclusion that none of these seems to be generated by anything involving probability. Rather, these paradoxes involve a conflict between our techniques for measuring infinite sets and the information we wish to preserve when doing so. More precisely, they are generated by the conflict between the coarseness of **CP**, our intuitions about the "relative sizes" of infinite sets,[8] and how these size relations ought to behave in different contexts. I thereby develop a unified framework in which to think about these seemingly distinct "probabilistic" paradoxes and suggest how these paradoxes might be resolved. The recognition that there are different—and inconsistent—techniques for measuring infinite sets will help to disabuse us of the intuitions underlying the paradoxes and will also help to make some of our claims more precise.

Finally, a more general theme that emerges from this investigation is that there is an inextricable indeterminacy to our theories of infinite counting.[9] However, though one might find this undesirable, I believe it provides a flexibility that allows us to select properties suited to particular contexts (e.g., probability theory, set theory, number theory, etc.). This will lead us back to questions about what a "good" theory of infinite counting should look like and whether there is any such.

4.2 The Theory of Numerosities

It is undeniable that Cantorian set theory provides us with much insight into the nature of mathematical infinity. Indeed, Cantor's theory of cardinal numbers offered the first systematic generalization of arithmetic from the finite to the infinite case. As we have seen, this is effected by accepting a particular intuitive criterion for "measuring" an infinite collection, **CP**. Crucial for our purposes here is the fact that 1-1 correspondence, viz., the "equipotency" equivalence relation, used to generate cardinality assignments is extremely coarse.

It is in large part this coarseness that produces counterintuitive verdicts, e.g., that all infinite sets of integers have the same "size." In this section, I wish to make clear how

[6] See especially [Mancosu, 2009], [Parker, 2013], [Mancosu, 2017] and the references provided below concerning applications of numerosity theory to probability.

[7] See [Benci et al., 2013] and [Benci et al., 2018]. These papers will be discussed below. One should also consult [Mancosu and Massas, 2023].

[8] By this I mean judgements like, "The even numbers comprise half of the natural numbers."

[9] I use "infinite counting" as synonymous with "measuring" for infinite sets.

the inclusion of finer-grained mathematical structure in the theory of numerosities allows us to validate the **PW** intuition rather than **CP**. I will primarily follow the most recent development of numerosity theory presented in [Benci and Nasso, 2019]. There is much here of both technical and philosophical interest, but the main point for our purposes is the following. In contrast to Cantorian cardinalities, the numerosity of a given set A can depend quite sensitively upon two things: (i) how the elements of A are "labelled"; (ii) how one constructs a model in which numerosity assignments are made.[10] I claim no originality in the presentation of the following technical results. Rather, I wish to show how results from [Benci and Nasso, 2019] might be employed to help to resolve our philosophical concerns.

4.2.1 Counting Systems

Following [Benci and Nasso, 2019],[11] let a *counting system*[12] be a triple $(\mathfrak{U}, \mathfrak{N}, \mathfrak{n})$ where $\mathfrak{U}$ is a family of sets to be counted, $\mathfrak{N}$ is a linearly ordered set of numbers, and $\mathfrak{n} : \mathfrak{U} \twoheadrightarrow \mathfrak{N}$ is a surjective function assigning a "size" to each set in $\mathfrak{U}$. Intuitively, letting A and B be infinite sets, we would like any such counting system to satisfy the following:

1. **Cantor's Principle (CP)**: $\mathfrak{n}(A) = \mathfrak{n}(B)$ iff there is a bijection between A and B.
2. **Part-Whole (PW)**: If $A \subset B$, then $\mathfrak{n}(A) < \mathfrak{n}(B)$.
3. The usual algebraic properties of the natural numbers, e.g., commutativity of sum and product operations.[13]

Importantly, these intuitive properties are gotten by considering our experience of counting finite entities. If we restrict ourselves to only finite sets we obtain:

Example 4.2.1 (Finite Counting System)**.** Let $\mathfrak{U}_{\text{fin}}$ be the class of finite sets, $\mathfrak{N} = \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and $|\cdot|_{\text{fin}}$ the finite cardinality function. Then our counting system $(\mathfrak{U}_{\text{fin}}, \mathbb{N}_0, |\cdot|_{\text{fin}})$ satisfies (1)-(3) above.

Unfortunately, there is no *prima facie* reason for thinking that all these properties will generalize to the infinite case. And indeed, as is well known, such a generalization fails. We cannot have a counting system containing infinite elements that satisfies (1) and (2) together. Cantorian cardinalities provide an example of a counting system at infinity satisfying (1):

Example 4.2.2 (Cantorian Cardinalities)**.** $(\mathcal{V}, \mathsf{Card}, |\cdot|)$ is the counting system of Cantorian cardinals where $\mathcal{V}$ is the universal class of all sets, Card is the class of cardinal numbers, and $|\cdot|$ assigns to a given set its equipotent cardinal.

[10] The theory of numerosities can be presented without appeal to labelled sets. In particular, one can begin with the primitive notion of an "equisize" equivalence relation (instead of "equipotency") and then show that the theory of counting based upon this relation is equivalent to the theory of numerosities in terms of labelled sets. This will not affect the main claims of the chapter.

[11] Note that, in the last two decades, various developments of the theory have been given. See [Benci and Nasso, 2003b] and [Benci and Baglini, 2021].

[12] See Appendix E for the formal properties of a counting system.

[13] See [Benci and Nasso, 2003b], p. 52 for further discussion of this condition.

It is possible to devise another counting system with infinite elements that satisfies (part of) (1) and rejects (2). This is given by:

Example 4.2.3 (Cantorian Ordinals)**.** $(\mathcal{WO}, \mathsf{Ord}, \mathrm{ot})$ is the counting system of Cantorian ordinals where $\mathcal{WO}$ is the class of well-ordered sets, Ord is the class of ordinals, and $\mathrm{ot}(\cdot)$ is the order-type function.

Cantorian ordinals represent a more complex way of counting than that of Cantorian cardinals. This is because we must keep track of both the elements in some $A \in \mathcal{WO}$ *and* their order. Note that this way of counting still rejects criterion (2), since, for example, $\mathrm{ot}(\mathbb{N}) = \omega$ and $\mathrm{ot}(\mathbb{N}_0) = 1 + \omega = \omega$, even though $\mathbb{N} \subset \mathbb{N}_0$.[14] However, Cantorian ordinals vindicate only part of (1). This is because, in order for two sets to have the same order type, there must be a bijection between the sets that, additionally, preserves order. Thus, if $\mathrm{ot}(A) = \mathrm{ot}(B)$ for $A, B \in \mathcal{WO}$, then A and B stand in 1-1 correspondence. However, the converse fails: there are many pairs A, B that are 1-1 correlated but have different order-types. Finally, it is worth noting that (3) is satisfied by neither cardinals nor ordinals. Indeed, ordinal addition and multiplication even fails to be commutative, e.g., $1 + \omega = \omega < \omega + 1$.

Benci and Di Nasso's aim was to develop a mathematically adequate counting system, the theory of numerosities, that rejects (1) instead of (2) and also satisfies (3). Let us see how to do this for the case of countable infinities.[15] In what follows, I will consider only countably infinite sets, unless otherwise noted.

Type 1 Sensitivity

We wish to construct a *numerosity counting system* $(\mathfrak{U}, \mathfrak{N}, \mathfrak{n})$, viz., a counting system that satisfies **PW** for infinite sets rather than **CP**. In order to obtain **PW** for this system, the presentation of numerosities with which we are concerned arranges elements of sets A and B into smaller sets to be counted. This is done by assigning each element of A (resp. B) a "label" via a labelling function ℓ_A (resp. ℓ_B). We are not, in contrast to Cantorian cardinalities, simply considering "bare" elements of sets. We are, instead, like Cantorian ordinals, preserving additional information. However, the process by means of which we do this is even more complicated.

We begin with a countably infinite set A,[16] partition it via a labelling function into countably many finite subsets, and then analyze a sequence of approximations. More precisely: consider a pair $\mathbf{A} := (A, \ell_A)$, called a *labelled set*, with $\ell_A : A \to \mathbb{N}_0$ such that, for any $n \in \mathbb{N}_0$, $\ell_A(a) = n$ for finitely many $a \in A$. Thus, A can be written as the union of the non-decreasing sequence of finite sets

$$A_0 \subseteq A_1 \subseteq \cdots \subseteq A_n \subseteq A_{n+1} \subseteq \cdots \tag{4.2.1}$$

where $A_n = \{a : \ell_A(a) \leq n\}$. The finite cardinality $|A_n|$ can then be thought of as the nth approximation to the numerosity of $\mathbf{A}$.

[14] Note that adding an element "to the right" of $\mathbb{N}$ does result in a larger order-type.

[15] See [Benci et al., 2006], [Benci et al., 2007], [Nasso and Forti, 2010], [Blass et al., 2012] for generalizations. As of yet, some central aspects of numerosities for uncountable sets remain unsettled.

[16] Again, see remark in footnote 15.

In virtue of this additional structure, we require a new notion of "sameness" for labelled sets. That is, under what conditions are two labelled sets $\mathbf{A}$ and $\mathbf{B}$ the same for the purposes of counting? And how precisely does this sameness relate to the assignment of numerosities? Consider first the notion of isomorphism between labelled sets:[17]

Definition 4.2.4. Two labelled sets $\mathbf{A} = (A, \ell_A)$ and $\mathbf{B} = (B, \ell_B)$ are isomorphic, written as $\mathbf{A} \cong \mathbf{B}$, if there exists a bijection $\varphi : A \to B$ that preserves the labellings of $\mathbf{A}$ and $\mathbf{B}$, i.e., $\ell_B \circ \varphi = \ell_A$.

Clearly, now, we wish to say that two labelled sets that are "the same," i.e., isomorphic in the above sense, will have the same numerosity. Indeed, $\mathbf{A} \cong \mathbf{B} \Rightarrow \mathfrak{n}(\mathbf{A}) = \mathfrak{n}(\mathbf{B})$. However, for our purposes below, it is more instructive to see how assignments of numerosities can *differ*. There are two ways in which labelled sets can fail to be isomorphic. First, if there is not a bijection $\varphi : A \to B$, then clearly the labelled sets $\mathbf{A}$ and $\mathbf{B}$ will not be isomorphic. This condition encodes the coarsest counting condition, viz., 1-1 correspondence. However, we now take a step further and say that $\mathbf{A} \ncong \mathbf{B}$ if our labellings are not preserved, i.e., $\ell_B \circ \varphi \neq \ell_A$ (with φ now assumed to be a bijection). Either failure of isomorphism will result in a change in the natural "counting function" that serves to provide the finite approximations of the numerosity of a labelled set (see Proposition 4.2.6 below):

Definition 4.2.5. The counting function $\gamma_{\mathbf{A}} : \mathbb{N}_0 \to \mathbb{N}_0$ of labelled set $\mathbf{A} = (A, \ell_A)$ is given by

$$\gamma_{\mathbf{A}}(n) = |A_n| = |\{a \in A : \ell_A(a) \leq n\}| . \tag{4.2.2}$$

As we have already seen, A is given by the increasing union $\bigcup_{n \geq 0} A_n$, and so A_n is the nth approximation of the size of $\mathbf{A}$.

There is then the following relationship between isomorphic labelled sets and their counting functions:

Proposition 4.2.6. *Let $\mathbf{A} = (A, \ell_A)$ and $\mathbf{B} = (B, \ell_B)$ be labelled sets. Then the counting functions $\gamma_{\mathbf{A}}$ and $\gamma_{\mathbf{B}}$ are precisely identical iff $\mathbf{A} \cong \mathbf{B}$.*

Proof. See Appendix E. □

With these preliminary definitions and results in hand, [Benci and Nasso, 2019] show that their "Alpha-Theory," an alternative development of nonstandard analysis, provides a natural definition of the numerosity of a labelled set:[18]

Definition 4.2.7. The numerosity of a labelled set $\mathbf{A}$ is given by

$$\mathfrak{n}(\mathbf{A}) = \lim_{n \uparrow \alpha} \gamma_{\mathbf{A}}(n), \tag{4.2.3}$$

[17] See [Benci and Nasso, 2003b], [Benci and Nasso, 2019].

[18] In [Benci and Nasso, 2019], the numerosity of a labelled set is written as $\mathfrak{n}_\alpha(\cdot)$ to make explicit the relationship to the alpha-limit. I will suppress the α-subscript.

where $\lim_{n\uparrow\alpha}$ is the notion of an "alpha-limit."[19]

Just as the (ϵ, δ)-limit is the foundational notion of classical analysis, so too is the alpha-limit the foundational notion of nonstandard analysis via Alpha-Theory. Loosely, the alpha-limit of a sequence is the value taken by the sequence at an ideal, infinite number, α. Then we have

Definition 4.2.8. Let $\mathfrak{L}$ be the class of (countable) labelled sets, and let $\mathfrak{N} \subseteq \mathbb{N}_0^*$ be the range of the surjective function $\mathfrak{n}$, where $\mathbb{N}_0^*$ is the set of hypernatural numbers. Then $(\mathfrak{L}, \mathfrak{N}, \mathfrak{n})$ is a numerosity counting system, viz., a counting system satisfying **PW**.

In [Benci and Nasso, 2019], the existence of alpha-limits for arbitrary sequences, as well as the existence of α, are given as axioms of the theory.[20] The authors then construct a model for these axioms, thereby guaranteeing their consistency. This process is entirely analogous to the axiomatic introduction of the real numbers: we postulate the existence of the real numbers as a complete ordered field, and then construct a model via either equivalence classes of Cauchy sequences of rational numbers or Dedekind cuts of rational numbers. Finally, we define field operations and an order relation on the sets of the model and verify that the resulting structure satisfies the properties of a complete ordered field. The details of the model construction for the axioms of Alpha-theory is rather more involved and are not especially relevant here. However, we can rest assured that the use of alpha-limits is entirely safe.[21]

In contrast to the classical limit, the alpha-limit has some rather strange properties. Most striking is the fact that alpha-limits always exist. Furthermore,

Proposition 4.2.9. *Let $f, g : \mathbb{N}_0 \to \mathbb{N}_0$. Then, if $f(n) \neq g(n)$ for all but finitely many n,*[22]

$$\lim_{n\uparrow\alpha} f(n) \neq \lim_{n\uparrow\alpha} g(n). \tag{4.2.4}$$

Proof. See Appendix E. □

Thus, we notice the following. By Proposition 4.2.6 we know that a failure of isomorphism between labelled sets $\mathbf{A}$ and $\mathbf{B}$ means that their counting functions $\gamma_{\mathbf{A}}$ and $\gamma_{\mathbf{B}}$ are not identical. Let us assume that these functions differ on all but finitely many n. Then by

[19] See [Benci and Nasso, 2003a] and [Benci and Nasso, 2019]. Theorem 16.14 of [Benci and Nasso, 2019] shows that this way of defining numerosities ensures that $(\mathfrak{L}, \mathfrak{N}, \mathfrak{n})$ with $\mathfrak{N} \subseteq \mathbb{N}_0^*$ (the hypernaturals) satisfies all the desired properties for a numerosity counting system.

[20] See Appendix E.

[21] In [Benci and Nasso, 2019], the authors first introduce axioms for what they call *Alpha-Calculus* (see Appendix E) which deals with alpha-limits of real-valued sequences. A model for the axioms of Alpha-Calculus is then given in Section 2.11 using non-principal maximal ideals (equivalent to non-principal ultrafilters) in the ring of real-valued sequences. (An alternative model construction for Alpha-Calculus is given in Chapter 11.4 via ultrapowers.) In Chapter 4, the authors develop *Alpha-Theory* as an extension of Alpha-Calculus insofar as we take alpha-limits of sequences taking values in any set (not just $\mathbb{R}$). A model for the axioms of Alpha-Theory is then given in Chapter 6.7.

[22] Note that the sequences for which the alpha-limit is defined can take values on any set, not just $\mathbb{N}_0$. I restrict myself to $\mathbb{N}_0$ because we will be interested in alpha-limits of counting functions taking values in $\mathbb{N}_0$.

Proposition 4.2.9 and the definition of numerosity:

$$\lim_{n\uparrow\alpha}\gamma_{\mathbf{A}}(n) \neq \lim_{n\uparrow\alpha}\gamma_{\mathbf{B}}(n) \Longrightarrow \mathfrak{n}(\mathbf{A}) \neq \mathfrak{n}(\mathbf{B}). \tag{4.2.5}$$

We will exploit this chain of relationships when we consider the Label Invariance paradox below.

Type 2 Sensitivity

Finally, we should note that there is a second way that the assignment of numerosities might be considered "sensitive" or perhaps "arbitrary." This has been examined to some extent in the literature,[23] but here I will provide a different (though equivalent) presentation of this sensitivity. My presentation will facilitate comparison with a possible resolution of God's Lottery discussed below.

Benci and Di Nasso construct a model for the axioms of Alpha-Calculus (see Section 2.11 of [Benci and Nasso, 2019]). They also show that models for Alpha-Calculus are highly non-unique in the sense that a model is described by specifying the family of qualified sets (see Appendix E). Furthermore, it is shown that every non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ is the family of qualified sets that describes a model for Alpha-Calculus (Theorem 2.22 and Theorem 11.18 of [Benci and Nasso, 2019], respectively). Recall the following definition:[24]

Definition 4.2.10 (Filter; Ultrafilter)**.** Let I be a nonempty set. Then a nonempty collection $\mathcal{U} \subseteq \mathscr{P}(I)$ is said to be a *filter* over I if $\mathcal{U}$ is closed under supersets and finite intersections. $\mathcal{U}$ is a *proper* filter if $\emptyset \notin \mathcal{U}$. Finally, an *ultrafilter* is a proper filter $\mathcal{U}$ satisfying

1. For every $A \subseteq I$, either $A \in \mathcal{U}$ or $A^c \in \mathcal{U}$ with $A^c = I \setminus A$.

Finally, if no finite subsets belong to $\mathcal{U}$, then $\mathcal{U}$ is said to be *non-principal* or *free*.

Theorem 2.22 of [Benci and Nasso, 2019] proves that a family of qualified sets $\mathcal{Q}$ satisfies the above conditions for a non-principal ultrafilter, and Theorem 11.18 proves the "converse" that, given a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, there is a family of qualified sets $\mathcal{Q}$ such that $\mathcal{Q} = \mathcal{U}$.

Since we are defining our numerosity counting systems via the Alpha-Calculus, it should be clear that numerosities will inherit the sensitivities of the models of Alpha-Calculus. However, this becomes even more complex and raises interesting foundational questions. This is because Benci and Di Nasso suggest that the following principle, the *Cauchy Infinitesimal Principle* (CIP), be added to the axioms of Alpha-Calculus (or the more general Alpha-Theory):

> *Every positive infinitesimal number is the alpha-limit of some decreasing infinitesimal sequence.*

CIP is in fact independent of the axioms of Alpha-Calculus, and it can be shown that the new theory produced by adjoining CIP to the Alpha-Calculus axioms admits models.

[23] See [Mancosu, 2009] and [Parker, 2013]
[24] See [Goldblatt, 1998].

However, the existence of these models requires highly non-constructive mathematics; in particular, to prove their existence one must go beyond the axioms of ZFC. This connection is made evident by the following theorem (Theorem 14.12 of [Benci and Nasso, 2019]):

Theorem 4.2.11. *Alpha-Calculus theory proves that CIP holds iff the non-principal ultrafilter $\mathcal{Q}$ of qualified sets over $\mathbb{N}$ is selective.*[25]

It is known that the existence of a selective ultrafilter is independent of ZFC.[26] Thus, the existence of models of Alpha-Calculus/Theory in which CIP holds is independent of ZFC. But what, precisely, has this to do with numerosities? The connection is provided by the following (Theorem 16.33 of [Benci and Nasso, 2019]):

Theorem 4.2.12. *Assume the axioms of Alpha-Theory. Then the following are equivalent:*

1. *Cauchy Infinitesimal Principle (CIP).*
2. *The numerosity counting system $(\mathfrak{L}, \mathfrak{N}, \mathfrak{n})$ is Zermelian, i.e., it satisfies **PW** and the following property: $\mathfrak{n}(A) \leq \mathfrak{n}(B)$ iff $\mathfrak{n}(A) = \mathfrak{n}(A')$ for some subset $A' \subseteq B$.*
3. *The set of numerosities $\mathfrak{N}$ is precisely the set of hypernaturals $\mathbb{N}_0^*$. That is, for every $\nu \in \mathbb{N}_0^*$ there exists a labelled set $\mathbf{A}$ such that $\mathfrak{n}(\mathbf{A}) = \nu$.*

This complicated chain of dependencies relates the new presentation of numerosity theory within Alpha-Theory to the earlier work in [Benci and Nasso, 2003b]. There numerosities are taken to be equivalence classes of nondecreasing functions $f : \mathbb{N} \to \mathbb{N}$ equivalent modulo a selective ultrafilter. Then, it is shown that a numerosity function $\mathfrak{n}$ from the class of countable labelled sets $\mathfrak{L}$ to $\mathbb{N}_0^*$ exists iff there exists a selective ultrafilter.[27] This is precisely what we have seen above, except via a detour through Alpha-Theory, qualified sets, and CIP. Once more, this will help us to see more clearly how numerosity theory might be applied to the paradoxes below.

In virtue of the dependence of the numerosity function on the selective ultrafilter and the fact that there are many selective ultrafilters, the assignments of numerosities are not unique. As we shall see below, this is closely related to the arbitrariness of properties of α since the properties of α will depend on our family of qualified sets.

This sensitivity of numerosities has been deemed either a disadvantage of the theory ([Parker, 2013]) or, at least, no less of a problem than the underdetermination present in ZFC, e.g., the fact that ZFC cannot resolve basic questions about size assignments like the Continuum Hypothesis ([Mancosu, 2009]). The details of this debate would take us too far afield, but I wish to note that this second sensitivity can be turned into an advantage of the theory. In particular, we might wish to solve a particular problem that requires $\mathfrak{n}(\mathbb{N}) = \alpha$ to have various properties. This can be achieved by selecting appropriate sets to be members of our selective ultrafilter/family of qualified sets. This idea will feature significantly in the resolution of God's Paradox. More generally, the sensitivity of numerosities suggests

[25] There are many conditions for classifying selective ultrafilters. See Proposition 4.1 of [Benci and Nasso, 2003b] and [Benci and Nasso, 2019], p. 291.

[26] See, e.g., [Booth, 1970] and [Kunen, 1976].

[27] See [Benci and Nasso, 2003b], pp. 62-3.

that what constitutes a "good" theory of infinite counting may in fact be highly context-dependent.

In conclusion, we have seen the following. The theory of numerosities validates **PW**, not **CP**, and has well-behaved algebraic properties. It was required that our counting theory preserve more information than either Cantorian cardinalities or ordinals. In particular, the numerosity of a set A depends on both the elements of A and how the elements of A are labelled (under the presentation of numerosity theory with which I have been concerned). Call this **Type 1** sensitivity of numerosities. This will help us to identify what is philosophically important in some famous paradoxes below and will provide independent arguments for their resolution. We have also seen that numerosities exhibit a sensitivity to the underlying selective ultrafilter. Call this **Type 2** sensitivity. This will help us to resolve some of the paradoxes in question.

4.3 The Label Invariance Paradox and God's Lottery

In this section, I begin by examining two paradoxes from infinitary probability theory (Label Invariance and God's Lottery) and provide a single framework that both diagnoses and dissolves the paradoxes. Much of this will turn, of course, on the details of the paradoxes in question. However, my main contention is that these paradoxes are, ultimately, not about probability theory *per se*. Rather, they involve a mismatch between: (i) our intuitions about the relative sizes of infinite subsets of countably infinite sets; (ii) the coarse mathematical framework typically used to analyze these relationships. In particular, the framework of cardinalities (and its underlying reliance upon bijective correspondence alone) obliterates information that must be preserved in order to validate our intuitions about relative sizes of infinite sets. It is this mismatch the produces the paradoxes and not, ultimately, anything about the nature of likelihood. I conclude by examining Timothy Williamson's much discussed argument that infinitesimals cannot save a regularity constraint in infinitary probability. I show that this involves assumptions common to the previous paradoxes and that his argument can be disrupted by, once more, jettisoning the Cantorian framework.[28]

4.3.1 The de Finetti Lottery and Countable Additivity

I will begin with the Label Invariance paradox. The most focused discussion of the paradox is found in [Bartha, 2004], so I follow the basics of the set-up found there.[29] I would like to note at the outset that I agree with much of what Bartha says; however, I think his analysis is not sufficiently general insofar as it is tethered to the formalism of probability theory and does not acknowledge the full significance of infinitary considerations.

[28] Ultimately, I think his argument is highly indeterminate and must be precisified before any evaluation of it can be made.

[29] There is also a brief discussion of the Label Invariance intuition in [Wenmackers and Horsten, 2013]. [Gyenis and Rédei, 2015] discuss similar ideas in a rather different context. I take this up below in my analysis of Bertrand's Paradox. For a discussion of this intuition in the context of inflationary cosmology, see [Norton, 2021], [Parker, 2020], and [Wenmackers, 2023].

Bartha's primary objective in this paper is to dispense with difficulties arising from the *de Finetti Lottery*.[30] I will describe this lottery in some detail as it will be relevant for many sections of the chapter. De Finetti argued that we should be able to make sense of a uniform probability distribution over a countably infinite set. Consider a lottery in which the number of tickets issued is countably infinite with each ticket having an equal (subjective) probability of winning.[31] However, as de Finetti showed, the assumption of equiprobability for each ticket is incompatible with a standard axiom of infinitary probability theory: Countable Additivity (CA) (see Appendix F). Let a_i, a_j be tickets in our lottery with $i, j \in \mathbb{N}$ and let $P(a_i)$, $P(a_j)$ be real numbers representing the probability of drawing a_i or a_j, respectively. Equiprobability is then given by $P(a_i) = P(a_j)$ for all i, j. CA is given by $P(a_1) + P(a_2) + \cdots = 1$, since the probability that some ticket wins is unity, and this should be the infinite sum of probabilities for each individual ticket. Evidently, if we have $P(a_i) = P(a_j) = r$ for $r \in (0, 1]$, then the series in CA diverges, and CA must be abandoned. Similarly, if $P(a_i) = P(a_j) = 0$, then CA also fails. Thus, it would seem, we must abandon either Equiprobability or CA. De Finetti chooses to abandon CA and, ultimately, retreats to a probabilistic finitism by letting all $P(a_i) = 0$ to retain Equiprobability.

Bartha agrees with de Finetti that Equiprobability should be retained but wishes to show that, in many other cases, one can coherently retain CA. That is, the use of CA in other applications of subjective probabilities is entirely unproblematic. He begins by surveying some arguments for dropping Equiprobability in the de Finetti lottery. The first is that any mechanism that might be employed will be biased in some sense, and so the conflict between Equiprobability and CA never gets off the ground.[32] If we think this response cogent, we might still object that the de Finetti lottery deals with *subjective* probabilities (credences) rather than *objective* probabilities (chances), and so worries about a "mechanism" are not apropos.[33] Still, we might reason that our credences should reflect our knowledge of chances,[34] and so we should still want a truly random mechanism. In any case, after pursuing a number of such lines of thought, Bartha concludes (rightly, I think) that such arguments for dropping Equiprobability are not conclusive.

The second, somewhat more worrisome, argument against Equiprobability goes as follows. We have a Dutch Book argument for CA as a constraint on subjective probability as with other axioms of probability theory.[35] More precisely, if we assign positive, real-valued probabilities to a_i for all i, then $\sum_{i=1}^{\infty} P(a_i) = 1$ on pain of being Dutch Booked. Thus,

[30] See [de Finetti, 1974].

[31] When discussing [Bartha, 2004], I restrict myself to subjective probabilities because his paper is concerned with whether Countable Additivity serves as a constraint on subjective interpretations of probability (following [de Finetti, 1974] and [Kelly, 1996]). Furthermore, parts of Bartha's discussion deal with Dutch Book Arguments, which are expressly concerned with whether there are rational constraints on subjective probabilities. See [Hájek, 2009] for discussion. However, my interpretation of the Label Invariance paradox applies equally well to both subjective and objective probabilities. I discuss this further below in the context of God's lottery.

[32] See [Spielman, 1977], [Howson and Urbach, 1993] and the references therein.

[33] See [Williamson, 1999] for extended discussion. See footnote 31 for the restriction to subjective probabilities.

[34] As codified, for example, by Lewis's *Principal Principle*, i.e., that the subjective probability for event E, given that the chance of E is r, should be precisely r. See [Lewis, 1980].

[35] See [Howson and Urbach, 1993] and [Williamson, 1999].

since CA holds, and Equiprobability is incompatible with CA in the de Finetti lottery, reject Equiprobability.

Bartha sees two possible ways for avoiding the conclusion that we must adopt CA and thereby drop Equiprobability. The first is to use non-standard probability measures. This option has garnered much attention in recent years, and I will discuss it at length below.[36] The second is simply to deny that we have a real-valued credence (or, in the context of a Dutch Book, a fair betting quotient) for the proposition that some ticket a_i wins. Namely,

> The crucial point [...] is that we might lack betting quotients for these propositions taken in isolation. If we have no betting quotients for these propositions, and hence no subjective probabilities, then countably additivity is inapplicable rather than violated. This is not the bland observation that if there are no betting quotients, there are no subjective probabilities at all. Rather, we shall see that it is possible to define a type of betting quotient that is faithful to all the standard axioms of the probability calculus except countable additivity ([Bartha, 2004], 307).

Many of the finer details of Bartha's proposal are inessential for our purposes, so I will briefly summarize his construction. In essence, he wishes to show that we can express Equiprobability between two outcomes (and later two sets of outcomes; more on this below) without appealing to the existence of a probability function to which CA applies. He starts by defining a *relative betting quotient* for proposition B relative to A, written $RBQ(B;A)$, as $k \in \mathbb{R}^+$ such that a particular bet[37] is subjectively fair. This is a generalization of a special case in which A and B are assigned well-defined betting quotients $p, q \in [0,1]$. In this case, k is simply q/p, and yields a particular betting outcome. Bartha obtains the same such outcome *without* postulating the existence of p, q. The key point is that, in doing so, we have guaranteed that the relative betting quotient need not have an upper bound. Without the existence of an upper bound, CA need not apply, as the series in question need not converge.

Finally, from $RBQ(B;A)$ alone we can define *relative probabilities*. Let $R(B,A)$ be the probability of B relative to A. Then, we can express Equiprobability as the condition that $R(B,A) = 1$, and, by the above reasoning, we need not fear incompatibility with CA, which will not be applicable in general.[38] Finally, we can express the probability of B relative to the entire outcome space X, written as $R(B,X)$, which in turn induces a probability measure $Pr_R(B)$. With this set-up in hand, we can now describe the Label Invariance paradox.

[36] See [Bartha and Hitchcock, 1999], [Benci et al., 2013], [Wenmackers and Horsten, 2013], [Easwaran, 2014], [Benci et al., 2018], and further references therein.

[37] See [Bartha, 2004], p. 308.

[38] Note that there are cases in which a version of CA for relative betting quotients will apply. For instance, let $RBQ(B_i;A)$ be defined for all i, let the B_i be exclusive, and let $RBQ(B;A)$ be defined with $B = B_1 \vee B_2 \vee \cdots$. Then we will have $\sum_{i=1}^{\infty} RBQ(B_i;A) = RBQ(B;A)$. If we let B_i be the event that "ticket i wins" and A the event that "some ticket wins," then we will encounter the same puzzle in the case when each $RBQ(B_i;A) = 0$. However, Bartha's analysis ensures that none of *these* $RBQ(B_i;A)$ is defined since any relative betting quotient is some *positive* real number k. The conceptual effect of this restriction is to say that "...the proposition 'ticket [i] wins' and 'some ticket wins' are probabilistically incommensurable" ([Bartha, 2004], 310).

4.3.2 The Label Invariance Paradox

This paradox threatens Bartha's solution because it seems to show that assuming Equiprobability over a countably infinite set, even when Equiprobability is defined using relative probabilities, will lead to inconsistency. That is, the Label Invariance Paradox suggests that Equiprobability is indeed the troublesome assumption and not CA. Somewhat surprisingly, we shall see that another intuition could be the culprit.

Begin by labeling the elements of countably infinite set A as $a_1, a_2, a_3, \ldots$. Assume that for any pair $\{a_i, a_j\}$, the (subjective) probability of choosing a_i is equal to the (subjective) probability of choosing a_j, i.e., $P(a_i) = P(a_j)$ for all $i, j \in \mathbb{N}$. Once more, call this assumption *Equiprobability*. Now assume that we have a well-defined relative probability function R. We can then use this to generate an induced probability measure Pr_R for an event relative to the entire outcome space. This allows us to express subjective probabilities for infinite subsets of our countably infinite outcome space. In particular, Bartha claims we should have that

$$Pr_R(\mathsf{Even}) = Pr_R(\mathsf{Odd}) = \frac{1}{2}, \tag{4.3.1}$$

where Even is the event that the selected a_i has an even label and Odd is the event that the selected a_i has an odd label. I take it that Equation 4.3.1 has the status of a reasonable intuition in Bartha's discussion. One might argue for the intuition as follows. Since Even (*mutatis mutandis* Odd) expresses that the selected a_i belongs to a set consisting of *every other* member in our original list, the event should have a relative betting quotient, and hence relative probability, of $1/2$ with respect to the full set of outcomes.[39] Similarly, we should have

$$Pr_R(\mathsf{One}) = Pr_R(\mathsf{Two}) = Pr_R(\mathsf{Three}) = Pr_R(\mathsf{Four}) = \frac{1}{4} \tag{4.3.2}$$

where, for $m = 0, 1, 2, 3, \ldots$,

$$\begin{aligned} \mathsf{One} &:= \{a_i \in A : i = 4m + 1\}, \quad \mathsf{Two} := \{a_i \in A : i = 4m + 2\}, \\ \mathsf{Three} &:= \{a_i \in A : i = 4m + 3\}, \quad \mathsf{Four} := \{a_i \in A : i = 4m + 4\}. \end{aligned}$$

As above, we understand, say, One, to be the event that the selected a_i has a label of the form $i = 4m + 1$. The reasoning underlying Equation 4.3.2 is the same as that for Equation 4.3.1, viz., One expresses that the selected a_i belongs to a set consisting of *every fourth* member in the original list. Thus, the event should have relative betting quotient, and hence relative probability, of $1/4$ with respect to the full set of outcomes.

Now a second crucial intuition enters the scene. Call *Label Invariance* the intuition that the way in which we label the elements of A should not affect the probabilities of selecting some a_i from Even, Odd, One, etc. Let us then re-label the original list $a_1, a_2, a_3, \ldots$ to produce a new list $b_1, b_2, b_3, \ldots$ according to the following rules:

- Rule I: $a_{i=2m} \mapsto b_{i=4m}$
- Rule II: $a_{i=4m+3} \mapsto b_{i=4m+2}$

[39] Bartha does not provide any reasoning for Equation 4.3.1.

- Rule III: $a_{i=4m+1} \mapsto b_{i=2m+1}$

Now let $\mathsf{Odd-New}$ be the event that the selected b_i has an odd label and let $\mathsf{Even-New}$ be the event that the selected b_i has an even label.

By Label Invariance, the same reasoning as that given for Equation 4.3.1 should yield

$$Pr_R(\mathsf{Odd-New}) = Pr_R(\mathsf{Even-New}) = \frac{1}{2}. \tag{4.3.3}$$

However, by Rule III, $\mathsf{Odd-New}$ represents that same subset of individuals as One, so we should have

$$Pr_R(\mathsf{Odd-New}) = Pr_R(\mathsf{One}) = \frac{1}{4}. \tag{4.3.4}$$

Contradiction.

Bartha claims that there are two key intuitions at work here: (i) Label Invariance and (ii) Equiprobability; furthermore, he identifies Label Invariance as the problematic intuition, a claim that I will make sense of in my own framework below. Despite our agreement here, I do not think he teases out what is of fundamental philosophical interest: a conflict between our infinitary intuitions and the information that must be carried by our theory of the infinite in order to validate these intuitions. Let me make this more precise.

Begin first with Equiprobability. Bartha initially describes this intuition as: for tickets a_i, a_j in our lottery with $i, j \in \mathbb{N}$ and $P(a_i)$, $P(a_j)$ real numbers representing the probability of drawing a_i or a_j, respectively, we have $P(a_i) = P(a_j)$ for all i, j. However, it is not quite this intuition that plays a role in the Label Invariance paradox. Rather, it is the fact that "...positing a relationship of equiprobability *between sets of outcomes* in a countably infinite set leads to paradox" ([Bartha, 2004], 310; emphasis my own).[40] The above construction of relative probability functions and induced probability measures allows Bartha to describe this relationship more precisely and yields claims like $Pr_R(\mathsf{Odd}) = Pr_R(\mathsf{Even}) = 1/2$.[41]

However, the relevant relationship between infinite subsets of A, in particular that between Even and Odd, is not a probabilistic one. It is, simply, the intuition that the even numbers (resp. odds) comprise one-half of the natural numbers. Call this the *Relative Size Intuition*.[42] The Relative Size Intuition requires, first, that we take **PW** on board. This is because, minimally, we want the size assignment for the evens to be strictly less than that for the naturals, given that the evens are a proper subset of the naturals. It also appears to require something like **FR**,[43] since the intuition that the even numbers comprise one-half of the naturals plausibly arises from the idea that every other natural number is even.[44] In what follows, it is sufficient to consider the weaker **PW** underlying the relative size intuition. In sum, Bartha is not ultimately interested in Equiprobability, but rather the Relative Size Intuition, which makes no essential appeal to anything involving probability.[45]

[40] Presumably, he would want to say "infinite sets of outcomes of equal cardinality."

[41] Again, see my brief justification of this intuition above.

[42] Indeed, this is why Bartha uses relative probability functions and measures induced from these.

[43] Recall that the **Frequency (FR)** intuition says that if infinite sets A and B occur "equally often" in an ambient set C, then A and B have the same size.

[44] Compare the reasoning behind Equation(s) 4.3.1 and 4.3.2.

[45] This is obscured by the fact that Bartha extends the "Equiprobability intuition for singletons" to some-

Now let us turn to Label Invariance. Like Equiprobability (and, ultimately, the Relative Size Intuition), this is stated in terms of the probabilistic formalism, but I think this inessential. All that we require is more fundamental information about infinite sets. There are two ingredients to the infinitary Label Invariance assumption: (i) there are relabellings of a countably infinite set A; (ii) these relabellings should not affect our probability assignments. Note that (i) is just saying that there are permutations of A, viz., bijections from A to itself. This is unobjectionable in many mathematical contexts.[46] However, since the only information these bijections carry is 1-1 correspondence between the elements of A, we must immediately abandon component (ii) of Label Invariance. This is because the "probabilities" mentioned in (ii), in the particular context we are treating, are supposed to reflect part-whole relations between infinite subsets of A, codified by the Relative Size Intuition, and this information is not preserved by Cantorian bijections.[47] It is important to note that, in this context, Label Invariance is a "mixed" intuition: it combines both **CP** and **PW**. This need not be the case if the "probabilities" mentioned in (ii) are preserved under 1-1 correspondence; in such a situation, the labellings are indeed irrelevant.[48] However, if further information is desired, say, either **PW** or **FR**, then we are in trouble. Thus, since both Label Invariance and the Relative Size Intuition can be interpreted in contexts which allow for the mixing of different infinitary intuitions, expressed by **PW** and **CP**,[49] contradictions can arise. And this is because **PW** conflicts with **CP** on infinite sets. Thus, we should conclude that the Label Invariance Paradox is not a "probabilistic" paradox at all. It simply reflects our conflicting intuitions concerning how to measure the "size" of infinite sets. This should be the diagnosis of the Label Invariance Paradox.

Now that we have our diagnosis, we can determine how to resolve the paradox. Clearly, we should consistently apply only one infinitary intuition. Let us say that we wish to validate **PW**. In order to do so, we should reject any sort of Label Invariance condition when we wish to preserve information about the relative sizes of infinite sets of the same cardinality. Consider a more general Label Invariance condition: (i′) there are relabellings of A; (ii′) these should not affect our relative size assignments of infinite sets. Clearly, again, (i′) is simply an existence condition that would be difficult to question. But if we wish to hold (i′) and (ii′) together, then the relabellings in question cannot just be Cantorian bijections; they must preserve the relative size assignments of infinite subsets of A. But Cantorian bijections will not do this on infinite sets of the same cardinality.

The presentation of numerosities in terms of labelled sets provides another way of seeing that Label Invariance must be rejected. Let us simplify Bartha's discussion somewhat by taking $A = \mathbb{N}$ (rather than some countably infinite set of arbitrary elements) and letting Odd and Even be the sets of odd natural numbers and even natural numbers, respectively.[50]

thing like "Equiprobability for infinite subsets."

[46] That is, excluding very weak mathematical contexts in which the existence of permutations of A might fail.

[47] It is also worth noting that Label Invariance and the use of Cantorian cardinalities contradict *Finite Additivity* (FA), one of Kolmogorov's axioms for classical probability theory. See Appendix F. By Cantorian cardinalities, Even, Odd, and $\mathbb{N}$ all have the same measure, and hence the same probability assignment. However, it is also the case that $\mathbb{N} = \mathsf{Even} \cup \mathsf{Odd}$. Thus $P(\mathbb{N}) = P(\mathsf{Even} \cup \mathsf{Odd}) \neq P(\mathsf{Even}) + P(\mathsf{Odd})$.

[48] Clearly, this is the case where A is finite.

[49] And sometimes also **FR**.

[50] Rather than, as in Bartha's paper, the sets comprised of $a_i \in A$ with odd and even labels, respectively.

Similarly, let $\mathsf{One} = \{1, 5, 9, 13, \ldots\}$. As in the Label Invariance paradox, select a bijection, or "relabelling," $\varphi : \mathsf{One} \to \mathsf{Odd}$ that takes $1 \mapsto 1$, $5 \mapsto 3$, $9 \mapsto 5$, etc.[51] This relabelling plays the same role as Rule III in the Label Invariance paradox. Now consider One and Odd as labelled sets, i.e., $(\mathsf{One}, \ell_{\mathsf{One}})$ and $(\mathsf{Odd}, \ell_{\mathsf{Odd}})$. Note that, since $\mathsf{One} \subseteq \mathbb{N}_0$ and $\mathsf{Odd} \subseteq \mathbb{N}_0$, they are equipped with the canonical label, i.e., $\ell_{\mathsf{One}}(n) = n$ for $n \in \mathsf{One}$ and $\ell_{\mathsf{Odd}}(n) = n$ for $n \in \mathsf{Odd}$.[52] From Definition 4.2.4 above, we saw that two labelled sets are isomorphic iff there is a bijection preserving their labels. Further, if two labelled sets are non-isomorphic, then their numerosities will differ.

We wish to see what happens when our relabelling φ interacts with ℓ_{One} and ℓ_{Odd}. That is, we wish to know whether, for all $a \in \mathsf{One}$, $\ell_{\mathsf{Odd}}(\varphi(a)) = \ell_{\mathsf{One}}(a)$. Let $a = 5$. Then $\varphi(5) = 3$ by the definition of φ. Now note that, by our canonical labelling, $\ell_{\mathsf{Odd}}(3) = 3$, but $\ell_{\mathsf{One}}(5) = 5$. Thus,

$$\ell_{\mathsf{Odd}}(\varphi(5)) \neq \ell_{\mathsf{One}}(5). \tag{4.3.5}$$

Thus, since our relabelling bijection φ does not preserve the canonical labels of $(\mathsf{One}, \ell_{\mathsf{One}})$ and $(\mathsf{Odd}, \ell_{\mathsf{Odd}})$, we conclude that these sets are non-isomorphic under φ (by Definition 4.2.4). By Proposition 4.2.6, the associated counting functions for these sets will differ, and by Proposition 4.2.9 and the definition of numerosity, $\mathfrak{n}(\mathsf{One}) \neq \mathfrak{n}(\mathsf{Odd})$. This disagreement in numerosities is sufficient to block the paradox. This is because the contradiction derived by Bartha, namely that we have both $Pr_R(\mathsf{Odd} - \mathsf{New}) = 1/2$ and $Pr_R(\mathsf{Odd} - \mathsf{New}) = Pr_R(\mathsf{One}) = 1/4$, ultimately derives from the fact that we collapse $\mathsf{Odd} - \mathsf{New}$ into One. But this is not validated by numerosity theory.

This provides us with an independent argument against the general Label Invariance intuition that makes no appeal whatsoever to probabilistic notions. We initially arrived at the rejection of Label Invariance by analyzing Bartha's argument couched in probabilistic formalism. However, we might have proceeded as follows: given some countably infinite set A, we assume the general Label Invariance condition. We know that the Relative Size Intuition requires that **PW** be satisfied. We then try to use our only theory of infinite counting that satisfies **PW**, the theory of numerosities. Following one presentation of the theory (in terms of labelled sets), we then see by the above reasoning that Label Invariance cannot be maintained in general, and this is because arbitrary bijections need not preserve the labels of labelled sets. Thus, we reject Label Invariance, and the paradox does not get off the ground, prior to the introduction of anything having to do with probability.

We will discover that the Label Invariance intuition is something of a silent player in many other so-called probabilistic paradoxes. It will cause difficulties below in both Williamson's infinite sequence of coin tosses and in Bertrand's paradox. I believe its pervasiveness arises from uncritical generalizations of finitary intuitions to the infinite case and from a reliance upon **CP**.

I followed Bartha in choosing my bijection φ to match Rule III in his version of the paradox. However, a natural thought immediately arises: perhaps we can construct bijections satisfying further conditions that will preserve relative size assignments? This is more or less Bartha's solution (and also what is accomplished by using isomorphisms of labelled sets in

[51] We need only consider Odd here, not some relabelled set $\mathsf{Odd} - \mathsf{New}$.

[52] Consider Example 16.3 of [Benci and Nasso, 2019].

numerosity theory); however, his solution simply bottoms out in favoring **PW** rather than **CP**, and thus supports my analysis of the Label Invariance paradox as a paradox of infinity.

Bartha's Solution

It is worth considering Bartha's own solution to the paradox since it is structurally analogous to what I have proposed. Let us see why.

My solution focused upon our background theory of infinite counting. In particular, we saw that, if we use numerosities (or any possible theory vindicating **PW**) to retain the possibility of validating the Relative Size Intuition, then we must reject Label Invariance. On the other hand, Bartha's solution focuses upon the background theory of "symmetries" employed in probability theory. His idea is, first, to identify what kinds of symmetries are "admissible" on our outcome space. He then uses these symmetries to validate Equiprobability (and, ultimately, the Relative Size Intuition via relative probabilities) and rejects Label Invariance. I find this approach somewhat less clean as it makes appeal to vexed notions like "symmetry" and the Principle of Indifference. More to the point, it ultimately seems to rely on favoring **PW** rather than **CP** and so reduces to my solution.

Bartha wishes to show that it is the Label Invariance intuition that causes issues. He defines a "symmetry" as a bijection $\theta : X \to X$, where X is our outcome space, that "preserves all features that bear on probability, [which] vary from case to case" ([Bartha, 2004], p. 312). He then makes a distinction between "coherent" and "incoherent" symmetries via the following condition:[53]

Definition 4.3.1 (Coherent/Regular Symmetries)**.** A set Θ of symmetries on X is *coherent* or *regular* if the following conditions hold:

1. Θ is a group under function composition.

2. There do not exist (i) a non-empty subset Y of X and (ii) $m, n \in \mathbb{N}$ with $m > n$ such that there are symmetries $\theta_1, \ldots, \theta_m$ and $\psi_1, \ldots, \psi_n$ in Θ satisfying

$$\bigsqcup_{i=1}^{m} \theta_i(Y) \subseteq \bigcup_{j=1}^{n} \psi_j(Y), \tag{4.3.6}$$

   where $\theta_i(Y)$ denotes the image of non-empty $Y \subset X$ under symmetry θ_i (similarly, for $\psi_j(Y)$).

The second condition might be called a "no collapse" condition in the sense that it rules out placing m disjoint copies of Y inside n copies of Y with $m > n$. It can be shown that such symmetries allow us to define the Equiprobability assumption and relative probabilities rigorously.[54] However, it is clear that the "symmetries" present in the Label Invariance Paradox fail to be coherent. Indeed, this is because the Label Invariance assumption is equivalent to claiming that any bijection on a countably infinite set is an admissible "symmetry."

[53] I have edited the condition slightly for clarity.

[54] See the Appendix to [Bartha, 2004] and [Bartha and Johns, 2001] for details.

However, not all relabellings/bijections satisfy the no collapse condition. Consider $Y = \mathsf{Odd}$ and $m = 2, n = 1$. There exist relabellings

$$\theta_1 : \mathsf{Odd} \to \mathsf{One} \tag{4.3.7}$$

$$\theta_2 : \mathsf{Odd} \to \mathsf{Three} \tag{4.3.8}$$

such that

$$\bigsqcup_{i=1}^{2} \theta_i(\mathsf{Odd}) = \psi_1(\mathsf{Odd}) \tag{4.3.9}$$

with ψ_1 the trivial relabelling, i.e., we have $\mathsf{One} \sqcup \mathsf{Three} = \mathsf{Odd}$. Hence, the no collapse condition is violated. Thus, we reject the Label Invariance assumption and accept the Equiprobability assumption via appeal to admissible symmetries on our outcome space.

It should be apparent that the no collapse condition proposed by Bartha just says that, given an infinite subset Y of a countably infinite outcome space X, an admissible symmetry cannot place Y in 1-1 correspondence with its proper subsets. But this is just a preference for **PW** over **CP**. Thus, I think any proposed solution will ultimately bottom out in my own in terms of infinitary intuitions.

4.3.3 God's Lottery

Once we reject the Label Invariance assumption (at any level of generality), we have dissolved the Label Invariance paradox. We are then left with the Relative Size Intuition, which Bartha glosses as Equiprobability for infinite subsets of A; the original Equiprobability intuition for singletons also remains. However, this returns us to the initial predicament discussed by Bartha: we cannot describe a uniform probability distribution over a countably infinite set using the axioms of Kolmogorov's probability theory. We are still faced with the conflict between Equiprobability and Countable Additivity. One possible way to deal with this conflict would be the solution described by Bartha, that is, provide a formalism with which to describe Equiprobability such that CA need not apply (but the rest of Kolmogorov's axioms do). The other possible way would be to use non-standard probability measures.[55] I will now consider this approach by examining the recently developed Non-Archimedean Probability (NAP) theory, which relies upon the theory of numerosities. In particular, NAP theory can describe "God's Lottery" and does so such that the probability assigned to an event is directly proportional to the numerosity of the subset representing the event. God's Lottery is a version of de Finetti's Lottery for objective probabilities (chance). This is made clear at the outset of [McCall and Armstrong, 1989]:

> [...] God's lottery was different in that the winning number would be chosen from the set of *all* positive integers. The winning number would, of course, be finite, but no upper limit could be placed on the range from which it would be selected.

[55] NAP theory is a recent iteration of this much older idea, viz., that the domain, range, or both domain and range of the probability function be non-Archimedean sets. See [Loeb, 1975], [Lewis, 1980], [Skyrms, 1980], [Cutland, 1983], and [Nelson, 1987].

> In an entirely fair and purely random way, God would simply choose a number. The fairness of the procedure would consist in the fact that the chance of any given number being chosen would be the same as that of any other number (*ibid.*, 223).

Again, the exact same set of concerns about the "sizes" of infinite subsets of a countably infinite set come into play here. This is unaffected by the distinction between credence and chance. However, other considerations may differ depending on whether one considers credence or chance. For instance, as we saw above, Bartha employs Dutch Book arguments, which, I take it, are irrelevant to discussions of chance.[56]

I now describe how NAP theory can be used to describe God's lottery and emphasize the role that the sensitivity of numerosities plays in providing the solution. First, I show how numerosity theory can validate the part-whole assumption that, e.g., the "size" of the evens should be greater than that of the "size" of set Two above. Then, I show how this relationship can be encoded by the NAP theory.

The Label Invariance paradox has made clear that an uncritical use of relabellings of infinite sets will produce contradictions unless we are very clear about the information we wish to be preserved under these relabellings. It is here that the "sensitivity" of numerosities, so often assumed to be a disadvantage of the theory, can be turned into an advantage. First, by examining the presentation of numerosity theory in terms of labelled sets (Type 1 sensitivity), we were able to produce an independent argument against Label Invariance without making any appeal to probability. Second, by choosing our ultrafilter appropriately, or, equivalently, which subsets $Q \subseteq \mathbb{N}$ are qualified[57] (Type 2 sensitivity), we show that $\mathbb{N}$ can be partitioned into k-many equinumerous sets for any $k \in \mathbb{N}$ and thereby validate our intuitions about the relative sizes of these partitions (for various choices of k).

Indeed, consider again the size relations encoded by the Relative Size Intuition. That is, we wish to say that Even comprises half of the natural numbers, Two comprises one-fourth of the naturals, etc.[58] Then, assuming Equiprobability for pairwise disjoint elements, we get the intuitive probability assignments $P(\mathsf{Even}) = 1/2 > P(\mathsf{Two}) = 1/4$. Now consider the following theorem (refer to Appendix E for discussion of the technical notions, e.g., the Qualified Set Axiom, QSA):

Theorem 4.3.2. *Assume* $(\mathsf{QSA})_Q$ *where* $Q = \{m! : m \in \mathbb{N}\}$. *Then, for every* $k \in \mathbb{N}$, *the number* $\alpha \in \mathbb{N}^*$ *is a multiple of* k *and the numerosity of the set of multiples of* k *is*

$$\mathfrak{n}(\{k, 2k, 3k, \ldots, nk, \ldots\}) = \frac{\alpha}{k}. \tag{4.3.10}$$

From this theorem, we are able to consistently formulate the Relative Sizes Intuition implicit in God's Lottery and the Label Invariance paradox. Given the properties of $\mathfrak{n}(\cdot)$ we have

$$\mathfrak{n}(\mathsf{Even}) + \mathfrak{n}(\mathsf{Odd}) = \mathfrak{n}(\mathsf{Even} \sqcup \mathsf{Odd}) = \mathfrak{n}(\mathbb{N}) = \alpha. \tag{4.3.11}$$

[56] See [Hájek, 2009] for an illuminating discussion of Dutch Book arguments.

[57] See Appendix E for discussion of qualified sets.

[58] Going forward, I will use Even, Odd, etc. to denote the set of evens, the set of odds, etc. rather than the event that an element of a countably infinite set A has an even label, odd label, etc.

Applying Theorem 4.3.2 we conclude that $\mathfrak{n}(\mathsf{Even}) = \mathfrak{n}(\mathsf{Odd}) = \alpha/2$. Furthermore, given that sets $\mathsf{One}, \mathsf{Two}$, etc. are all multiples of four, we can conclude

$$\mathfrak{n}(\mathsf{One}) = \mathfrak{n}(\mathsf{Two}) = \mathfrak{n}(\mathsf{Three}) = \mathfrak{n}(\mathsf{Four}) = \frac{\alpha}{4}. \tag{4.3.12}$$

Finally, we conclude that

$$\mathfrak{n}(\mathbb{N}) = \alpha > \mathfrak{n}(\mathsf{Even}) = \frac{\alpha}{2} > \mathfrak{n}(\mathsf{Two}) = \frac{\alpha}{4}, \tag{4.3.13}$$

thereby validating our intuitions about the relative sizes of infinite subsets of $\mathbb{N}$.

Now let us see how these size assignments can be reflected by our NAP probability assignments. I follow the presentation given in [Benci et al., 2013]. The key concept needed for applying NAP to God's Lottery is that of a Λ-limit.[59] Defining this limit takes a bit of work; refer to Appendix G for my discussion of the conceptual underpinnings of the formalism.

Once the notion of a Λ-limit is in hand, one can validate the relative size assignments of infinite subsets of $\mathbb{N}$ (given by numerosities) in NAP theory. We want the probability of drawing an even number from $\mathbb{N}$ to be the same as that of drawing an odd number from $\mathbb{N}$, reflecting the Relative Size Intuition that precisely one-half of the naturals are even (resp. odd). In NAP theory, we deal with triples (Ω, P_N, J), where Ω is the outcome space of elementary events, $P_N : \mathscr{P}(\Omega) \to \mathbb{F}^*$ is the NAP probability function[60] with $\mathbb{F}^*$ a non-Archimedean field, and J an algebra homomorphism. Using NAP theory, we can produce a general result with $P_N(\mathsf{Even}) = P_N(\mathsf{Odd})$ as a special case. For any infinite set of k-tuples of natural numbers, written as $\mathbb{N}_k$, we have

$$P_N(\mathbb{N}_k) = \frac{1}{k} \tag{4.3.14}$$

by choosing $\Lambda = \{m! : m \in \mathbb{N}\}$. The main conceptual point here is that the Λ-limit, which will itself provide the value of $P_N(\mathbb{N}_k)$, the NAP probability assignment for $\mathbb{N}_k$, depends upon our choice of Λ. (Once more, the technical details are involved; see the discussion of Axiom C.4 in Appendix G.) This choice of Λ as $\{m! : m \in \mathbb{N}\}$ is exactly as it should be, since this is precisely the set that yields Theorem 4.3.2, and we wish our probability assignment for event $E \in \mathscr{P}(\Omega)$ to be directly proportional to the numerosity of the subset representing

[59] Note that the Λ-limit is very similar to the alpha-limit discussed above and enjoys many of the same properties. The (slight) differences are spelled out in Appendix G. See also Section 3.6 of [Benci et al., 2018] for a discussion of the formal connections between numerosities and NAP theory.

[60] Distinguished from a Kolmogorov probability function by the N subscript.

E. We may then explicitly compute:[61]

$$P_N(\mathbb{N}_k) = \frac{\lim_{\lambda_n \in \Lambda} |\mathbb{N}_k \cap \{1, \ldots, m!\}|}{\alpha} \tag{4.3.15}$$

$$= \frac{\lim_{\lambda_n \in \Lambda} \left[\frac{n}{k}\right]}{\alpha} \tag{4.3.16}$$

$$= \frac{\lim_{\lambda_n \in \Lambda} \frac{n}{k}}{\alpha} \tag{4.3.17}$$

$$= \frac{1}{k} \tag{4.3.18}$$

as desired. In particular, letting $k = 2$, we get $\mathbb{N}_2 = \mathsf{Even}$ and $P_N(\mathsf{Even}) = \frac{1}{2}$. Thus, we are able to validate our basic intuitions about a fair lottery on $\mathbb{N}$, i.e., the probability of drawing an even number is precisely $\frac{1}{2}$, since the probability of drawing a natural number is precisely one and we have made an appropriate choice for Λ. Similarly, writing $\mathbb{N}_{k,l} = \{k - l, 2k - l, \ldots, nk - l, \ldots\}$ with $l \in \{0, \ldots, k - 1\}$, we obtain $\mathsf{Odd} = \mathbb{N}_{2,1}$ and find $P_N(\mathsf{Odd}) = \frac{1}{2}$, as desired.

This solution has been effected by an appropriate choice of Λ. Of course, if we had chosen Λ differently, we might have gotten a different probability assignment. However, as I noted above, this non-uniqueness is something of an advantage. As long as we are able to get clear on what probabilistic intuitions we wish to validate, we can use numerosities and NAP theory to satisfy those intuitions given our choice of ultrafilter or Λ, respectively. Thus, the sensitivity of these theories is in fact a powerful tool.

It might be interesting to see whether one could construct infinite subsets of the natural numbers that exhibit a more worrisome amount of indeterminacy. For instance, in [Benci et al., 2013] it is shown that different choices for Λ will yield either $P_N(\mathsf{Even}) = \frac{1}{2}$ or $P_N(\mathsf{Even}) = \frac{1}{2} - \frac{1}{2\alpha}$. That is, a different choice of Λ will leave the standard part of the probability fixed but induce infinitesimal differences. However, one might not be bothered by the infinitesimal differences. Perhaps, though, one could construct subsets whose probability assignments differ with respect to their *standard* parts according to choice of Λ. However, *prima facie*, my sense is that this project would be of limited philosophical interest (though not necessarily mathematical interest). This is because we were interested in using numerosities and NAP theory to capture basic intuitions we have about relatively well-behaved subsets: evens, odds, multiples of k. When we pass to even slightly more complicated subsets, e.g., arbitrary arithmetic progressions, it simply does not seem that we have any fixed intuitions about the relative size of such subsets in $\mathbb{N}$. It would seem that, in this and related cases, we have no such intuitions that prompted our investigations of God's Lottery. Thus, if there is indeterminacy in computing probabilities for these sets, so be it.

[61] See [Benci et al., 2013], pp. 140-143 for full details. It is interesting that the intuition validated here is not simply Part-Whole but rather a Density intuition. I will discuss this in other work.

4.3.4 Williamson's Infinite Sequence of Coin Tosses and Label Invariance

The Argument

Another case to which my analysis might be fruitfully applied is Williamson's "infinite sequence of coin tosses" ([Williamson, 2007]). Here Williamson produces an argument for the conclusion that infinitesimal probability assignments cannot save a "regularity" constraint on our probability theory. Intuitively, regularity says that the probability of a possible event should be strictly larger than the probability of an impossible event and that only necessary truths should have probability 1 and necessary falsehoods probability 0. Classical probability theory cannot accommodate regularity when we wish to deal with infinite (either countable or uncountable) sets of possibilities. However, as we have seen with NAP theory, why not then simply allow that our probability function take on infinitesimal values?

Against this, Williamson argues as follows. (I use his notation throughout.) Consider a fair coin tossed infinitely many times at one second intervals. Let $H(1\ldots)$ be the event that every toss comes up heads and let $H(2\ldots)$ be the event that "every toss after the first comes up heads." That is, $H(1\ldots)$ may be written as $H(1\ldots) := H(1) \wedge H(2\ldots)$, where $H(1)$ is the event that the first toss comes up heads. Let $P(X)$ denote the probability of event X and let the probability of X conditional on Y be given by the usual formula

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)}. \tag{4.3.19}$$

Given these definitions and letting $X = H(2\ldots)$ and $Y = H(1)$, we can write:

$$P(H(1\ldots)) = P(H(1) \wedge H(2\ldots)) = P(H(1)) \cdot P(H(2\ldots)|H(1)). \tag{4.3.20}$$

Because the toss is fair, we know that $P(H(1)) = \frac{1}{2}$. Furthermore, because the tosses are independent of one another, the conditional probability $P(H(2\ldots)|H(1))$ is simply $P(H(2\ldots))$. Substituting these equalities into Equation 4.3.20, we obtain:

$$P(H(1\ldots)) = \frac{1}{2}P(H(2\ldots)). \tag{4.3.21}$$

However, Williamson goes on to argue that $H(1\ldots)$ and $H(2\ldots)$ are "isomorphic events" and thus we should have $P(H(1\ldots)) = P(H(2\ldots))$. His defense of this isomorphism runs as follows:

> More precisely, we can map the constituent single-toss events of $H(1\ldots)$ one-one onto the constituent single-toss events of $H(2\ldots)$ in a natural way that preserves the physical structure of the set-up just by mapping each toss to its successor. $H(1\ldots)$ and $H(2\ldots)$ are events of exactly the same qualitative type; they differ only in the inconsequential respect that $H(2\ldots)$ starts one second after $H(1\ldots)$. That $H(2\ldots)$ is preceded by another toss is irrelevant, given the independence of the tosses. Thus $H(1\ldots)$ and $H(2\ldots)$ should have the same probability ([Williamson, 2007], 175).

Finally, we conclude that $P(H(1\ldots)) = \frac{1}{2}P(H(2\ldots))$ and $P(H(1\ldots)) = P(H(2\ldots)$. These

equations can only hold simultaneously when $P(H(1\ldots)), P(H(2\ldots))$ are identically zero. This is true whether we deal with infinitesimals or not. Thus, since $H(1\ldots)$ is, intuitively, a possible event, infinitesimals cannot preserve regularity.

My Analysis

I think many aspects of Williamson's argument are severely underdetermined. The two notions most in need of clarification are those of "event" and "isomorphism." It may simply be an unfortunate turn of phrase, but Williamson calls both the *sequences* of coin tosses (i.e., both $H(1\ldots)$ and $H(2\ldots)$) and the "constituent" single coin tosses "events." This produces confusion because it is unclear whether Williamson is taking "event" here in the technical probabilistic sense or in the more mundane sense of a physical occurrence.

As described in Appendix F, a Kolmogorov probability space is a triple $(\Omega, \mathscr{F}, P)$. Ω is the sample space consisting of "elementary events," viz., singletons representing single outcomes in Ω. The "events" of this probability space are measurable subsets of Ω and elements of $\mathscr{F}$. According to Williamson's set-up and other mathematical models of infinite sequences of coin tosses,[62] our Kolmogorov probability space should be $(\Omega, \mathscr{F}, P)$ with $\Omega = \{0,1\}^{\mathbb{N}}$, the set of countably infinite sequences of 0s and 1s, $\mathscr{F}$ the σ-algebra generated by "cylinder sets,"[63] and $P : \mathscr{F} \to [0,1]$ the unique probability measure that extends $\mu\left(C^{(i_1,\ldots,i_n)}_{(t_1,\ldots,t_n)}\right) = 2^{-n}$ (guaranteed by the Carathéodory extension theorem). Considered in this way, clearly only $H(1\ldots)$ and $H(2\ldots)$ are events. Furthermore, if we consider these events as unstructured sets (as Williamson's language of "one-one" correspondence would suggest; more on this below), then these events cannot be isomorphic. In particular, since the outcome space is $\Omega = \{0,1\}^{\mathbb{N}}$, as a subset of this outcome space, $H(1\ldots)$ must be represented as a singleton $\{s\}$, namely, the countably infinite sequence with all 1s. On the other hand, as we can see from Williamson's description, $H(2\ldots)$ should not be a singleton. Indeed, he says that "[...] $H(2\ldots)$ is the event that every toss after the first comes up heads" ([Williamson, 2007], 175). This suggests that $H(2\ldots)$ is a *disjunctive* event, since the first toss could be heads or tails. This event is given by the pair $\{s,t\}$ with s as above and t the sequence in which the first term is 0, followed by all 1s. Thus, since obviously a singleton cannot be isomorphic to a pair, Williamson's argument falls apart.[64]

But perhaps Williamson intends "event" to mean "physical occurrence." This is suggested, first, by his talk of "preserving the physical structure of the set-up" and, second, by the fact that he indiscriminately calls sequences of tosses and individual tosses "events." By the principle of charity, I think we should adopt this interpretation. Without it, we must saddle Williamson with negligence concerning both the set-theoretic issue above and his claims that there are "constituent events" (a single coin toss cannot be an "event" in $\{0,1\}^{\mathbb{N}}$). Under this interpretation, Williamson's argument can get off the ground but, I must admit, remains perplexing. For then it would appear the argument is really about

[62] See, for instance, [Benci et al., 2013] and [Benci et al., 2015].

[63] A cylinder set of codimension n is constructed using an n-tuple of indices $(i_1,\ldots,i_n)$ and an n-tuple of elements in $\{0,1\}$, viz., $(t_1,\ldots,t_n)$, where each t_j is either 0 or 1. We then define a cylinder set of codimension n to be $C^{(i_1,\ldots,i_n)}_{(t_1,\ldots,t_n)} = \{\omega \in \Omega : \omega_{i_j} = t_j\}$. This represents the event that for every $j = 1,\ldots,n$, the i_jth coin toss yields t_j as outcome.

[64] This same criticism can be found in [Howson, 2017], which I encountered after writing this.

what makes physical occurrences "qualitatively identical" rather than anything about our constraints on probability theory. Of course, these topics cannot be cleanly separated, but this would constitute, what seems to me, a shift of focus. This is because Williamson is ultimately concerned with a formal property of a mathematical theory of probability (regularity). It is true that he uses an example that admits of a physical description to establish his point; however, to my mind, the example is idealized to such an extent that it renders questions about the nature of physical events irrelevant. In any case, for an interesting discussion of Williamson's argument in terms of qualitative differences between physical events, see [Parker, 2021]; however, I find it hard to see that the disambiguation required for Williamson's argument comes from "[...] Special Relativity so that there is no absolute inclusion relation between times." This seems an *ad hoc* patch and rather far afield from where we started.

Before we attempt to find a more satisfying interpretation of Williamson's argument, we must get clearer on what he means by "isomorphism." For without the premiss that $H(1\ldots)$ and $H(2\ldots)$ are isomorphic in some sense, his argument fails to go through. Unfortunately, the premise is a weak one because there is no univocal meaning of "isomorphism," and Williamson does not elaborate on this notion. Thus, we must specify further the kind of structure we wish to preserve under the isomorphism in question. This is a familiar notion in mathematics: we must specify the category in which an isomorphism holds (e.g., Sets, Topological Spaces, Differentiable Manifolds, whatever). Even internal to the category of sets, we can impose additional structure to be preserved. This is what we witness in the theory of numerosities: we wished to preserve **PW** and so needed an isomorphism of labelled sets, not unstructured sets.

With these distinctions in hand, we can assess possible interpretations of the experiment. One possibility would be to assume Williamson really is just interested in the physics and metaphysics of events and how these topics relate to probability. In this case, the "isomorphism" in question would not be a mathematical notion at all.[65] Another possibility would be to assume Williamson is strictly interested in the formalism of probability theory. However, in this case, the isomorphism in question must be a 1-1 mapping, and no such mapping exists between singletons and pairs. An intermediate interpretation between these extremes is, I think, preferable.

Let us assume that Williamson intends "event" to mean "physical occurrence." However, this need not imply that his "isomorphism" has much to do with physical events at all. Williamson justifies the "isomorphism" between $H(1\ldots)$ and $H(2\ldots)$ by noting that "[...] we can map the constituent single-toss events of $H(1\ldots)$ one-one onto the constituent single-toss events of $H(2\ldots)$ in a natural way that preserves the physical structure of the set-up..." That is, the desired isomorphism comes from the one-one correspondence between the individual tosses $\{H(1), H(2), H(3), \ldots\}$ comprising the sequence $H(1\ldots)$ and the individual tosses $\{H(2), H(3), H(4), \ldots\}$ comprising the sequence $H(2\ldots)$. This is a cogent isomorphism between infinite sets. But why think this preserves physical structure? On this point, Williamson is silent, but perhaps he is thinking as follows. First, the isomorphism in question maps each constituent toss in $H(1\ldots)$ to its temporal successor in $H(2\ldots)$. This only changes the temporal index of each toss in $H(1\ldots)$. That is, we have merely "relabelled"

[65] See both [Parker, 2019] and [Parker, 2021] for attempts to make sense of this.

as above in the Label Invariance paradox, and so, according to Williamson, everything should remain the same, including probabilities. Furthermore, the way Williamson has set up the experiment guarantees that all other qualitative physical properties and relations are preserved (the coin tosses are from the same fair coin, the time intervals are exactly the same, etc.). Thus, the claim that $H(1\ldots)$ and $H(2\ldots)$ are isomorphic events (in the sense of physically occurring sequences of tosses) is undergirded by a preference for **CP**, a Label Invariance intuition, and the stipulations of the experiment.[66]

By why prefer **CP** in the first place? This question leads us to an earlier, important criticism of Williamson by Weintraub ([Weintraub, 2008]). She argues that, since we should construe $H(2\ldots)$ as a *subsequence* of $H(1\ldots)$, Williamson's "isomorphism" fails to preserve physical characteristics of the situation. Indeed, she says,

> Williamson's example shows that isomorphism doesn't preserve *all* basic physical properties. He claims that the two 'sequences of events[67] are of *exactly* the same qualitative type' [...] But although all the physical properties of the *constituent events* are preserved by the mapping, as are the temporal intervals between adjacent tosses, there is a *global* property (of the complex event) which is not preserved. The second sequence is a *proper subset* of the first (*ibid.*, p. 249).

Again, we must be careful here as it is not at all clear that Williamson intends $H(2\ldots)$ to be a proper subset of $H(1\ldots)$.[68] Nonetheless, Weintraub's concerns clearly do apply to the second iteration of the experiment in which there are two coins and two sequences running in parallel ([Williamson, 2007], pp. 175-6; discussed below), so the point is well taken.

What is puzzling here is the fact that Williamson validates **CP** in order to preserve the physical features of each sequence of flips. Both $H(1\ldots)$ and $H(2\ldots)$ are temporal sequences. Then, assuming that time is infinite (which we must do anyway to make sense of the experiment), we see that there are more constituent tosses in $H(1\ldots)$ than in $H(2\ldots)$. $H(1\ldots)$ will take more time to complete given that it starts one second earlier. This is a relevant physical difference as Weintraub notes; however, I believe the Label Invariance intuition blinds him to this fact, and not mere prejudice for **CP**.

The root of the problem, then, is a conflict between mathematical and physical intuitions, rather than obvious technical errors or some preference for an unarticulated metaphysics of events. Put differently, this conflict is, once more, that between the coarseness of **CP** as criterion for measuring infinite sets (in this case used to produce an isomorphism) and the information about the sets we wish to retain. Williamson immediately proposes that we use **CP** to define an isomorphism between $H(1\ldots)$ and $H(2\ldots)$ via 1-1 correlating their "constituent" coin tosses. But he also has the intuition that relabellings are "inconsequential."

[66] It is reasonable to think that a more straightforward interpretation might be better here. Namely, we should think that it is the qualitative identity of the constituent tosses *alone* that supports the isomorphism claim. I was not satisfied with this interpretation because of Weintraub's criticism. It seems quite clear that the change in temporal index constitutes some sort of qualitative physical difference. So, the question becomes: why does Williamson think this difference is not relevant? The Label Invariance intuition is supposed to provide an answer to this question.

[67] Again, there is an infelicity here. We have sequences of tosses but not sequences of events in the probabilistic sense. This is further support for the claim that Williamson is interested in physical events.

[68] For example, if we interpret "event" probabilistically, the opposite claim is now true: $H(1\ldots) = \{s\}$ is a proper subset of $H(2\ldots) = \{s, t\}$

This intuition derives from the fact that, when measuring the size of unstructured sets using **CP**, the labels are indeed irrelevant. However, as we have seen, if you wish your measure of the relative sizes of infinite sets to be reflected in your probability assignments, the relabellings will matter. Furthermore, in this case the relabellings have a determinate physical meaning: they are the time indices for each individual coin toss. And this is a physical difference between the events in question. Thus, Williamson should jettison his background Cantorian framework in order to preserve physical characteristics of the toss sequences. But then, since, suitably construed, $H(2\ldots)$ is a proper subset of $H(1\ldots)$, we will no longer have an "isomorphism" of these events, and thus $P(H(1\ldots)) \neq P(H(2\ldots))$.[69] Therefore, Williamson's argument fails.

Dispensing with the Notion of Isomorphism

Once we settle on what is meant by "event," the issue with Williamson's argument lies with his choice of isomorphism and his background theory of infinity determining the kind of isomorphism we wish to use. Again, from elementary probability, he shows that

$$P(H(1\ldots)) = \frac{1}{2}P(H(2\ldots)). \tag{4.3.22}$$

Then, from his claim that $H(1\ldots)$ and $H(2\ldots)$ are "isomorphic" in some sense that preserves the physical situation, he shows that $P(H(1\ldots)) = P(H(2\ldots))$. As we have seen, this notion of isomorphism is fraught with difficulties. Is there then a way to show that $P(H(1\ldots)) = P(H(2\ldots))$ without appealing to such an "isomorphism"? It would seem a stock example from the theory of measure-preserving systems can help.[70]

Define measure μ on $\{0,1\}$ as

$$\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}. \tag{4.3.23}$$

Now consider our outcome space $\Omega = \{0,1\}^{\mathbb{N}}$ equipped with the infinite product measure

$$\mu_\infty = \prod_{\mathbb{N}} \mu. \tag{4.3.24}$$

Since Ω is a compact metric space, our σ-algebra $\mathscr{F}$ is the Borel σ-algebra. We then have a probability space $(\Omega, \mathscr{F}, \mu_\infty)$ that serves as a reasonable representation of Williamson's experiment. Now let $T : \Omega \to \Omega$ be a *left shift map* defined by

$$T(x_1, x_2, \ldots) = (x_2, x_3, \ldots). \tag{4.3.25}$$

This seems an adequate representation of Williamson's "relabelling" of temporal indices. Then $(\Omega, \mathscr{F}, \mu_\infty, T)$ is a *Bernoulli system*. It is well-known that Bernoulli systems are

[69] See pp. 147ff. of [Benci et al., 2013] for the details of how to actually assign such probabilities using NAP theory.

[70] For the details of this example, see the excellent text [Einsiedler and Ward, 2011], especially Chapter 2 and Appendix A.2.

measure-preserving systems, viz., for any event $E \in \mathscr{F}$ we have $\mu(T^{-1}E) = \mu(E)$. Hence, assuming $(\Omega, \mathscr{F}, \mu_\infty, T)$ represents what Williamson is after, applying T to "shift" the temporal indices of our events will not affect the probabilities of $H(1\ldots)$ and $H(2\ldots)$. Thus, we are able to get the desired premise $P(H(1\ldots)) = P(H(2\ldots))$ without appealing to any notion of isomorphism.

Though not our concern here, this still does not save his argument. And this is because, if we allow our measure to take hyperreal values, the *standard parts* of $P(H(1\ldots))$ and $P(H(2\ldots))$ will agree, but the non-standard parts need not. Thus, this premise does not hold, and thus Williamson cannot derive his desired conclusion.

Embedding Events in the Outcome Space

The analysis of Williamson's argument given in [Benci et al., 2018] suggests that there is yet a third ambiguity present. In particular, the authors claim that Williamson has failed to provide a well-defined sample space for his experiment:

> [...] the assignment of probabilities does not make sense in the absence of a well-defined sample space that is applied in a consistent way. In the case of Williamson's argument, a crucial aspect of fixing the sample space is an answer to the question, 'When does the count of events start?' (*ibid.*, 529).

At times, the authors misleadingly say that the "sample space has changed." We must understand this in the following way: all the "models" provided in [Benci et al., 2018] use the same *set* as the sample space. (In their countable case, $\Omega = \mathbb{N}$, and in their uncountable case, $\Omega = \{0,1\}^{\mathbb{N}}$.) However, the subsets of these sample spaces, to which we assign probabilities, are understood differently. In particular, there is, "[...] a different correspondence between sets in the event space and the situations in the (hypothetical) world" (*ibid.*, 528). This can be seen by considering the "relabelling" of the time indices of the sequences of tosses. The authors insist, though, that "it is not the labelling itself that is essential, but rather the choice of sample space and the embedding of events therein" (*ibid.*, 527).

I believe that the authors are picking up on an important issue, one that will take center-stage when we turn to Bertrand's paradox. However, I do not think it is playing much of a role in Williamson's argument. First, I do think the relabelling is important; or, at least, the intuition that relabelling elements of events should not affect our understanding of the events in any significant way (Label Invariance) is important. As we have seen, **CP** and Label Invariance underlie Williamson's argument for the "isomorphism" between $H(1\ldots)$ and $H(2\ldots)$. Second, it is implausible that Williamson is making the mistake of applying the sample space inconsistently, for in his second version of the experiment he seems to distinguish the so-called "different models" of the situation. Let us examine this more carefully.

Benci et. al. begin by describing two "situations" independent of probabilistic formalism. In Situation T1, a fair coin is tossed on all of a countably infinite collection of occasions, whereas, in Situation T2, a fair coin is tossed on all but the first of a countably infinite collection of occasions.[71] Evidently, Situation T1 is supposed to describe $H(1\ldots)$ and Sit-

[71] Once more, it is not clear that this is how Williamson is understanding the event $H(2\ldots)$, but I'll let the gloss of [Benci et al., 2018] stand for the purposes of argument.

uation T2 is supposed to describe $H(2\ldots)$. They then introduce two different "models" in which we might try to formalize these situations. Model A deals with the NAP probability space $(\Omega_A = \{0,1\}^{\mathbb{N}}, \mathscr{P}(\Omega_A), P_N)$. Model B deals with the exact same NAP probability space, except that, "[...] this set [the sample space] is now used in a different way, namely, to reflect that the count of events starts at the first toss of $H(2\ldots)$" (*ibid.*, 529). The point is, then, that in order to obtain $P(H(1\ldots)) = P(H(2\ldots))$, Williamson exploits the intuition that $P_A(H(1\ldots)) = P_B(H(2\ldots))$. This amounts to comparing probabilities across different models in which the sample space is interpreted differently. And, as the authors point out, the NAP probability functions depend very delicately upon the sample space and how events are "embedded" in the sample space. As such, we cannot derive $P(H(1\ldots)) = P(H(2\ldots))$ from $P_A(H(1\ldots)) = P_B(H(2\ldots))$, and so Williamson's argument against infinitesimals does not succeed.

This is an interesting idea, but it suffers from a few issues. First, it would seem that the Bernoulli system provides a univocal model of Williamson's experiment. Still, Benci et. al. could object to the use of this model in the following way. NAP probability assignments reflect the ratio between the numerosity of an event E and the numerosity of all cases, provided that all cases are equiprobable. However, Parker has shown that numerosities (and, he claims, any theory of infinite counting satisfying **PW**) are not translation/shift invariant ([Parker, 2013]). And so, though Bernoulli systems are measure-preserving systems, they have been formulated in a classical Cantorian framework that does not preserve NAP probability assignments. Williamson could concede this point, but then appeal to a recent paper in which a non-Archimedean probability space is constructed that models infinite sequences of coin tosses ([Benci et al., 2015]). This model *does* preserve the relevant probability assignments, viz., the probability of event E is the ratio between the numerosity of E and the numerosity of the sample space. Furthermore, it is shown that the standard part of the NAP probability assignment for any E agrees with the usual measure-theoretic probability assignment. Thus, Williamson does have a univocal model available for his analysis.[72]

Second, in another iteration of his experiment, Williamson *does* distinguish what Benci et. al. call Models A and B. He says,

> [...] [S]uppose that another fair coin, qualitatively identical with the first, will also be tossed infinitely many times at one second intervals, starting at the same time as the second toss of the first coin, all tosses being independent. Let $H^*(1\ldots)$ be the event that every toss of the second coin comes up heads, and $H^*(2\ldots)$ the event that every toss after the first of the second coin comes up heads. [...] These two infinite sequences of tosses proceed in parallel, synchronically, and there is no qualitative difference between the coins ([Williamson, 2007], 175).

Diagrammatically, with T_i being the toss with time index i, the entire experiment is given by the following arrays. The first array represents the first coin and the second array

[72] For the original discussion of the relationship between numerosities and measures, see [Benci et al., 2014]. See Mancosu's forthcoming book, *The Wilderness of Infinity: Robert Grosseteste, William of Auvergne and mathematical infinity in the thirteenth century*, for an application of these ideas to mathematico-philosophical claims about the infinite from the 13th century.

represents the second coin:

$$\begin{array}{ccccc} T_1 & T_2 & T_3 & T_4 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \quad (\text{Event } H(1\ldots)) \\ & 1 & 1 & 1 & \cdots \quad (\text{Event } H(2\ldots)) \end{array}$$

$$\begin{array}{ccccc} T_1 & T_2 & T_3 & \cdots & \\ 1 & 1 & 1 & \cdots & (\text{Event } H^*(1\ldots)) \\ & 1 & 1 & \cdots & (\text{Event } H^*(2\ldots)) \end{array} \tag{4.3.26}$$

Once more, Williamson is trying to derive $P(H(1\ldots)) = P(H(2\ldots))$ using an isomorphism claim. He argues as follows: $H(1\ldots)$ and $H^*(1\ldots)$ are isomorphic events via 1-1 correspondence (implicit premise; preference for **CP**). Then $P(H(1\ldots)) = P(H^*(1\ldots))$ because "the probability that a coin comes up heads on every toss does not depend on when one starts tossing [...]" (Label Invariance) and the coins are qualitatively identical. By the same argument, $H^*(1\ldots)$ and $H(2\ldots)$ are isomorphic (implicit premise) and thus $P(H^*(1\ldots)) = P(H(2\ldots))$. Thus, by transitivity $P(H(1\ldots)) = P(H(2\ldots))$, and the argument is off and running.

Evidently, the array describing the first coin is the author's Model A and the array describing the second coin is their Model B. In particular, $P_A(H(1\ldots)) = P(H(1\ldots))$ and $P_B(H(2\ldots)) = P(H^*(1\ldots))$. Williamson has, however, made the appropriate distinctions, and so is not conflating different models of the experiment. Rather, as we saw, he obtains $P(H(1\ldots)) = P(H^*(1\ldots))$ (Benci et. al. 's $P_A(H(1\ldots)) = P_B(H(2\ldots))$) using the same isomorphism claim (and underlying intuitions of **CP** and Label Invariance) as in the first iteration of the experiment. Thus, once more, the crucial thing to understand is the isomorphism and the information it preserves.

In sum, the analysis given in [Benci et al., 2018] is not an appropriate way to understand Williamson's argument and its endemic problems. Nonetheless, I think their discussion is extremely useful in its emphasis on the fact that different representations of probabilistic situations and different interpretations of the very same formalism may disagree in subtle ways. This idea will feature significantly below.

4.4 Bertrand's Paradox

4.4.1 Introduction

Bertrand's paradox, originally proposed by J. Bertrand in [Bertrand, 1889], has been a central concern in discussions of classical probability theory and the Principle of Indifference.[73] It has received—and continues to receive—a great deal of attention from philosophers, mathematicians, and physicists alike.[74] Nonetheless, I believe that all extant descriptions of (and

[73] In essence, the Principle says that the possibilities of which we have equal ignorance should be assigned equal probabilities.

[74] See, for instance, [Keynes, 1921], [Jaynes, 1973], [Marinoff, 1994], [Shackel, 2007], [Bangu, 2010], [Rowbottom, 2013], [Klyve, 2013], [Aerts and de Bianchi, 2014], [Gyenis and Rédei, 2015], [Rizza, 2018], [Shackel and Rowbottom, 2020].

proposed solutions to) the paradox fail to recognize the issues that constitute its heart. It is, in its original formulation at least, a paradox of infinity produced by a conflict between our techniques for measuring the infinite and the properties we wish to preserve when doing so. In particular, an uncritical reliance upon **CP** helps to produce the paradox. Indeed, it remains so thorny and contested precisely because the underlying foundational context has remained unexamined. Analyses and resolutions have been given in terms of very ingenious constructions internal to measure-theoretic probability theory.[75] However, for all their ingenuity, these miss the point entirely.

I aim to show that Bertrand's paradox is generated by the fact that the relative sizes of infinite sets are not preserved under bijections and thus by a genuine conflict between **PW** and **CP**. In particular, the relative size of $S \subset X$, e.g., "S is precisely one-fourth the 'size' of X," depends crucially upon (i) how X and S are represented and (ii) our ambient measure of infinity. And this has nothing to do with probability theory *per se*. Furthermore, in light of the conflict between **PW** and **CP**, Bertrand's paradox is a genuine paradox, contrary to many recent analyses.

I begin by examining Bertrand's original text because this helps to make clear the central role of infinitary considerations at its essence. I will then provide a more streamlined, quasi-formal description of the paradox and demonstrate how our concerns with representation and infinity appear. Finally, I compare my findings with some recent analyses of the paradox that I find particularly insightful ([Gyenis and Rédei, 2015] and [Rizza, 2018]). I believe that there are a number of interesting points of agreement; however, the authors often complicate their analyses unnecessarily with their emphasis upon the details of classical probability theory and erroneously claim that there is no real paradox present.

4.4.2 Bertrand's Question and Three Procedures

Bertrand's original formulation of the paradox is geometric in nature. This is the formulation that I will follow throughout the paper; however, it is important to note that other statements have been given in quite different dress.[76] It would be interesting to see to what extent my analysis could be applied to these, but I will set this aside for future work.

Bertrand begins with a preliminary section on the infinite that is worth quoting:[77]

> A remark is now necessary: the infinite is not a number; we must not, without explanation, introduce it into our reasoning. The illusory precision of words can produce contradictions. To choose *at random* from an infinite number of possible cases is not a sufficient indication of what to do ([Bertrand, 1889], §4).

After presenting a simple number-theoretic paradox, Bertrand says, with tongue-in-cheek, "Contradictions of this sort can be multiplied *ad infinitum*." These remarks are interesting because they indicate that Bertrand himself quite clearly conceived of the central issue being the indeterminate nature of the infinite and not anything about the nature of probability.

[75] [Jaynes, 1973] and [Marinoff, 1994] are perhaps the best known possible resolutions. [Shackel, 2007] and [Gyenis and Rédei, 2015] are more recent discussions in this vein.

[76] See, for example, the "Water and Wine" paradox in [von Mises, 1957] and the "Cube Factory" in [van Fraassen, 1989]. See [Mikkelson, 2004] for a proposed resolution of the former.

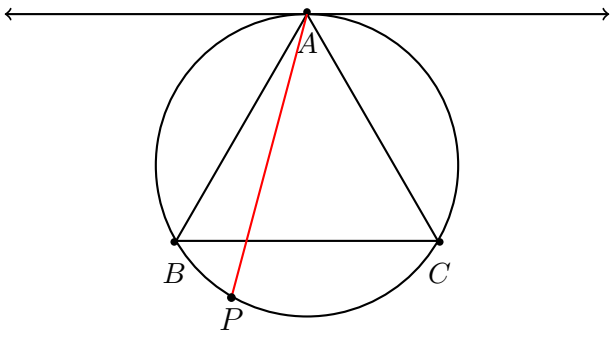
[77] All translations are my own from the French of [Bertrand, 1889]. See Appendix H.

Bertrand's discussion of the eponymous paradox starts by introducing a seemingly well-formed question:

> One draws, at random, a chord in a circle. What is the probability that [the chord] is shorter than the side of an inscribed equilateral triangle? ([Bertrand, 1889], §5)

He then provides three procedures by which the probabilities might be found. Unfortunately, each procedure generates a distinct value, only one of which, it seems, can be correct. Since the procedures actually focus on chords *longer* than the sides of the inscribed triangle,[78] I adopt the traditional terminology and call "Bertrand's question" the following: *What is the probability that the chord is longer than the side of an inscribed equilateral triangle?* In order to answer this question, note that there is only one equilateral triangle (up to rotations) that can be inscribed in a circle of radius r. By elementary geometry, the length of each side of the triangle is precisely $\sqrt{3}r$. Here, then, are modernized, though quite faithful, versions of Bertrand's three procedures:[79]

Procedure I: Label a point A on the circumference of the circle. Inscribe an equilateral triangle with one of its vertices at A and others at B and C. Fix A as an endpoint of the chords we will construct. Any such chord is then completely determined by the specification of its other endpoint and thus by the specification of the angle θ formed between the line tangent to A and the chord itself. Since each angle of an equilateral triangle is $60°$, a chord $\overline{AP}$ will be longer than a side of ΔABC iff $60° < \theta < 120°$. Since θ can take values in $(0°, 180°)$,[80] one-third of all possible angles will produce a chord $\overline{AP}$ (one possible construction drawn below) with length greater than $\sqrt{3}r$. Thus, the probability of drawing a chord longer than the side of an inscribed equilateral triangle is precisely $\frac{1}{3}$.



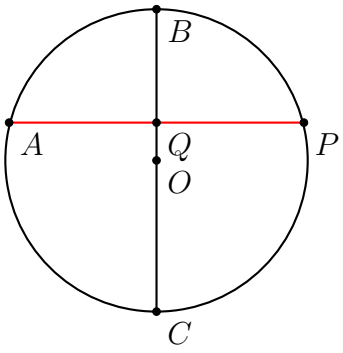
Procedure II: Let chord $\overline{AP}$ be given. Let $\overline{BC}$ be the diameter of the circle perpendicular to $\overline{AP}$. This diameter is composed of two radii $\overline{BO}$ and $\overline{OC}$ of length r with the O the center of the circle. Now consider the length of $\overline{QO}$, written as $\ell(\overline{QO})$, with Q the point

[78] Bertrand obviously required that we perform a final step in subtracting the probability of drawing a longer chord from unity to obtain the probability of drawing either a shorter chord or a chord of equal length. Given this set-up, we would then also have to compute the probability of selecting a chord at random that is precisely equal to the side of the inscribed equilateral triangle.

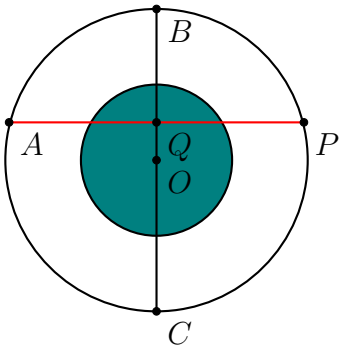
[79] Again, the reader is invited to compare with my translation in Appendix H.

[80] I exclude the endpoints because $\theta = 0°, 180°$ would produce tangents, not chords.

of intersection of $\overline{AP}$ and $\overline{BO}$ (w.l.o.g. $\overline{OC}$). If $0 \leq \ell(\overline{QO}) < \frac{r}{2}$, then $\overline{AP}$ is longer than the side of an inscribed equilateral triangle.[81] Finally, since $\overline{QO}$ can take values in $[0, r)$, we see that the above range comprises precisely half the possibilities. Thus, the probability of drawing a chord longer than $\sqrt{3}r$ is precisely $\frac{1}{2}$.



Procedure III: As in Procedure II, we note that an arbitrarily selected chord $\overline{AP}$ will have length greater than $\sqrt{3}r$ iff the distance from O to its midpoint Q is less than $\frac{r}{2}$. We now think about what this means in terms of areas of concentric circles. A smaller circle of radius $\frac{r}{2}$ (shaded in teal below) will have area $\frac{\pi r^2}{4}$, which is $\frac{1}{4}$ the area of the original circle of radius r. Any point inside the circle will be the midpoint of some chord, and our desired chords, by the above reasoning, will have midpoints inside the circle of radius $\frac{r}{2}$. Thus, the probability of drawing a chord longer than $\sqrt{3}r$ is precisely the ratio of the area of the smaller circle to the area of the original, i.e., $\frac{1}{4}$.



After presenting these procedures, Bertrand suggests that none of the putative answers is either correct or incorrect but rather that the "question is ill-posed." This response can be read, especially in light of his remarks in §4, as a preference for finitism. This reading is supported by the description of §4 in the table of contents: "The number of cases cannot be infinite. Contradiction resulting from forgetting this condition" (*Le nombre des cas ne doit*

[81] This is shown as follows. Inscribe equilateral triangle ΔXYZ in a circle of radius r with center O. The sides of this triangle have length $\sqrt{3}r$. Let M be the midpoint of side $\overline{XY}$. Then the length of $\overline{MY}$ is $\frac{\sqrt{3}}{2}r$. We also know that the length of $\overline{OY}$ is r, and so the length of $\overline{OM}$ is $\frac{r}{2}$. Now construct a radius $\overline{OR}$ through M. Let our chord $\overline{AP}$ intersect $\overline{OR}$ at Q. We then see that $\overline{AP}$ will be longer than $\sqrt{3}r$ provided thar $0 \leq \ell(\overline{QO}) < \frac{r}{2}$.

pas être infini. Contradiction résultant de l'oubli de cette condition). Thus, Bertrand's preferred response to his paradox is a probabilistic finitism: in order for probability assignments to be well-defined, the number of cases must be finite. The draw of probabilistic finitism did not lose its luster in the subsequent century. As noted above, de Finetti proposed that Kolmogorov's postulate of Countable Additivity be rejected in favor of Finite Additivity.[82] This latter connection to Bertrand has obviously been noted[83] but has served to obscure the significance of Bertrand's paradox as a paradox of infinity and not expressly of probability.

In any case, though perfectly coherent, finitism is a restrictive position and should only be accepted as a last resort. I believe my analysis can make good sense of the paradox without dispensing with the use of infinitary resources.

4.4.3 My Analysis

I will first provide a quasi-formal framework in which all three procedures can be described. This will help to make the discussion more uniform. Let X be the set of all chords that can be drawn in a circle C of radius r. This set is uncountably infinite. Let S be the proper subset of X that has chords longer than $\sqrt{3}r$ as its elements. Given this information, what Bertrand in effect does is produce three different representations of S and X, with each representation corresponding to the procedures described above. Intuitively, each representation should be constituted by an ambient representation space, a bijection from X to this representation space, and the image of S under the bijection. Unfortunately, as we will see, there are a number of complications that arise for each representation. In particular, for each case, we only obtain a bijection by imposing further conditions on the representations, some of which are quite *ad hoc*. I will note these in turn.

Representation I: Let our ambient representation space be given by

$$R_1 := \{\text{the set of angles in the interval } (0^\circ, 180^\circ)\}\,. \tag{4.4.1}$$

Let our bijection be $\varphi_1 : X \to R_1$ and let $\varphi_1(S) \subset R_1$, the image of S under φ_1, be the set of angles taking values in $(60^\circ, 120^\circ)$. The size of $\varphi_1(S)$ should be one-third the size of R_1.

Remark 4.4.1. Unfortunately, this cannot be quite right. By selecting a point A as the endpoint of the chords to be chosen, we have implicitly restricted our attention to a proper subset of X. We have, in fact, produced a partition of $X := \bigcup_t P_t$ with each part determined by parameter $t \in [0, 2\pi r)$. Given a choice of endpoint A, we then deal with only one of the parts. Thus, we have a family of bijections $\{\varphi_1^t\} : P_t \subset X \to R_1$, not simply $\varphi_1 : X \to R_1$.[84]

Even worse, this partition will only be well-defined if we exclude diameters. Consider a diameter with A as one of its endpoints. Let this be in partition P_s. Now apply ρ, where ρ is a 180° rotation. When dealing with a chord that is not a diameter, this ρ will produce a new chord. However, in the case of our diameter with endpoint A, when we apply ρ we will double-count this diameter, assigning it to both P_s and some distinct P_q.

[82] See, again, [de Finetti, 1974].
[83] See [Shackel, 2007], p. 156.
[84] See Section 4.4.3 below for a discussion of whether this poses problems for Bertrand.

Representation II: Let our ambient representation space be given by

$$R_2 := \{\text{the set of lengths in the interval } [0, r)\}\,. \tag{4.4.2}$$

Let our bijection be $\varphi_2 : X \to R_2$. Finally, let $\varphi_2(S) \subset R_2$ be the image of S under φ_2, which is given by the set of lengths in $[0, \frac{r}{2})$. By the reasoning in Procedure II, the "size" of $\varphi_2(S)$ should be one-half the size of R_2.

Remark 4.4.2. As above, there is a further complication. Bertrand has again produced a partition of X, where each P_t is determined by the initial chord drawn and the radius perpendicular to it. Two chords will lie in the same partition iff they are parallel and their midpoints lie on the same radius. Thus, we really have a family of bijections $\{\varphi_2^t\}$ from $P_t \subset X \to R_2$.

Representation III: Let our ambient representation space be given by[85]

$$R_3 := \{\text{the set of interior points of the circle of radius } r\}\,. \tag{4.4.3}$$

Let our bijection be $\varphi_3 : X \to R_3$. Finally, let $\varphi_3(S) \subset R_3$ be the set of points in the interior of the circle centered at O with radius $\frac{r}{2}$. By the reasoning in Procedure III, the "size" of $\varphi_3(S)$ should be one-fourth the size of R_3.

Remark 4.4.3. Unfortunately, as in Representation I, we have an issue with the diameters. In particular, though any point inside the circle *besides* the center O is the midpoint of a unique chord, O serves as the midpoint for uncountably many chords (the diameters). Thus, φ_3 will only be a bijection if we exclude the center O as a midpoint of a chord.[86]

Parts, Wholes, and Symmetry

Thus, we see that a careful analysis of the paradox is rather more complicated than many commentators have realized. In particular, a common theme that emerges from my remarks is that, even before we pass to the representation spaces, Bertrand deals not with X directly but rather with a *proper subset* of X. In Representations I and II, we deal with a part P_s of one of the partitions, and in Representation III we deal with the subset of X that excludes O as a midpoint determining a chord.[87] Bertrand proceeds to write down bijections between these subsets and the representation spaces. But why think that representing these subsets of X via one of the R_i will yield the relative size information we desire, viz., the relative size of S in X? Some commentators have argued that there is no issue here precisely because the sets we consider are all (uncountably) infinite. Jaynes, for instance, suggests that considering just one set of possible chords (e.g., one of the partitions in Procedure I) is sufficient to produce the same answer as would have been produced had we considered the entire set of chords. This is problematic. While it is true that a 1-1 correspondence between

[85] I take it that we would exclude the points on the circumference of the circle from being midpoints of chords since a "chord" in this case would just be a point.

[86] This is noted in [Shackel, 2007].

[87] This is noted by [Shackel, 2007] (pp. 156ff.) and [Rowbottom, 2013].

infinite sets is a necessary condition on the adequacy of our representation spaces, it cannot be sufficient.

In order to infer the relative size of S in X by restricting our analyses to those proper subsets of X mentioned above, we would have to show that each subset is representative of or "essentially the same as" the entire set of chords. Perhaps this is so in light of the fact that the partitions in Representations I and II will form a group under the symmetries of the circle. For instance, we can produce distinct parts P_s and P_q in Representation II by rotating an initially given chord (an element of P_s) by angle θ and then constructing the diameter perpendicular to this rotated chord. Then, all chords parallel to the rotated chord and with midpoints on the new diameter will be in P_q. A similar idea holds for Representation I. It is plausible that these symmetry considerations will be sufficient.[88]

For the time being, I think we can proceed under the assumption that each part (of one of the partitions) is an adequate representative for X as long as we recognize that this does not follow from the fact that the sets are all uncountably infinite. The adequacy of the parts follows, rather, from symmetry considerations. Thus, for the sake of simplicity, I assume that we have legitimate bijections $\varphi_i : X \to R_i$ instead of the more cumbersome families of bijections described above. It is also the case that the conflicting intuitions driving the paradox appear regardless of the precise details of the domain and co-domain of the mappings we consider. It is, however, important to clarify what Bertrand is doing in his initial set-up; recognizing that he deals with "representative" parts of X is crucial for this. However, for my ultimate claims about the conflict between **PW** and **CP**, putting things in terms of partitions would only serve to obfuscate matters.

The Crux of the Paradox

We are now in a position to describe how the paradox arises. First, note that our original sets X and S, though perfectly well-defined, do not come equipped with an obvious way to measure their relative sizes. All we know, *prima facie*, is that X and S are uncountably infinite. We do not have enough information to determine the size of S relative to X. Thus, we must either represent our sets more determinately or pass to a measure of infinity finer than cardinalities.

The second thing to notice is that each representation is judged to be acceptable, at least initially, because we are able to write down a bijection $\varphi_i : X \to R_i$ for $i = 1, 2, 3$. It is the case that each representation space is acceptable[89] but *only insofar as there is a 1-1 correspondence between X and R_i*. We should not necessarily expect that other properties we are interested in will be adequately represented. This is a lesson we have now encountered repeatedly: if our bijection is merely a Cantorian bijection, then we should not expect it to preserve more nuanced information. Indeed, information like part-whole relations will depend upon how our mathematical entities are presented. In this case, the part-whole relationships will depend upon how the representation spaces are given and the "natural"

[88] Bertrand himself may have been thinking along these lines. See his remarks in Appendix H concerning "the symmetry of the circle." See also [Klyve, 2013]. Rizza believes that we cannot decide as to whether Bertrand in fact considers the whole ensemble of chords or merely a proper part without first providing a "mathematical determination" ([Rizza, 2018], 379).

[89] Keeping in mind the caveats about representative parts.

measure carried by each space.

Indeed, Bertrand uses R_1, R_2, R_3 because, unlike X, they come equipped with intuitive measures: ranges of angle measures,[90] interval length, and areas, respectively. These intuitive measures are required for us to judge the size relations between the R_i and $\varphi_i(S)$; however, there is no *prima facie* reason to think that the size relations will remain invariant between the representation spaces.[91] Indeed, as we have now seen, the size relations change as we deal with different representation spaces: for example, $\varphi_1(S)$ is one-third of R_1, whereas $\varphi_2(S)$ is one-half of R_2. But why think in the first place that the relative size of $\varphi_i(S)$ in R_i would be the same for all i? Once more, I believe that it is our coarse, Cantorian intuitions that have led us astray. Indeed, note that we can write down bijections between the R_i by composing the φ_i and their inverses. For example, we can construct a bijection $\varphi_2 \circ \varphi_1^{-1} : R_1 \to R_2$. In virtue of these bijections, we are inclined to think that the size relations between the R_i and their subsets $\varphi_i(S)$ will remain invariant. But this is not the case: each representation space comes equipped with its own intuitive measure of set size, and these produce different relative size assignments.

Thus, we can dissolve the paradox by jettisoning our Cantorian intuitions about infinite sets. Once we notice that the relative size assignments of our sets will depend upon *how* we choose to represent our original set X, we must simply stick to one representation. This shows that there is some truth to what Bertrand says: perhaps there is no univocal "correct" answer to his question. This is, however, part and parcel of what we have seen in our discussion of Label Invariance and numerosities. Once we pass to measuring infinite collections, there will always be some indeterminacy present.

This account of Bertrand's paradox has some affinities with the description of Williamson's infinite sequence of heads provided in [Benci et al., 2018]. The authors placed great emphasis on the fact that Williamson implicitly switched between representations of an "intuitive" situation, and it was this that produced the appearance of paradox. We have found something similar; however, I would like to emphasize that the root of the confusion in Bertrand's paradox is not merely that we implicitly move between representations. It is our predilection for **CP**. This is because we are inclined to accept each model precisely because we can write down bijections φ_i and $\varphi_j \circ \varphi_i^{-1}$ $(i \neq j)$ and we expect that these will preserve the information we care about. Unfortunately, when dealing with the infinite, the situation is rather more nuanced.[92]

In sum, we have been able to express the fundamental conflicting intuitions at play in Bertrand's Paradox without making use of probability theory. I have presented the paradox in the following way: given the set of all chords X in some circle C or radius r, what is the relative size of S, the set of chords with length greater than $\sqrt{3}r$? If one wishes to

[90] It seems that the intuitive measures on R_1 could be variously described. We might also try to use arc length on the circumference.

[91] Nor will appeal to a "standard" measure across the R_i help. Since we are dealing with intervals and regions in $\mathbb{R}$ and $\mathbb{R}^2$, we might immediately try to use Lebesgue measure. This will not work. For instance, let's try to use Lebesgue measure μ in Representation I. Intuitively, we wish to show that arc BC is one-third the circumference of the circle. However, for the unit circle S^1, $\mu(S^1) = 0$, since $S^1 \subset \mathbb{R}^2$ and the Lebesgue measure of any entity of dimension strictly less than the ambient space is always zero.

[92] There are points at which [Shackel, 2007] seems to circle around this idea, but I am not sure if this is what he had in mind.

honor the letter of Bertrand's question, we can proceed from this to an intuitive probability assignment by invoking the basic principle that *The probability of some event E is the ratio of the "number" of elementary events in E to the total "number" of elements in our sample space* Ω. Obviously, E is the event that we draw a chord from S and $\Omega = X$. Thus, the probability of our requested event should just be the ratio of the size of S to the size of X. But, of course, producing such a probability assignment requires that we have a determinate way to measure the sizes of S and X. And this is what we lack. Thus, one pillar of Bertrand's paradox is this: how to measure the sizes of infinite sets such that we preserve their relative sizes.

Unfortunately, as we saw, there is no obvious way to do this for our X and S. And this fact introduces further complications involving representations. In order to produce tractable size assignments, we passed to new representation spaces equipped with their own intuitive measures. These measures produced different size assignments (for $\varphi_i(S) \subset R_i$), which led to the paradox because we expected these relative sizes to be preserved under bijections.

Now yet another problem emerges. Let us simply pick a representation space. We wish to lift the relative size of a subset of this space back to X and thereby determine the relative size of S in X. However, why think that our measurement of the relative size of a subset of the representation space will produce values that agree with the "correct" relative size of S in X (assuming there is such)? This becomes especially pressing if we no longer expect relative sizes to be invariant under bijections. Let us make this more concrete. I will use Representation II, since we have genuine bijections from partitions of X to R_2 (that is, we do not have the issues with diameters present in Representations I and III). Let us work with $\varphi_2^t : P_t \subset X \to R_2$. We then see that $\varphi_2^t(P_t) = \left\{\text{the set of lengths in } [0, \frac{r}{2})\right\}$. The "intuitive" measure on R_2 is simply interval length.[93] This tells us that half the chords we can draw through $[0, r)$ are longer than $\sqrt{3}r$. But why now expect that this measure of the size of $\varphi_2^t(P_t)$ in R_2 can be lifted back to S and X thereby yielding a probability assignment of $\frac{1}{2}$ for our desired event? There is nothing in the mathematics of the problem that guarantees this can be done. Indeed, our above discussions should incline us to think otherwise: the size of infinite sets are highly sensitive to presentation when we wish to ascertain part-whole relations. Nonetheless, a precise answer to this question depends, ultimately, on whether we can specify the sizes of X and S independently of Bertrand's procedures. Let us now turn to this possibility.

4.4.4 Bertrand's Paradox and the Problem of Mathematical Determination

In a recent paper [Rizza, 2018], Rizza provides a novel analysis of Bertrand's paradox that can be fruitfully compared with my own account. His central contention is that the paradox fundamentally expresses a "determination problem." That is, the paradox (disagreement of probability values) is generated because Bertrand did not have the mathematical tools available to specify the size of infinite collections of geometrical objects, viz., the size of the set of all chords in a circle of radius r. Thus, once we provide a numerical specification for

[93] If one wishes to use a more general measure, the Lebesgue measure μ will agree with interval length, as it should. Indeed, $\mu([0, r/2)) = \mu([0, r/2])$ since $[0, r/2] = [0, r/2) \cup \{r/2\}$ and $\mu(\text{point}) = 0$. Then, $\mu([0, r/2]) = r/2 - 0 = r/2$. Finally, by the same reasoning, $\mu([0, r)) = r$.

the infinite collections involved, we can judge the adequacy of Bertrand's drawing procedures and dissolve the paradox. I will briefly describe Rizza's findings and then compare them to my own analysis.

After a discussion of a recent debate concerning Bertrand's paradox ([Rowbottom, 2013], [Klyve, 2013]), which Rizza takes to support his main thesis, he proceeds to extend the existing apparatus of probability theory with a new computational theory: Sergeyev's *Gross-one*.[94] The basic idea here is to operate with an enriched numeral system, viz., one employing an infinite base η (read "Gross-one"). η is then supposed to refer to an infinitely large integer, greater than any natural number represented as a finite-base numeral. One of the advantages of this theory is that we can obtain numerical measures of infinite collections for which the usual arithmetical properties hold. A second advantage is that we are able to numerically discriminate between the size of an infinite collection and its infinite subcollections, i.e., **PW** holds. As Rizza notes, these properties are also had by numerosities; however, it is worth remarking that both in terms of their aims and mathematical constructions, the theory of Gross-one and the theory of numerosities are rather different.

In any case, using η, Rizza provides a parameterization of the chords in a circle of radius r. We have already seen that the specification of a chord will be given by the specification of two points on the circumference of the circle. This suggests that we label η many points on the circumference and compute the number of chords from there. Fixing an endpoint, say A in Procedure I above, we find that there are $\eta - 1$ discriminable chords through A. As our fixed endpoint ranges throughout the η many points on the circumference, we find that there are $\eta(\eta - 1)$ total chords. However, we have double-counted, since distinct orderings (e.g., $\overline{AB}$ and $\overline{BA}$) of the labelled endpoints have been considered distinct chords. Thus, the total number of chords (with X again denoting the total collection of chords) is the infinitely large integer

$$|X| = \frac{\eta^2 - \eta}{2}. \tag{4.4.4}$$

Obviously, this "determination" was unavailable to Bertrand. It easily follows that we can introduce a uniform, discrete probability distribution on the sample space X of chords with the above size. The probability of drawing a given chord is then $\frac{2}{(\eta^2-\eta)}$. Using these values, we can compute the probabilities of drawing a chord equal to $\sqrt{3}r$ $(P(e))$, shorter than $\sqrt{3}r$ $(P(s))$, and longer than $\sqrt{3}r$ $(P(l))$.[95] Given these values of $P(e), P(s), P(l)$, we can then judge the adequacy of Bertrand's proposed drawing procedures. Rizza shows that the finite values of these probabilities all agree (again given our specification of η many points on the circumference), though they do differ by infinitesimal amounts.[96] Rizza concludes

> [...] when such homogeneous distribution is fixed as the geometrical configuration of reference, the drawing methods proposed by Bertrand are in finite agreement and only generate infinitely small discrepancies ([Rizza, 2018], 391).

Remark 4.4.4. It is worth noting that the finite values obtained are *not* those computed

[94] See [Sergeyev, 2003], [Sergeyev, 2009a], [Sergeyev, 2009b].
[95] See [Rizza, 2018], pp. 385-6 for the computations.
[96] See [Rizza, 2018], pp. 386-391 for computations.

by Bertrand's original procedures with the exception of Procedure I. This is because, as Rizza's claims, only Procedure I considers a subset $S \subset X$ that is a scale model (or, in my terminology, "representative") of the total space. However, there is room to disagree here as the way Rizza construes Procedure I is not obviously equivalent to the way Bertrand construes it. In short, it is somewhat unclear whether Rizza is dealing with the same issues that Bertrand is; I will discuss this further below.

In any case, the above conclusion at least is unimpeachable; it follows directly from the mathematics of the situation. It is nonetheless worth asking: Is this really a solution to Bertrand's paradox? I would like to argue that our answer to this question should be no. Rizza is absolutely correct in his assertion that Bertrand lacked a "sufficient mathematical instrumentality" to address his question. However, Bertrand's response to this lack was to provide the various geometric representation spaces (see above), which themselves carry intuitive measures of size. In so doing, he was able to produce numerical values for his desired probabilities. As I have argued, it is precisely this use of different geometric representation spaces, along with our expectation that the relative sizes of infinite sets ought to be preserved under bijections, that produces the paradox. *This* is the paradox, and it remains entirely untouched by Rizza's computations. Indeed, he seems to realize this, when he says

> [...] If, on the other hand, one replaces the geometrical configuration attached to the parameterisation of chords as pairs of labelled endpoints *with other geometrical ensembles* (points on a diameter, interior points), which in turn lead to distinct random selection processes, then probability values proliferate (*ibid.*, 391).

Thus, we must be aware that, in some sense, he has entirely side-stepped what makes the paradox paradoxical. Nonetheless, this does not detract from the value of his contribution to our understanding of the paradox. To conclude, I would like to sketch our points of agreement as these make clear what we have learned from Bertrand's paradox.

First, our accounts agree that Bertrand's paradox is not ultimately about probability theory (the Principle of Indifference, etc.) but is rather about the infinitary resources at our disposal. Rizza stresses that the "canonical instrumentality" of classical probability theory must be supplemented in order to assess Bertrand's procedures. My proposal is somewhat more radical as I think we can make sense of what is at issue entirely in terms of the properties of infinite sets we wish to preserve.

Second, our accounts agree that Cantorian techniques are entirely useless in making sense of the paradox. This is because we need a way to measure infinite collections such that field-theoretic operations hold and such that part-whole relations are preserved. Rizza does not make much of this, but the insufficiency of Cantorian techniques (in virtue of their coarseness) has been a theme of this essay. I take it that it would be entirely possible to achieve a similar sort of parameterization of chords and subsequent computation of probabilities using numerosities, since these satisfy the above properties.[97]

[97] I would like to pursue this in future work following the developments in [Benci et al., 2014] and [Benci et al., 2015].

4.4.5 Defusing Bertrand's Paradox

[Gyenis and Rédei, 2015] provides another interesting and subtle analysis of Bertrand's paradox. The authors' primary contention is that Bertrand's paradox is not, in fact, a genuine "unresolvable" paradox. Rather, it expresses a non-trivial mathematical fact, and this has not been realized because all previous analyses have not been posed at an appropriate level of abstraction. They claim that, once we adequately describe Bertrand's paradox in terms of measure-theoretic probability theory, we will be able to see what is really at stake.[98] Thus, the project is very much in line with Rizza's: once we provide a "sufficient mathematical instrumentality" for the paradox, it will be defused.

Besides this connection to Rizza's work, [Gyenis and Rédei, 2015] is of interest because the authors discuss a particular version of an intuition that has played a role in this chapter: Label Invariance. They show that, in the category of Haar probability measure spaces, their preferred context for describing Bertrand's paradox, Label Invariance does not hold. (What the authors mean by "Label Invariance" will be discussed below.) This is the mathematical fact that the paradox purportedly expresses. Rizza takes this as grist for his own mill, suggesting that the violation of Label Invariance, "[...] may be regarded as a pointer to differences between the probability models that cannot be fully detected by the canonical instrumentality [in this case, measure-theoretic probability]" ([Rizza, 2018], 393). I would like to continue to suggest, however, that Label Invariance, in the context of infinitary probability theory, derives its plausibility in large part from our predilection for Cantorian techniques. Furthermore, against Rizza, I would like to emphasize that what is of philosophical interest here is not the "problem of mathematical determination," but rather, once more, the conflict between different infinitary intuitions and our desire to retain particular information (in the Bertrand case, the relative size of set S in X).

Bertrand's Paradox in the Category of Haar Probability Spaces

Gyenis and Rédei attempt to show that the common construal of the paradox as a violation of the Principle of Indifference is incorrect.[99] Internal to what they call the "classical interpretation of probability theory," their argument shows that

> [...] Bertrand's paradox does not affect the principle of indifference and does not, in and by itself, undermine the classical interpretation of probability; the classical interpretation, the principle of indifference, and labelling invariance are independent ideas ([Gyenis and Rédei, 2015], 351).

By the "classical interpretation," the authors mean Kolmogorov's measure-theoretic development of probability theory along with a "link" between this formalism and the world. That is (following their notation), we deal with probability spaces $(X, \mathcal{S}, p)$ where X is the outcome space of elementary random events, $\mathcal{S}$ is a Boolean σ-algebra of subsets of X, and $p : \mathcal{S} \to [0, 1]$ is a countably additive measure satisfying $p(X) = 1$. In addition, we deal

[98] As the authors note, this is also done in [Shackel, 2007]; however, Shackel's discussion is posed at too high a level of abstraction and is thus not sufficiently determinate.

[99] Nonetheless, in order to discharge this claim, they must provide an adequate interpretation of the Principle of Indifference and the paradox itself. See below.

with (here real-valued) random variables, viz., measurable functions $f : X \to \mathbb{R}$, where measurability is the condition that $f^{-1}(B) \in \mathcal{S}$ for any Borel set B of $\mathbb{R}$.

To their credit, the authors note that this understanding of probability theory in conjunction with the Kolmogorov axioms is not the only way to approach probabilistic phenomena.[100] However, measure-theoretic probability is the most widely-used and important approach, and so they take this as an appropriate context to describe Bertrand's paradox. I think this is uncontroversial, but the alternative possibilities raise interesting questions about the generality of their solution. In particular, I do wonder about their claim that

> The significance of probability theory being *part of* measure theory is that foundational-conceptual problems of probability theory, such as Bertrand's paradox, can *best* be analysed in terms of measure theoretic concepts (*ibid.*, 352; emphasis my own).

This seems too strong.[101] I will return to this below; however, it is certainly true that framing Bertrand's paradox in terms of measure theory is a reasonable approach.

With this formalism in place, the authors state the "elementary classical interpretation," viz., the classical interpretation for finite X:

> In the case of a finite number of elementary events, the probabilities of events are given by the measure p_u that is uniform on the set of elementary events, and (frequency link:) the numbers $p_u(A)$ will be (approximately) equal to the relative frequency of A occurring in a series of trials producing elementary random events from X (*ibid.*, 353).

However, as the authors note, this interpretation is not "maintainable" because only in special circumstances will $p_u(A)$ be indicative of the frequencies with which A occurs. These special circumstances are codified by the Principle of Indifference. In short, this principle must express some sort of epistemic neutrality with respect to the elementary events in order for the classical interpretation to work. The authors' way of describing this epistemic neutrality is via permutations of X: given Π_n, the group of permutations on set X with n elements, we are epistemically neutral with respect to the elementary events if, for all $\pi \in \Pi_n$

$$p_u(\{x_i\}) = p_u(\{x_{\pi(i)}\}) \tag{4.4.5}$$

with $i \in \{1, \ldots, n\}$. Clearly, this is a Label Invariance condition for finite sample spaces.

The authors now claim that, in order to provide an adequate interpretation of Bertrand's paradox, we must formulate a Principle of Indifference for infinite sample spaces. That is, we need to find a permutation group with respect to which we can express a Label Invariance condition (for some heretofore undetermined measure). This requires imposing

[100] They mention different axiomatizations by Rényi, Popper, and Keynes.

[101] Indeed, compare to the more cautious claim found in Folland's discussion of probability theory, "Although measure spaces are a natural setting for the study of probability, it is hardly accurate to say that probability is a branch of measure theory, for its central ideas and many of its techniques are distinctively its own" ([Folland, 1999], 313).

further structure on X. Begin by replacing Π_n with some group G equipped with a group-action $\alpha_g : X \to X$.[102] Our desired Label Invariance condition can then be expressed as: for all $g \in G$ and all $A \in \mathcal{S}$

$$p(A) = p(\alpha_g(A)), \tag{4.4.6}$$

i.e., the probability measure is invariant under the action of G. Unfortunately, for a general probability space $(X, \mathcal{S})$ with X uncountable, there is no guarantee that we can find such a G leading to a unique probability measure p. Nonetheless, we *can* find such a G and p if we require that X be a topological group[103] satisfying further conditions. In particular, let X be a compact topological group. Then there is a unique (up to scalar multiplication) measure p_H, the *Haar measure*, on the Borel sets of X that is invariant under group action.

In sum, then, the authors claim that we should consider triples $(X, \mathcal{S}, p_H)$ where X is a compact topological group with continuous group action, $\mathcal{S}$ is the Borel σ-algebra on X, and p_H is the Haar measure on $\mathcal{S}$. When we do so, we get a "general classical interpretation" of probability theory along with a General Principle of Indifference, where the latter is given by $p_H(A) = p_H(\alpha_g(A))$ for all $A \in \mathcal{S}$ and $g \in G$.

With this formalism in hand, we can make precise a Label Invariance condition for the "general classical interpretation." Given two probability spaces $(X, \mathcal{S}, p_H)$ and $(X', \mathcal{S}', p'_H)$ used to describe the same phenomena, we call a map $h : X \to X'$ a relabelling if it is a bijection between X and X' and both h and h^{-1} are measurable. That is, for all $A \in \mathcal{S}$, $h(A)$, the image of A under h, is in $\mathcal{S}'$. Similarly, for all $A' \in \mathcal{S}'$, $h^{-1}(A')$ is in $\mathcal{S}$. The measurability of h and h^{-1} ensures that, for any event $A \in \mathcal{S}$ with a probability value, its relabelled version will also receive a probability value. Label Invariance in this context then is the requirement that for all $A \in \mathcal{S}$ and $A' \in \mathcal{S}'$

$$p'_H(h(A)) = p_H(A) \tag{4.4.7}$$
$$p_H(h^{-1}(A')) = p'_H(A'). \tag{4.4.8}$$

That is, the probabilities assigned are invariant under relabellings. Of course, this is just another way of saying that $(X, \mathcal{S}, p_H)$ and $(X', \mathcal{S}', p'_H)$ are isomorphic as probability spaces, written as $(X, \mathcal{S}, p_H) \cong (X', \mathcal{S}', p'_H)$. Gyenis and Rédei, then, have dealt with my complaint about needing to specify what is meant by an "isomorphism."[104] They have very clearly delineated the structure that must be preserved and have thus given us an excellent handle on Label Invariance in their general classical interpretation.

Finally, we can express what the authors call

> **General Bertrand's Paradox**: Let $(X, \mathcal{S}, p_H)$ and $(X', \mathcal{S}', p'_H)$ be probability spaces where X and X' are compact topological groups of uncountable cardinality and p_H, p'_H are Haar measures on the Borel σ-algebras of $\mathcal{S}$ and $\mathcal{S}'$, respectively. Then Label Invariance fails because either (i) there is no relabelling between X

[102] See, e.g., [Lang, 2002], Chapter I.5 for the definition of a group action. Lang defines it in a different (though equivalent) manner.

[103] A topological group is a group G endowed with a topology such that the group operations $G \times G \to G$ given by $(x, y) \mapsto xy$ and $G \to G$ given by $x \mapsto x^{-1}$ are continuous.

[104] See Section 4.3.4.

and X' or because (ii) if there is a relabelling, then another relabelling can be found such that Label Invariance does not hold.

Described in this way, Bertrand's "paradox" is a consequence of the following result[105]

Theorem 4.4.5 (Rudin, 1993)**.** *If G is an infinite compact topological group, with Haar measure m_G, then there is an autohomeomorphism[106] h of G such that*

$$m_G(h(E)) \neq m_G(E) \tag{4.4.9}$$

for some open set $E \subset G$.

Thus, Gyenis and Rédei conclude that, once we have properly formalized the classical interpretation of probability theory and an infinitary Principle of Indifference, Bertrand's Paradox is no paradox at all. It is simply an expression of a non-trivial mathematical fact: the failure of Label Invariance in the category of Haar probability spaces.

Remarks on Formalizations of Bertrand's Paradox

Both [Rizza, 2018] and [Gyenis and Rédei, 2015] propose that the central issue at work in Bertrand's paradox is the lack of an appropriate mathematical framework.[107] Their respective plans of attack are, however, somewhat different. Rizza's is much closer to Bertrand's original discussion insofar as he asks a general (he would say indeterminate) question about the probability of drawing a chord longer than the sides of an inscribed equilateral triangle and proceeds to consider three different representations of this situation. Rizza's idea is that we can make Bertrand's question determinate by providing a more expressive "numerical instrumentality" to the original geometrical set-up, i.e., we label points on the circumference of the circle using Gross-one and determine the size of X (the set of all chords) and the size of S (the set of all desired chords) using this labelling. We then consider how this numerical framework is affected by each of Bertrand's drawing procedures or representations of the original set-up. Rizza is getting at something important, viz., the paradox inextricably involves questions about measuring the size of infinite sets, but, as I have noted above, seems to side-step the paradox entirely.

Gyenis and Rédei, on the other hand, discuss neither Bertrand's original set-up nor his drawing procedures anywhere in their paper. They seem to think these irrelevant in light of the fact that Bertrand's paradox was formulated prior to our best current mathematical

[105]The authors write the result differently and cite both [van Douwen, 1984] and [Rudin, 1993]. More precisely: van Douwen shows that there is an infinite, compact, totally disconnected Hausdorff space X with positive Borel measure μ such that, for all Borel sets $E \subset X$ and all autohomeomorphisms h on X, $\mu(h(E)) = h(E)$. He then asks whether this result might hold when we replace X with an infinite, compact, connected group G and the Borel measure μ with Haar measure m_G. (Van Douwen shows that this is false when G is totally disconnected.) Rudin demonstrates that the answer to van Douwen's question is *always* negative.

[106]An autohomeomorphism h of G is a bijective map $h : G \to G$ such that h and h^{-1} are continuous. Continuous maps are Borel measurable, so autohomeomorphisms are relabellings in the sense of [Gyenis and Rédei, 2015].

[107]Though only Gyenis and Rédei are explicit about this, none of these authors seems to think it a genuine paradox.

probability theory; the "paradox" is an artefact of the historical situation, viz., that Bertrand was writing before the foundations of measure-theoretic probability theory (and further mathematics) were laid. Again, their discussion is very nice, and their points well-taken, but it too seems not to address the paradox. Or, perhaps better, it addresses a particular iteration of the paradox in the context of measure-theoretic probability theory (with the further conditions required above to make sense of Label Invariance in the infinitary context). The effect of their discussion is to provide a mathematical framework in which one *could* describe Bertrand's original proposal and to note very general properties about this framework that would be inherited by the proposal. Furthermore, one must buy at the outset of their discussion that describing Bertrand's paradox in terms of the Principle of Indifference is the right way to go. Obviously, I wish to resist this move. In sum, the authors' approach only works if we are working in the context of measure-theoretic probability and accept that an adequate description of the paradox requires the Principle of Indifference. And as I made clear above, and as Rizza's discussion shows, there are many ways we might describe what Bertrand is doing.

Thus, I think that the possibility of quite different formalizations of Bertrand's paradox suggests that it deals with something more general than either Rizza or Gyenis-Rédei realize. However, in both papers, there are hints at this generality. Rizza's recognition of the need for measures of infinity satisfying **PW** and the usual field operations is one such hint. The hint provided by Gyenis and Rédei arises in their discussion of their precise notion of Label Invariance and a more conceptual desideratum of probability theory, which they call *Label Irrelevance*. In the tightly circumscribed context of Haar probability spaces, Bertrand's paradox expresses the failure of Label Invariance (that is, relabellings need not be isomorphisms of Haar probability spaces). On the other hand, Label Irrelevance says that "[...] the specific way the random events are named is irrelevant from the perspective of the value of their probability" ([Gyenis and Rédei, 2015], 350). Then, the reason that Bertrand's paradox "appears paradoxical" is that we easily conflate Label Invariance with Label Irrelevance, despite the fact that these notions are distinct. They continue,

> Labelling irrelevance is respected in probabilistic modelling perfectly well—but it is respected not by labelling invariance holding true: If the elements in the pair of sets $(X, \mathcal{S})$ label the (elementary, respectively, general) random events of some random phenomenon, then one is free to use another pair of sets $(X', \mathcal{S}')$ to label the events *as long as no random events are lost* in X' and $\mathcal{S}'$, i.e., as long as there is a re-labelling h between X and X'. Labelling irrelevance says that the choice of $(X, \mathcal{S})$ or $(X', \mathcal{S}')$ does not affect the probabilities of the random events, and this is in harmony with the fact that fixing either $(X, \mathcal{S})$ or $(X', \mathcal{S}')$ does not determine any probability measure on either [...] (*ibid.*, 366; my emphasis).

This quote makes clear that the seeds of the paradox are already present in the notion of Label Irrelevance. The key thing here is that, for Label Irrelevance to hold, no random events can be "lost."[108] This is very easy to satisfy in the finite case; however, when we pass to spaces of infinite size, this condition will be determined by the ambient measure of infinity

[108]This comes up again in Gyenis and Rédei's discussion of "descriptive accuracy." I also think the idea here is precisely the same as Bartha's no-collapse condition.

that we use, which is, in turn, determined by the information we wish to preserve in our analysis. In Bertrand's paradox, the information that interests us is the relative size of infinite sets, in particular, how large the set S of chords longer than $\sqrt{3}r$ is inside X, the set of all chords. Given that there is not a canonical way of imposing such a measure on X, Bertrand passed to more tractable geometric representations of the situation. However, we are led into thinking that these representations are adequate given our uncritical reliance upon **CP** and Cantorian cardinalities; that is, we can write down bijections (relabellings) between X and each representation space and between pairs of representation spaces. But these bijections need not preserve relative size assignments. Thus, we have a conflict between **PW** and **CP**, made obscure by the different spaces in question and the imposition of probabilistic structure. Indeed, the probabilistic structure (in this case at least) simply reflects the relative sizes of the infinite sets. (Note that Gyenis and Rédei's Label Irrelevance crops up before probability measures are imposed.) Thus, once more, Bertrand's paradox is a paradox of infinity and this is, in fact, implicitly borne out by the analyses of Rizza and Gyenis-Rédei. However, against both these analyses, we should conclude that Bertrand's paradox *is* a genuine paradox insofar as it is produced by a genuine conflict between **CP** and **PW**.

4.4.6 Conclusion

Bertrand asked *What is the probability that a given chord is longer than the side of an equilateral triangle inscribed in a circle of radius r?* Given that he had no way to determine the size of the set of all chords X and the subset S with which he was concerned, he passed to different geometric representations of the problem. As we saw, there is a real question as to whether any of these representations is adequate in virtue of the fact that Bertrand implicitly restricts our attention to proper subsets of X (the partitions discussed above). Nonetheless, in virtue of symmetry considerations, I think there is a reasonable argument that these subsets can play the role of X. As such, we get bijections $\varphi_i : X \to R_i$ for representation spaces R_i $(i = 1, 2, 3)$ as well as bijections $\varphi_j \circ \varphi_i^{-1} : R_i \to R_j$ $(i \neq j)$. And now the paradox emerges. Each space R_i is equipped with its own intuitive geometric measure that yields a particular part-whole relationship between itself (representing X) and $\varphi_i(S)$ (representing S). The part-whole relationships depend *crucially* upon the particular representation space. We are then tempted to think, in virtue of the bijections between all these spaces, that the part-whole relations will remain the same across representations. But this is precisely what fails: as we have repeatedly seen, Cantorian bijections do not preserve this more nuanced part-whole information, and thus we get different intuitive probability assignments for each representation.

I then considered what I take to be two exemplary recent discussions of the paradox. These are interesting and informative but fail to recognize the heart of Bertrand's discussion and therefore that he has identified a *genuine* paradox. Nonetheless, I do think that both discussions can serve as implicit confirmation of my analysis.

4.5 Summary and Concluding Remarks

In this chapter, I have examined the interplay between infinitary intuitions and probability theory. We began by examining the theory of numerosities and saw that, as a theory of infinite counting, numerosities validate **PW** rather than **CP**. In order to do this, the construction of our theory had to be rather complicated; in particular, given a particular presentation of numerosity theory, it was crucial that each element of a set A being counted be given a "label." Furthermore, we saw that, if this labelling were to change, the size assignment for A would also change. I called this sensitivity to labelling the **Type 1** sensitivity of numerosities. Furthermore, we saw that numerosities exhibit what I called **Type 2** sensitivity, viz., sensitivity to the model construction in which size assignments are made (more precisely: the sets present in our underlying selective ultrafilter).

I then put these characteristics of numerosities to work in probability theory. My first aim was to show that various paradoxes in infinitary probability theory (Label Invariance, God's Lottery, Bertrand's paradox) can all be fruitfully conceptualized as paradoxes of the infinite. In one way or other, they all involve a conflict between our infinitary intuitions and the coarse framework based upon 1-1 correspondence in which we try to formalize these intuitions. I used the Type 1 sensitivity of numerosities to provide an independent argument that the Label Invariance assumption (at any level of generality) had to be jettisoned. This idea emerged again in the discussion of Williamson's Coin: there I argued that Williamson's argument can be disrupted by jettisoning his background Cantorian assumptions and rejecting his "re-labelling" isomorphism. Finally, I emphasized the utility of Type 2 sensitivity of numerosities in dissolving God's Lottery. In particular, this sensitivity can actually be turned into an advantage: once we ascertain the intuitions we wish to validate, we can precisely calibrate our model to produce results consonant with these intuitions.

Finally, I considered Bertrand's paradox. I argued that Bertrand's paradox is generated by a conflict between our intuition that the relative sizes of infinite sets should be preserved under bijections and the actual relative size verdicts of geometric measures. This paradox is made even more complicated due to the presence of multiple representation spaces, each of which yields a different verdict. As with the other paradoxes above, we are misled by our uncritical reliance upon Cantorian intuitions when we in fact require a much more nuanced formal framework.

I then considered one recent proposal to dissolve the paradox via such a framework, viz., the theory of Gross-one. Here the **PW** intuition emerged once more in different dress: in order to produce a workable model of Bertrand's initial problem, we required a way of manipulating infinite quantities such that **PW** and the usual arithmetical operations are satisfied. It seems entirely possible that numerosities could be used to this effect; however, as I noted, this way of proceeding side-steps what actually produces the paradox. Thus, though my analysis of Bertrand's paradox exhibits some affinities with my analysis of the paradoxes in Section 4.3, the utility of new infinitary techniques here is rather less clear.

I hope to have shown that reflection upon the conceptual underpinnings of infinity sheds light on some rather vexing puzzles. In particular, these so-called "probabilistic" paradoxes are no such thing. Rather, they are all generated by a conflict between our long comfort with **CP** and cardinalities and our intuitions about part-whole and relative size relations among infinite sets. The disagreement of probability values merely reflects this deeper conflict.

A final theme that has emerged is the indeterminacy present in our theories of infinite counting. This phenomenon is made explicit by our examination of numerosities: it is fascinating that the theory of numerosities, in order to validate a fundamental infinitary intuition (**PW**), is so very *sensitive*. This inclines me to think that there is an inextricably indeterminate character to the infinite, an indeterminacy rather different from the classical idea that the infinite is "indefinitely extensible." That is, even very basic statements about infinite quantities, e.g., $\alpha = \mathfrak{n}(\mathbb{N})$ is odd, fail to have determinate truth-values. I have argued that, because we know how to force a particular truth-value in the context of numerosities, this indeterminacy can serve as a theoretical advantage. Nonetheless, it is somewhat troubling that *all* our theories of infinite counting exhibit this indeterminacy ($2^{\aleph_0} = \aleph_1$? Is $\alpha = \mathfrak{n}(\mathbb{N})$ odd?, etc.). This will have to serve as food for future thought.

Bibliography

[Aerts and de Bianchi, 2014] Aerts, D. and de Bianchi, M. S. (2014). Solving the hard problem of Bertrand's paradox. *Journal of Mathematical Physics*, 55(083503).

[Almheiri et al., 2020] Almheiri, A., Hartman, T., Maldacena, J., Shaghoulian, E., and Tajdini, A. (2020). Replica wormholes and the entropy of Hawking radiation. *Journal of High Energy Physics*, 5:1–42.

[Arana, 2015] Arana, A. (2015). On the Depth of Szemerédi's Theorem. *Philosophia Mathematica*, 23(2):163–176.

[Arana, 2017] Arana, A. (2017). On the alleged simplicity of pure proof. In Kossak, R. and Ording, P., editors, *Simplicity: Ideals of Practice in Mathematics and the Arts*, pages 207–229. Springer.

[Arana, 2019] Arana, A. (2019). Elementarity and purity. In Alvarez, C. and Arana, A., editors, *Analytic Philosophy and the Foundations of Mathematics*. Palgrave-Macmillan.

[Arana and Mancosu, 2012] Arana, A. and Mancosu, P. (2012). On the relationship between plane and solid geometry. *The Review of Symbolic Logic*, 5(2):294–353.

[Aurich et al., 2021] Aurich, R., Buchert, T., France, M. J., and Steiner, F. (2021). The variance of the CMB temperature gradient: a new signature of a multiply connected Universe. *Classical and Quantum Gravity*, 38(225005).

[Aurich et al., 2008] Aurich, R., Janzer, H., Lustig, S., and Steiner, F. (2008). Do we live in a 'small universe'? *Classical and Quantum Gravity*, 25(125006):1–12.

[Aurich et al., 2004] Aurich, R., Lustig, S., Steiner, F., and Then, H. (2004). Hyperbolic universes with a horned topology and the cosmic microwave background anisotropy. *Classical and Quantum Gravity*, 21:4901–4925.

[Avigad, 2003] Avigad, J. (2003). Number theory and elementary arithmetic. *Philosophia Mathematica*, 11(3):257–284.

[Avigad, 2009] Avigad, J. (2009). The metamathematics of ergodic theory. *Annals of Pure and Applied Logic*, 157:64–76.

[Avigad et al., 2010] Avigad, J., Gerhardy, P., and Towsner, H. (2010). Local stability of ergodic averages. volume 362. Transactions of the American Mathematical Society.

[Avigad and Simic, 2006] Avigad, J. and Simic, K. (2006). Fundamental notions of analysis in subsystems fo second-order arithmetic. *Annals of Pure and Applied Logic*, 136:138–184.

[Avigad and Towsner, 2010] Avigad, J. and Towsner, H. (2010). Metastability in the Furstenberg-Zimmer tower. *Fundamenta Mathematicae*, 210:243–268.

[Baker, 2005] Baker, A. (2005). Are there Genuine Mathematical Explanations of Physical Phenomena? *Mind*, 114:223–238.

[Baker, 2016] Baker, A. (2016). Non-Optional Projects. In Leibowitz, U. D. and Sinclair, N., editors, *Explanation in Ethics and Mathematics: Debunking and Dispensability*, chapter 12, pages 220–235. Oxford: Oxford University Press.

[Bangu, 2010] Bangu, S. (2010). On Bertrand's paradox. *Analysis*, 70(1):30–35.

[Barnes, 1969] Barnes, J. (1969). Aristotle's Theory of Demonstration. *Phronesis*, 14(2):123–152.

[Barnes, 1993] Barnes, J. (1993). *Aristotle: Posterior Analytics*. Clarendon Aristotle Series. Oxford: Oxford University Press, second edition.

[Bartha, 2004] Bartha, P. (2004). Countable Additivity and the de Finetti Lottery. *British Journal for the Philosophy of Science*, 55:301–321.

[Bartha and Hitchcock, 1999] Bartha, P. and Hitchcock, C. (1999). The Shooting-Room Paradox and Conditionalizing on "Measurably Challenged" Sets. *Synthese*, 118:403–437.

[Bartha and Johns, 2001] Bartha, P. and Johns, R. (2001). Probability and Symmetry. *Philosophy of Science Supplement*, 68(3):S109–S122.

[Barwise and Schlipf, 1975] Barwise, J. and Schlipf, J. (1975). On recursively saturated models of arithmetic. In Saracino, D. and Weispfenning, V., editors, *Model Theory and Algebra*, number 498 in Lecture Notes in Mathematics, pages 42–55. Springer.

[Batterman, 2013] Batterman, R., editor (2013). *Oxford Handbook of Philosophy of Physics*. Oxford University Press.

[Beisbart, 2009] Beisbart, C. (2009). Can We Justifiably Assume the Cosmological Principle in Order to Break Model Underdetermination in Cosmology? *Journal for General Philosophy of Science*, 40:175–205.

[Beleznay and Foreman, 1996] Beleznay, F. and Foreman, M. (1996). The complexity of the collection of measure-distal transformations. *Ergodic Theory Dynamical Systems*, 16:929–962.

[Ben-Menahem, 2001] Ben-Menahem, Y. (2001). Convention: Poincaré and Some of His Critics. *The British Journal for the Philosophy of Science*, 52(3):471–513.

[Benacerraf, 1965] Benacerraf, P. (1965). What Numbers Could not Be. *The Philosophical Review*, 74(1):47–73.

[Benacerraf, 1973] Benacerraf, P. (1973). Mathematical Truth. *The Journal of Philosophy*, 70(19):661–679.

[Benci and Baglini, 2021] Benci, V. and Baglini, L. L. (2021). Euclidean Numbers and Numerosities. *The Journal of Symbolic Logic*, pages 1–35.

[Benci et al., 2014] Benci, V., Bottazzi, E., and Nasso, M. D. (2014). Elementary numerosity and measures. *Journal of Logic and Analysis*, 6(3):1–14.

[Benci et al., 2015] Benci, V., Bottazzi, E., and Nasso, M. D. (2015). Some applications of numerosities in measure theory. *Rend. Lincei Mat. Appl.*, 26:37–47.

[Benci et al., 2013] Benci, V., Horsten, L., and Wenmackers, S. (2013). Non-Archimedean Probabilities. *Milan Journal of Mathematics*, 81:121–151.

[Benci et al., 2018] Benci, V., Horsten, L., and Wenmackers, S. (2018). Infinitesimal Probabilities. *British Journal for the Philosophy of Science*, 69:509–552.

[Benci and Nasso, 2003a] Benci, V. and Nasso, M. D. (2003a). Alpha-Theory: An Elementary Axiomatics for Nonstandard Analysis. *Expositiones Mathematicae*, 21:355–386.

[Benci and Nasso, 2003b] Benci, V. and Nasso, M. D. (2003b). Numerosities of labelled sets: A new way of counting. *Advances in Mathematics*, 173:50–67.

[Benci and Nasso, 2019] Benci, V. and Nasso, M. D. (2019). *How to Measure the Infinite: Mathematics with Infinite and Infinitesimal Numbers*. Singapore: World Scientific.

[Benci et al., 2006] Benci, V., Nasso, M. D., and Forti, M. (2006). An Aristotelian Notion of Size. *Annals of Pure and Applied Logic*, 143:43–53.

[Benci et al., 2007] Benci, V., Nasso, M. D., and Forti, M. (2007). A Euclidean measure of size for mathematical universes. *Logique et Analyse*, 50(197):43–62.

[Bertrand, 1889] Bertrand, J. (1889). *Calcul des probabilités*. Paris: Gauthier-Villars.

[Betti, 2010] Betti, A. (2010). Explanation in Metaphysics and Bolzano's Theory of Ground and Consequence. *Logique et Analyse*, 56(211):281–316.

[Bishop and Bridges, 1985] Bishop, E. and Bridges, D. (1985). *Constructive Analysis*. Number 279 in Grundlehren der mathematischen Wissenschaften. Springer.

[Blass et al., 2012] Blass, A., Nasso, M. D., and Forti, M. (2012). Quasi-selective ultrafilters and asymptotic numerosities. *Advances in Mathematics*, 231:1462–1486.

[Bolzano, 1972] Bolzano, B. (1972). *Theory of Science*. Oxford: Basil Blackwell.

[Bolzano, 1975] Bolzano, B. (1975). *Paradoxien des Unendlichen*. Hamburg: Felix Meiner Verlag.

[Booth, 1970] Booth, D. (1970). Ultrafilters on a Countable Set. *Annals of Mathematical Logic*, 2:1–24.

[Bourbaki, 1950] Bourbaki, N. (1950). The Architecture of Mathematics. *The American Mathematical Monthly*, 57(4):221–232.

[Brown and Simpson, 1993] Brown, D. and Simpson, S. G. (1993). The Baire Category Theorem in Weak Subsystems of Second-Order Arithmetic. *Journal of Symbolic Logic*, 58(2):557–578.

[Brundit and Ellis, 1979] Brundit, G. and Ellis, G. (1979). Life in the Infinite Universe. *Quarterly Journal of the Royal Astronomical Society*, 20:37–41.

[Buchholz et al., 1981] Buchholz, W., Feferman, S., Pohlers, W., and Sieg, W. (1981). *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretic Studies*. Number 897 in Lecture Notes in Mathematics. Springer.

[Butterfield, 2014] Butterfield, J. (2014). On under-determination in cosmology. *Studies in History and Philosophy of Modern Physics*, 46:57–69.

[Caldon and Ignjatovic, 2005] Caldon, P. and Ignjatovic, A. (2005). On mathematical instrumentalism. *Journal of Symbolic Logic*, 70(3):778–794.

[Chiu and Hoffman, 1964] Chiu, H.-Y. and Hoffman, W., editors (1964). *Gravitation and Relativity*. Physical Investigations of the Universe. W.A. Benjamin, Inc.

[Cinti and Fano, 2021] Cinti, E. and Fano, V. (2021). Careful with those scissors, Eugene! Against the observational indistinguishability of spacetimes. *Studies in History and Philosophy of Science*, 89:103–113.

[Cornell et al., 2000] Cornell, G., Silverman, J. H., and Stevens, G., editors (2000). *Modular Forms and Fermat's Last Theorem*. Springer.

[Cornish et al., 1998] Cornish, N., Spergel, D., and Starkman, G. (1998). Circles in the sky: finding topology with the microwave background radiation. *Classical and Quantum Gravity*, 15:2657–2670.

[Cornish et al., 2004] Cornish, N., Spergel, D., Starkman, G. D., and Komatsu, E. (2004). Constraining the topology of the universe. *Phys. Rev. Lett.*, 92:201302.

[Cutland, 1983] Cutland, N. (1983). Nonstandard measure theory and its applications. *Bulletin of the London Mathematical Society*, 15:529–589.

[de Finetti, 1974] de Finetti, B. (1974). *Theory of Probability*. London: Wiley.

[de la Vallée Poussin, 1896] de la Vallée Poussin, C. (1896). Recherches analytiques sur la théorie des nombres premiers. *Ann. Soc, Sci. Bruxelles*, 20:183–256.

[Dean and Walsh, 2017] Dean, W. and Walsh, S. (2017). The Prehistory of the Subsystems of Second-Order Arithmetic. *The Review of Symbolic Logic*, 10(2):357–396.

[Detlefsen, 1990] Detlefsen, M. (1990). On an alleged refutation of Hilbert's program using Gödel's first incompleteness theorem. *Journal of Philosophical Logic*, 19(4):343–377.

[Detlefsen, 2008] Detlefsen, M. (2008). *The Philosophy of Mathematical Practice*, chapter 7: Purity as an Ideal of Proof, pages 179–198. Oxford: Oxford University Press.

[Detlefsen and Arana, 2011] Detlefsen, M. and Arana, A. (2011). Purity of methods. *Philosophers' Imprint*, 11(2):1–20.

[Earman, 1993] Earman, J. (1993). Underdetermination, realism, and reason. *Midwest Studies in Philosophy*, XVIII:19–38.

[Earman, 1995] Earman, J. (1995). *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes*. Oxford University Press: New York.

[Earman, 2001] Earman, J. (2001). Lambda: The Constant That Refuses to Die. *Archive for the History of Exact Sciences*, 55(3):189–220.

[Eastaugh, 2015] Eastaugh, B. (2015). *Reverse Mathematics: A Philosophical Account*. PhD thesis, University of Bristol.

[Eastaugh, 2019] Eastaugh, B. (2019). Set existence and closure conditions: unravelling the standard view of reverse mathematics. *Philosophia Mathematica*, 27(2):153–176.

[Easwaran, 2014] Easwaran, K. (2014). Regularity and Hyperreal Credences. *Philosophical Review*, 123:1–41.

[Easwaran et al., 2023] Easwaran, K., Hájek, A., Mancosu, P., and Oppy, G. (2023). Infinity. *The Stanford Encyclopedia of Philosophy*.

[Einsiedler and Ward, 2011] Einsiedler, M. and Ward, T. (2011). *Ergodic Theory with a view towards Number Theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag.

[Einstein, 1950] Einstein, A. (1950). *The Meaning of Relativity*. Princeton University Press, third edition.

[Ellis, 1971] Ellis, G. (1971). Topology and cosmology. *General Relativity and Gravitation*, 2:7–21.

[Ellis, 1975] Ellis, G. (1975). Cosmology and verifiability. *Quarterly Journal of the Royal Astronomical Society*, 16:245–264.

[Ellis, 1980] Ellis, G. (1980). Limits to verification in cosmology. *Annals New York Academy of Sciences*, 336:130–160.

[Ellis, 2007] Ellis, G. (2007). Issues in the Philosophy of Cosmology. In Butterfield, J. and Earman, J., editors, *Handbook of the Philosophy of Science. Philosophy of Physics: Part B*, pages 1185–1285. Elsevier B.V.

[Ellis, 2014] Ellis, G. (2014). On the philosophy of cosmology. *Studies in History and Philosophy of Modern Physics*, 46:5–23.

[Ellis and Hawking, 1973] Ellis, G. and Hawking, S. (1973). *The Large Scale Structure of Space-Time*. Cambridge Monographs in Mathematical Physics. Cambridge University Press.

[Ellis and Larena, 2020] Ellis, G. and Larena, J. (2020). The case for a closed universe. *Astronomy and Geophysics*, 61(1):38–40.

[Ellis and Schreiber, 1986] Ellis, G. and Schreiber, G. (1986). Observational and dynamic properties of small universes. *Physics Letters A*, 115:97–107.

[Ellis and Stoeger, 2009] Ellis, G. and Stoeger, W. R. (2009). A note on infinities in eternal inflation. *General Relativity and Gravitation*, 41:1475–1484.

[Ellis et al., 2018] Ellis, G. F., Meissner, K. A., and Nicolai, H. (2018). The physics of infinity. *Nature Physics*, 14:770–772.

[Fabre et al., 2013] Fabre, O., Prunet, S., and Uzan, J.-P. (2013). Topology beyond the horizon: how far can it be probed? *Physical Review D*, 92(4).

[Fagundes, 1983] Fagundes, H. (1983). The Compactification of Friedmann's Hyperbolic Model. *Phys. Rev. Lett.*, 51(517).

[Feferman, 1964] Feferman, S. (1964). Systems of Predicative Analysis, i. *The Journal of Symbolic Logic*, 29:1–30.

[Feferman, 1968] Feferman, S. (1968). Systems of Predicative Analysis, ii. *Journal of Symbolic Logic*, 33:193–220.

[Feferman, 1987] Feferman, S. (1987). Infinity in mathematics: is Cantor necessary? In di Francia, G. T., editor, *L'infinito nella scienza*, pages 151–209. Istituto della Enciclopedia Italiana, Rome.

[Feferman, 1996] Feferman, S. (1996). Kreisel's unwinding program. In Odifreddi, P., editor, *Kreiseliana: About and Around Georg Kreisel*. A.K. Peters.

[Feferman, 1998a] Feferman, S. (1998a). *In the Light of Logic*. Oxford: Oxford University Press.

[Feferman, 1998b] Feferman, S. (1998b). What Rests on What? The Proof-Theoretic Analysis of Mathematics. In *In the Light of Logic*, chapter 10, pages 187–208. New York: Oxford University Press.

[Feferman and Sieg, 1981] Feferman, S. and Sieg, W. (1981). Proof theoretic equivalences between classical and constructive theories for analysis. In Buchholz, W., Feferman, S., Pohlers, W., and Sieg, W., editors, *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretic Studies*, number 897 in Lecture Notes in Mathematics, chapter 2, pages 78–142. Springer.

[Ferreirós, 1999] Ferreirós, J. (1999). *Labyrinth of Thought: A history of set theory and its role in modern mathematics*. Basel: Birkhäuser.

[Field, 1980] Field, H. (1980). *Science without Numbers*. Princeton, NJ: Princeton University Press.

[Field, 1989] Field, H. (1989). *Realism, Mathematics and Modality*. Oxford: Basil Blackwell.

[Folland, 1999] Folland, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics. Wiley-Interscience, second edition.

[Freivogel et al., 2006] Freivogel, B., Kleban, M., Martinez, M., and Susskind, L. (2006). Observational consequences of a landscape. *Journal of High Energy Physics*, 3(39).

[Friedman, 1974] Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy*, 71(1):1–19.

[Friedman, 1992] Friedman, M. (1992). *Kant and the Exact Sciences*. Harvard University Press.

[Friedman, 1999] Friedman, M. (1999). *Reconsidering Logical Positivism*. Cambridge University PRess.

[Friedman, 2000] Friedman, M. (2000). Geometry, Construction, and Intuition in Kant and His Successors. In Sher, G. and Tieszen, R., editors, *Between Logic and Intuition*, pages 186–219. Cambridge University Press.

[Friedman, 2001] Friedman, M. (2001). *The Dynamics of Reason*. CSLI Publishing.

[Friedman, 2007] Friedman, M. (2007). Coordination, Constitution, and Convention: the Evolution of the A Priori in Logical Empiricism. In Richardson, A. and Uebel, T., editors, *The Cambridge Companion to Logical Empiricism*, pages 91–116. Cambridge University Press.

[Friedman, 2009] Friedman, M. (2009). Einstein, Kant, and the Relativized *A Priori*. In et. al., M. B., editor, *Constituting Objectivity*, pages 253–267. Springer.

[Furstenberg, 1977] Furstenberg, H. (1977). Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *Journal d' Analyse Mathématique*, 34:204–256.

[Furstenberg, 1981] Furstenberg, H. (1981). Poincaré recurrence and number theory. *Bulletin of the American Mathematical Society*, 5(3).

[Furstenberg and Katznelson, 1979] Furstenberg, H. and Katznelson, Y. (1979). An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.*, 34:275–291.

[Furstenberg et al., 1982] Furstenberg, H., Katznelson, Y., and Ornstein, D. (1982). The Ergodic Theoretical Proof of Szemerédi's Theorem. *Bulletin of the American Mathematical Society*, 7(3).

[Furstenberg and Weiss, 1978] Furstenberg, H. and Weiss, B. (1978). Topological dynamics and combinatorial number theory. *Journal d' Analyse Mathématique*, 34:61–85.

[Garriga and Vilenkin, 2001] Garriga, J. and Vilenkin, A. (2001). Many Worlds in One. *Physics Review D*, 64(043511):1–5.

[Gentzen, 1969] Gentzen, G. (1969). The Concept of Infinity in Mathematics. In *The Collected Works of Gerhard Gentzen*, 223-233, chapter 6. Amsterdam: North-Holland.

[Geroch, 1970] Geroch, R. (1970). Domain of Dependence. *Journal of Mathematical Physics*, 11(2):437–449.

[Geroch, 1967] Geroch, R. P. (1967). Topology in General Relativity. *Journal of Mathematical Physics*, 8:782–786.

[Geroch and Horowitz, 1979] Geroch, R. P. and Horowitz, G. T. (1979). Global structures of spacetimes. In Israel, W. and Hawking, S., editors, *General Relativity: An Einstein Centenary Survey*, chapter 5, pages 212–293. Cambridge University Press.

[Gibbons et al., 1983] Gibbons, G., Hawking, S., and Siklos, S., editors (1983). *The Very Early Universe*. Cambridge University Press.

[Gimbel, 2004] Gimbel, S. (2004). Un-conventional wisdom: theory-specificity in Reichenbach's geometric conventionalism. *Studies in History and Philosophy of Modern Physics*, 35:457–481.

[Girard, 1987] Girard, J.-Y. (1987). *Proof Theory and Logical Complexity*, volume 1. Bibliopolis.

[Glymour, 1972] Glymour, C. (1972). Topology, Cosmology, and Convention. *Synthese*, 24:195–218.

[Glymour, 1977] Glymour, C. (1977). Indistinguishable Space-Times and the Fundamental Group. In Earman, J., Glymour, C., and Stachel, J., editors, *Foundations of Space-Time Theories*, volume 8 of *Minnesota Studies in the Philosophy of Science*, pages 50–60. Minneapolis: University of Minnesota Press.

[Gödel, 1931] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica under verwandter Systeme I. *Monatsh. Math. Phys.*, XXXVIII:173–198.

[Gödel, 1933] Gödel, K. (1933). Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines Mathematischen Kolloquiums*, 4:34–38.

[Gödel, 1958] Gödel, K. (1958). Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287.

[Goldblatt, 1998] Goldblatt, R. (1998). *Lectures on the Hyperreals*. Number 188 in Graduate Texts in Mathematics. Springer-Verlag.

[Gowers, 2001] Gowers, T. (2001). A New Proof of Szemerédi's Theorem. *Geometric and Functional Analysis*, 11:465–588.

[Granville, 2008] Granville, A. (2008). Analytic number theory. In Timothy Gowers, J. B.-G. and Leader, I., editors, *The Princeton Companion to Mathematics*, chapter 17, pages 332–348. Princeton University Press.

[Gray and Ferreirós, 2021] Gray, J. and Ferreirós, J. (2021). Epistemology of Geometry. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Fall 2021 edition.

[Green and Tao, 2008] Green, B. and Tao, T. (2008). The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics*, 167:481–547.

[Guth, 1981] Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23(2):347–356.

[Guth, 2007] Guth, A. H. (2007). Eternal inflation and its implications. *Journal of Physics A: Mathematical and Theoretical*, 40:6811–6826.

[Gyenis and Rédei, 2015] Gyenis, Z. and Rédei, M. (2015). Defusing Bertrand's paradox. *British Journal for the Philosophy of Science*, 66:349–373.

[Hadamard, 1896] Hadamard, J. (1896). Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bull. Soc. Math. France*, 24:199–220.

[Hafner and Mancosu, 2008] Hafner, J. and Mancosu, P. (2008). *The Philosophy of Mathematical Practice*, chapter 6: Beyond Unification, pages 151–178. New York: Oxford University Press.

[Hájek, 2009] Hájek, A. (2009). Dutch book arguments. In Anand, P., Pattanaik, P., and Puppe, C., editors, *The Handbook of Rational and Social Choice*, chapter 7, pages 173–195. Oxford: Oxford University Press.

[Hallett, 2008] Hallett, M. (2008). Reflections on the purity of method in Hilbert's *Grundlagen der Geometrie*. In Mancosu, P., editor, *The Philosophy of Mathematical Practice*, pages 198–255. Oxford: Oxford University Press.

[Hayward and Twamley, 1990] Hayward, G. and Twamley, J. (1990). Large scale structure in a spatially compact hyperbolic universe. *Physics Letters A*, 149(2,3):84–90.

[Hilbert, 1967] Hilbert, D. (1967). On the Infinite. In van Heijenoort, J., editor, *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, pages 367–392. Cambridge, MA: Harvard University Press.

[Howson, 2017] Howson, C. (2017). Regularity and infinitely tossed coins. *European Journal for Philosophy of Science*, 7:97–102.

[Howson and Urbach, 1993] Howson, C. and Urbach, P. (1993). *Scientific Reasoning: The Bayesian Approach*. La Salle: Open Court Press, second edition.

[Hussey, 1983] Hussey, E. (1983). *Aristotle's Physics: Books III and IV*. Clarendon Aristotle Series. Oxford: Oxford University Press.

[Ingham, 2008] Ingham, A. (2008). Review 10,595c (mr0029411). *Mathematical Reviews*, pages 651–654.

[Isaacson, 1996] Isaacson, D. (1996). Arithmetical truth and hidden higher-order concepts. In Hart, W., editor, *The Philosophy of Mathematics*, pages 203–224. New York: Oxford University Press.

[Ivanova, 2015] Ivanova, M. (2015). Conventionalism, structuralism and neo-Kantianism in Poincaré's philosophy of science. *Studies in History and Philosophy of Modern Physics*, 52:114–122.

[J. Richard Gott, 1980] J. Richard Gott, I. (1980). Chaotic cosmologies and the topology of the Universe. *Mon. Not. R. Astr. Soc.*, 193:153–169.

[Jaynes, 1973] Jaynes, E. (1973). The Well-Posed Problem. *Foundations of Physics*, 3:477–493.

[Kant, 1961] Kant, I. (1961). *Critique of Pure Reason*. Translated by Norman Kemp Smith. London: Macmillan, Second edition.

[Katz, 1981] Katz, F. (1981). *Sets and Their Sizes*. PhD thesis, MIT.

[Katz and Reimann, 2018] Katz, M. and Reimann, J. (2018). *An Introduction to Ramsey Theory: Fast Functions, Infinity, and Metamathematics*, volume 87 of *Student Mathematical Library*. American Mathematical Society.

[Kaye and Wong, 2007] Kaye, R. and Wong, T. L. (2007). On Interpretations of Arithmetic and Set Theory. *Notre Dame Journal of Formal Logic*, 48(4):497–510.

[Kelly, 1996] Kelly, K. (1996). *The Logic of Reliable Inquiry*. Oxford: Oxford University Press.

[Keynes, 1921] Keynes, J. M. (1921). *A Treatise on Probability*. Dover Publications.

[Kim, 1988] Kim, J. (1988). Explanatory Realism, Causal Realism, and Explanatory Exclusion. *Midwest Studies in Philosophy*, XII:225–239.

[Kitcher, 1975] Kitcher, P. (1975). Bolzano's ideal of algebraic analysis. *Studies in History and Philosophy of Science*, 6:229–269.

[Kitcher, 1976] Kitcher, P. (1976). Explanation, Conjunction, and Unification. *Journal of Philosophy*, 73(8):207–212.

[Kitcher, 1981] Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science*, 48:507–531.

[Kitcher, 1984] Kitcher, P. (1984). *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

[Kitcher, 1989] Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World (sections 1-4.5). In Kitcher, P. and Salmon, W., editors, *Scientific Explanation*, volume Minnesota Studies in the Philosophy of Science, XIII, pages 410–437. Minneapolis: University of Minnesota Press.

[Klyve, 2013] Klyve, D. (2013). In defense of Bertrand: The non-restrictiveness of reasoning by example. *Philosophia Mathematica*, 21(3):365–370.

[Knobe et al., 2006] Knobe, J., Olum, K. D., and Vilenkin, A. (2006). Philosophical Implications of Inflationary Cosmology. *British Journal for the Philosophy of Science*, 57:47–67.

[Kohlenbach, 2008] Kohlenbach, U. (2008). *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics. Springer.

[Kra, 2007] Kra, B. (2007). Ergodic methods in additive combinatorics. In *Additive Combinatorics*, volume 43 of *CRM Proc. Lecture Notes*, pages 103–143. Amer. Math. Soc.

[Kunen, 1976] Kunen, K. (1976). Some points in $\beta(\mathbb{N})$. *Math. Proc. Cambridge Philos. Soc.*, 80:385–398.

[Lang, 1993] Lang, S. (1993). *Real and Functional Analysis*. Number 142 in Graduate Texts in Mathematics. New York: Springer-Verlag, third edition.

[Lang, 2002] Lang, S. (2002). *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer, revised third edition.

[Lange, 2017] Lange, M. (2017). *Because Without Cause*. Oxford: Oxford University Press.

[Laudan, 1990] Laudan, L. (1990). Demystifying underdetermination. In Savage, C., editor, *Scientific Theories*, pages 267–297. Minneapolis: University of Minnesota Press.

[Lehoucq et al., 1996] Lehoucq, R., Lachièze-Rey, M., and Luminet, J. (1996). Cosmic crystallography. *Astronomy and Astrophysics*, 313:339–346.

[Leibowitz and Sinclair, 2016] Leibowitz, U. D. and Sinclair, N., editors (2016). *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press.

[Leng, 2005] Leng, M. (2005). Mathematical explanation. In Cellucci, C. and Gillies, D., editors, *Mathematical Reasoning and Heuristics*, pages 167–189. London: King's College Publications.

[Levin, 2002] Levin, J. (2002). Topology and the cosmic microwave background. *Physics Reports*, 365:251–333.

[Lewis, 1980] Lewis, D. K. (1980). A subjectivist's guide to objective chance. In Carnap, R. and Jeffrey, R. C., editors, *Studies in Inductive Logic and Probability*, pages 263–293. Berkeley: University of California Press.

[Linde, 1982] Linde, A. D. (1982). A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108:389–393.

[Linnebo and Shapiro, 2018] Linnebo, Ø. and Shapiro, S. (2018). Actual and Potential Infinity. *Noûs*, 53(1):160–191.

[Lipton, 2004] Lipton, P. (2004). *Inference to the Best Explanation*. International Library of Philosophy. Routledge, second edition.

[Loeb, 1975] Loeb, P. (1975). Conversion from nonstandard to standard measure spaces and applications in probability theory. *Transactions of the American Mathematical Society*, 211:113–122.

[Luminet, 2003] Luminet, J. (2003). Dodecahedral space topology as an explanation for weak wide-angle temperature correlations in the cosmic microwave background. *Nature*, 425:593–595.

[Luminet, 2015] Luminet, J. (2015). Cosmic Topology. *Scholarpedia*, 10(8):31544.

[Luminet, 2008] Luminet, J. P. (2008). *The Wraparound Universe*. A.K. Peters.

[Luminet and Lachièze-Rey, 1995] Luminet, J. P. and Lachièze-Rey, M. (1995). Cosmic Topology. *Physics Reports*, 254.

[Macintyre, 2011] Macintyre, A. (2011). The impact of Gödel's incompleteness theorems on mathematics. In Baaz, M., editor, *Kurt Gödel and the Foundations of Mathematics: Horizons of Truth*, pages 3–25. Cambridge University Press.

[Maddy, 1998] Maddy, P. (1998). $V = L$ and Maximize. In *Logic Colloquium '95 (Haifa)*, volume Lecture Notes Logic, pages 134–152. Springer, Berlin.

[Malament, 1977] Malament, D. (1977). Observationally Indistinguishable Space-Times. In Earman, J., Glymour, C., and Stachel, J., editors, *Foundations of Space-Time Theories*, volume 8 of *Minnesota Studies in the Philosophy of Science*, pages 61–80. Minneapolis: University of Minnesota Press.

[Manchak, 2009] Manchak, J. B. (2009). Can we know the global structure of spacetime? *Studies in History and Philosophy of Modern Physics*, 40:53–56.

[Manchak, 2011] Manchak, J. B. (2011). What Is a Physically Reasonable Space-Time? *Philosophy of Science*, 78:410–420.

[Manchak, 2013] Manchak, J. B. (2013). Global Spacetime Structure. In Batterman, R., editor, *Oxford Handbook of Philosophy of Physics*, chapter 16. Oxford University Press.

[Mancosu, 1996] Mancosu, P. (1996). *Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century*. Oxford University Press.

[Mancosu, 1998] Mancosu, P. (1998). *From Brouwer to Hilbert*. New York, NY: Oxford University Press.

[Mancosu, 1999] Mancosu, P. (1999). Bolzano and Cournot on mathematical explanation. *Revue d'histoire des sciences*, 52(3/4):429–455.

[Mancosu, 2008a] Mancosu, P. (2008a). Mathematical Explanation: Why it Matters. In Mancosu, P., editor, *The Philosophy of Mathematical Practice*, chapter 5, pages 134–150. New York: Oxford University Press.

[Mancosu, 2008b] Mancosu, P., editor (2008b). *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press.

[Mancosu, 2009] Mancosu, P. (2009). Measuring the Size of Infinite Collections of Natural Numbers: Was Cantor's Theory of Infinite Number Inevitable? *The Review of Symbolic Logic*, 2(4):612–646.

[Mancosu, 2017] Mancosu, P. (2017). *Abstraction and Infinity*. Oxford: Oxford University Press.

[Mancosu, 2018] Mancosu, P. (2018). "Explanation in Mathematics", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). <https://plato.stanford.edu/entries/mathematics-explanation/>.

[Mancosu and Massas, 2023] Mancosu, P. and Massas, G. (2023). Totality, Regularity, and Cardinality in Probability Theory. *Philosophy of Science*, pages 1–20.

[Mancosu and Vailati, 1991] Mancosu, P. and Vailati, E. (1991). Torricelli's Infinitely Long Solid and Its Philosophical Reception in the Seventeenth Century. *Isis*, 82(1):50–70.

[Marinoff, 1994] Marinoff, L. (1994). A Resolution of Bertrand's Paradox. *Philosophy of Science*, 61:1–24.

[Martin, 1998] Martin, D. A. (1998). Mathematical evidence. In Dales, H. and Oliveri, G., editors, *Truth in Mathematics*, pages 215–231. Oxford: Oxford University Press.

[Matet, 2007] Matet, P. (2007). Shelah's proof of the Hales-Jewett theorem revisited. *European Journal of Combinatorics*, 28(6):1742–1745.

[McCabe, 2004] McCabe, G. (2004). The structure and interpretation of cosmology: Part i–general relativistic cosmology. *Studies in History and Philosophy of Modern Physics*, 35:549–595.

[McCall and Armstrong, 1989] McCall, S. and Armstrong, D. (1989). God's Lottery. *Analysis*, 49(4):223–224.

[McLarty, 2010] McLarty, C. (2010). What does it take to prove Fermat's Last Theorem? Grothendieck and the logic of number theory. *The Bulletin of Symbolic Logic*, 16(3):359–377.

[Mikkelson, 2004] Mikkelson, J. (2004). A Resolution of the Wine/Water Paradox. *The British Journal for the Philosophy of Science*, 55:137–145.

[Misner et al., 2017] Misner, C., Thorne, K., and Wheeler, J. (2017). *Gravitation*. Princeton University Press.

[Montalbán, 2011] Montalbán, A. (2011). Open Questions in Reverse Mathematics. *The Bulletin of Symbolic Logic*, 17(3):431–454.

[Moore, 2019] Moore, A. (2019). *The Infinite*. Routledge, third edition.

[Morrison, 2000] Morrison, M. (2000). *Unifying Scientific Theories*. Cambridge University Press.

[Nasso and Forti, 2010] Nasso, M. D. and Forti, M. (2010). Numerosities of Point Sets Over the Real Line. *Transactions of the American Mathematical Society*, 362(10):5355–5371.

[Nelson, 1987] Nelson, E. (1987). *Radically Elementary Probability Theory*. Princeton: Princeton University Press.

[Nolan, 1997] Nolan, D. (1997). Quantitative Parsimony. *British Journal for the Philosophy of Science*, 48(3):329–343.

[Nolan, 2022] Nolan, D. (2022). Space, time and parsimony. *Noûs*, pages 1–21.

[Norton, 2011] Norton, J. D. (2011). Observationally Indistinguishable Spacetimes: A Challenge for Any Inductivist. In Morgan, G. J., editor, *Philosophy of Science Matters: The Philosophy of Peter Achinstein*, chapter 13, pages 164–176. New York: Oxford University Press.

[Norton, 2021] Norton, J. D. (2021). Eternal inflation: when probabilities fail. *Synthese*, 198(S16):S3853–S3875.

[Parker, 2009] Parker, M. (2009). Philosophical method and Galileo's paradox of infinity. In van Kerkhove, B., editor, *New Perspectives on Mathematical Practices*, pages 76–113. World Scientific.

[Parker, 2020] Parker, M. (2020). Comparative infinite lottery logic. *Studies in History and Philosophy of Science*, 84:28–36.

[Parker, 2013] Parker, M. W. (2013). Set Size and the Part-Whole Principle. *The Review of Symbolic Logic*, 6(4):589–612.

[Parker, 2019] Parker, M. W. (2019). Symmetry arguments against regular probability: A reply to recent objections. *European Journal for Philosophy of Science*, 9(8).

[Parker, 2021] Parker, M. W. (2021). Weintraub's response to Williamson's coin flip argument. *European Journal for Philosophy of Science*, 11(71).

[Parsons, 1970] Parsons, C. (1970). On a number-theoretic choice schema and its relation to induction. In Kino, A., Myhill, J., and Vesley, R., editors, *Intuitionism and Proof Theory*, Studies in Logic and Foundations of Mathematics, pages 459–473. North-Holland.

[Parsons, 1983] Parsons, C. (1983). *Mathematics in Philosophy: Selected Essays*. Cornell University Press.

[Poincaré, 2015] Poincaré, H. (2015). *The Foundations of Science*. Cambridge Library Collection. Cambridge University Press.

[Poincaré, 2017] Poincaré, H. (2017). *La science et l'hypothèse*. Champs sciences. Flammarion.

[Rebouças and Gomero, 2004] Rebouças, M. and Gomero, G. (2004). Cosmic topology: A brief overview. *Brazilian Journal of Physics*, 34.

[Reichenbach, 1920] Reichenbach, H. (1920). *Relativitätstheorie und Erkenntnis A Priori*. Berlin: Julius Springer.

[Reichenbach, 1928] Reichenbach, H. (1928). *Philosophie der Raum-Zeit-Lehre*. Berlin: de Gruyter.

[Reichenbach, 1957] Reichenbach, H. (1957). *The Philosophy of Space and Time*. Dover Publications.

[Riemann, 2004] Riemann, B. (2004). *Bernhard Riemann: Collected Papers*. Kendrick Press; Translated by Roger Baker, Charles Christenson, and Henry Orde.

[Riemann, 2017] Riemann, B. (2017). *The Collected Works of Bernhard Riemann*. Dover Books on Mathematics. Dover Publications.

[Rizza, 2018] Rizza, D. (2018). A Study of Mathematical Determination through Bertrand's Paradox. *Philosophia Mathematica*, 26(3):375–395.

[Rosen, 2021] Rosen, J. (2021). Aristotle's Actual Infinities. *Oxford Studies in Ancient Philosophy*, 59:133–186.

[Ross and Minio-Paluello, 1964] Ross, W. and Minio-Paluello, L., editors (1964). *Aristotelis Analytica priora et posteriora*. Oxford.

[Rota, 1997] Rota, G.-C. (1997). *Indiscrete Thoughts*. Boston: Birkhäuser.

[Rovelli, 2008] Rovelli, C. (2008). Quantum gravity. *Scholarpedia*, 3(5):7117.

[Rowbottom, 2013] Rowbottom, D. (2013). Bertrand's paradox revisited: Why Bertrand's "solutions" are all inapplicable. *Philosophia Mathematica*, 21(3):110–114.

[Rudin, 1993] Rudin, W. (1993). Autohomeomorphisms of Compact Groups. *Topology and its Applications*, 52:69–70.

[Ryckman, 2007] Ryckman, T. (2007). Logical Empiricism and the Philosophy of Physics. In Richardson, A. and Uebel, T., editors, *The Cambridge Companion to Logical Empiricism*, chapter 8, pages 193–227. Cambridge: Cambridge University Press.

[Sergeyev, 2003] Sergeyev, Y. (2003). *The Arithmetic of Infinity*. Rende: Edizioni Orizzonti Meridionali.

[Sergeyev, 2009a] Sergeyev, Y. (2009a). Numerical computations and mathematical modelling with infinite and infinitesimal numbers. *Journal of Applied Mathematics and Computation*, 29:177–195.

[Sergeyev, 2009b] Sergeyev, Y. (2009b). Numerical point of view on calculus for functions assuming finite, infinite, and infinitesimal values over finite, infinite, and infinitesimal domains. *Nonlinear Analysis Series A: Theory, Methods and Applications*, 71:1688–1707.

[Shackel, 2007] Shackel, N. (2007). Bertrand's Paradox and the Principle of Indifference. *Philosophy of Science*, 74:150–175.

[Shackel and Rowbottom, 2020] Shackel, N. and Rowbottom, D. (2020). Bertrand's Paradox and the Maximum Entropy Principle. *Philosophy and Phenomenological Research*, CI(3):505–523.

[Shelah, 1988] Shelah, S. (1988). Primitive recursive bounds for van der Waerden numbers. *Journal of the American Mathematical Society*, 1:683–697.

[Simpson,] Simpson, S. Open Problems in Reverse Mathematics. Write-up of invited talk at the AMS-INS-SIAM conference *Computability Theory and Applications*, June 13-17, 1999, Boulder, CO.

[Simpson, 1999] Simpson, S. (1999). *Subsystems of Second Order Arithmetic*. Perspectives in Mathematical Logic. Springer.

[Sklar, 1974] Sklar, L. (1974). *Space, Time, and Spacetime*. University of California Press.

[Skyrms, 1980] Skyrms, B. (1980). *Causal Necessity*. New Haven: Yale University Press.

[Smeenk, 2013] Smeenk, C. (2013). Philosophy of Cosmology. In Batterman, R., editor, *Oxford Handbook of Philosophy of Physics*, chapter 17, pages 607–652. Oxford: Oxford University Press.

[Sober, 2015] Sober, E. (2015). *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press.

[Sorensen, 2014] Sorensen, R. (2014). Parsimony for empty space. *Australasian Journal of Philosophy*, 92:215–230.

[Spergel et al., 2007] Spergel, D., Bean, R., Doré, O., Nolta, M., Bennett, C., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H., Verde, L., Halpern, M., Hill, R., Kogut, A., Limon, M., Meyer, S., Odegard, N., Tucker, G., Weiland, J., Wollack, E.,

and Wright, E. (2007). Three-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. *The Astrophysical Journal Supplement*, 170:377–408.

[Spielman, 1977] Spielman, S. (1977). Physical Probability and Bayesian Statistics. *Synthese*, 36:236–269.

[Starkman, 1998] Starkman, G. D. (1998). Topology and cosmology. *Classical and Quantum Gravity*, 15:2529–2538.

[Steiner, 1978] Steiner, M. (1978). Mathematics, Explanation, and Scientific Knowledge. *Nous*, 12:17–28.

[Szemerédi, 1975] Szemerédi, E. (1975). On sets of integers containing no k elements in arithmetic progression. *Acta Arithmetica*, XXVII:199–245.

[Tait, 1981] Tait, W. W. (1981). Finitism. *Journal of Philosophy*, 78:524–546.

[Tallant, 2013] Tallant, J. (2013). Quantitative parsimony and the metaphysics of time: motivating presentism. *Philosophy and Phenomenological Research*, 87:688–705.

[Tao, 2006] Tao, T. (2006). The dichotomy between structure and randomness, arithmetic progressions, and the primes. In *Proceedings of the International Congress of Mathematicians*, volume I, pages 581–608.

[Tao, 2007] Tao, T. (2007). The ergodic and combinatorial approaches to Szemerédi's theorem. In Granville, A., Nathanson, M., and Solymosi, J., editors, *Additive Combinatorics*, pages 143–193. Providence, RI: American Mathematical Society.

[Tenenbaum, 1995] Tenenbaum, G. (1995). *Introduction to Analytic and Probabilistic Number Theory*. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press.

[Torretti, 1978] Torretti, R. (1978). *Philosophy of Geometry from Riemann to Poincaré*, volume 7 of *Episteme*. D. Reidel: Dordrecht, Holland.

[Towsner, 2008] Towsner, H. (2008). *Some Results in Logic and Ergodic Theory*. PhD thesis, Carnegie Mellon University.

[van der Waerden, 1928] van der Waerden, B. (1928). Beweis einer Baudetschen Vermutung. *Nieuw. Arch. Wisk.*, 15:212–216.

[van der Waerden, 1998] van der Waerden, B. (1998). Wie der Beweis der Vermutung von Baudet gefunden wurde. *Elemente der Mathematik*, 53:139–148.

[van Douwen, 1984] van Douwen, E. K. (1984). A Compact Space with a Measure that Knows which Sets are Homeomorphic. *Advances in Mathematics*, 52:1–33.

[van Fraassen, 1989] van Fraassen, B. (1989). *Laws and Symmetry*. Clarendon: Oxford.

[van Heijenoort, 1967] van Heijenoort, J., editor (1967). *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Cambridge, MA: Harvard University Press.

[Vilenkin and Winitzki, 1997] Vilenkin, A. and Winitzki, S. (1997). Probability distribution for Omega in open-universe inflation. *Physical Review D*, 55(2):548–559.

[von Mises, 1957] von Mises, R. (1957). *Probability, Statistics and Truth*. George Allen and Unwin.

[Wald, 1984] Wald, R. (1984). *General Relativity*. Chicago: University of Chicago Press.

[Weintraub, 2008] Weintraub, R. (2008). How Probable Is an Infinite Sequence of Heads? A Reply to Williamson. *Analysis*, 68(3):247–250.

[Wenmackers, 2023] Wenmackers, S. (2023). Uniform probability in cosmology. *Studies in History and Philosophy of Science Part A*, 101(C):48–60.

[Wenmackers and Horsten, 2013] Wenmackers, S. and Horsten, L. (2013). Fair infinite lotteries. *Synthese*, 190:37–61.

[Weyl, 2012] Weyl, H. (2012). Levels of Infinity. In Pesic, P., editor, *Levels of Infinity: Selected Writings on Mathematics and Philosophy*, pages 17–31. Dover Publications.

[Wheeler, 1964] Wheeler, J. (1964). Mach's principle as boundary condition for Einstein's equations. In Chiu, H.-Y. and Hoffman, W., editors, *Gravitation and Relativity*, Physical Investigations of the Universe. W.A. Benjamin, Inc.

[Wigner, 1960] Wigner, E. (1960). The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Communications in Pure and Applied Mathematics*, 13(1):1–14.

[Williams, 1969] Williams, N. (1969). On Grothendieck universes. *Compositio Mathematica*, 21(1):1–3.

[Williamson, 1999] Williamson, J. (1999). Countable Additivity and Subjective Probability. *British Journal for the Philosophy of Science*, 50:401–416.

[Williamson, 2007] Williamson, T. (2007). How probable is an infinite sequence of heads? *Analysis*, 67(3):173–180.

[Wittgenstein, 1978] Wittgenstein, L. (1978). *Remarks on the Foundations of Mathematics*. Oxford: Basil Blackwell, revised edition.

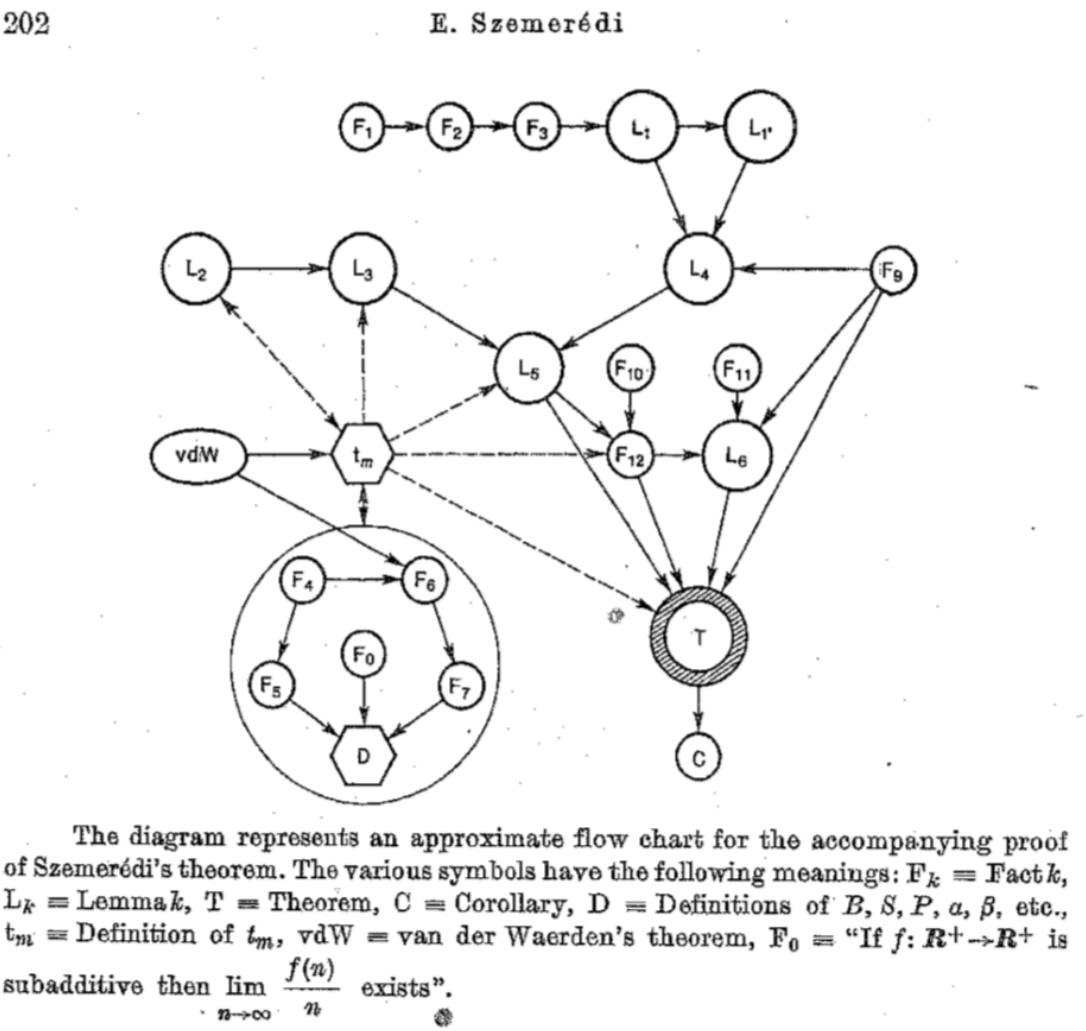
[Wolf, 1967] Wolf, J. A. (1967). *Spaces of constant curvature*. London: McGraw-Hill.

[Yokoyama, 2010] Yokoyama, K. (2010). On Π^1_1 conservativity for Π^1_2 theories in second order arithmetic. In et. al., T. A., editor, *Proceedings of the 10th Asian logic conference, Kobe, Japan, September 1-6, 2008*, pages 375–386. Hackensack, NJ: World Scientific.

[Zhao,] Zhao, Y. Szemerédi's Theorem via Ergodic Theory. Completed by the author in partial fulfillment of the requirements of the Part III Tripos in Mathematics at Cambridge University.

Appendices

A Excerpts from Szemerédi's Proof



The diagram represents an approximate flow chart for the accompanying proof of Szemerédi's theorem. The various symbols have the following meanings: $F_k \equiv$ Fact k, $L_k \equiv$ Lemma k, T $\equiv$ Theorem, C $\equiv$ Corollary, D $\equiv$ Definitions of B, S, P, α, β, etc., $t_m \equiv$ Definition of t_m, vdW $\equiv$ van der Waerden's theorem, $F_0 \equiv$ "If $f: \boldsymbol{R}^+ \to \boldsymbol{R}^+$ is subadditive then $\lim_{n\to\infty} \frac{f(n)}{n}$ exists".

Lemma A.1. *(Szemerédi Regularity Lemma; Lemma 1 in [Szemerédi, 1975]) Let A and B be disjoint sets, let I be a fixed subset of $[A,B] := (\{x,y\} : x \in X, y \in Y)$, $k(u) = \{v \in A \cup B : \{u,v\} \in I\}$, and $\beta(X,Y) = k(X,Y)|X|^{-1}|Y|^{-1}$. For all $\epsilon_1, \epsilon_2, \delta, \rho, \sigma$, there*

exist m_0, n_0, M, N*, such that for all* I *with* $|A| = m > M$*,* $|B| = n > N$*, there exist disjoint* $C_i \subseteq A$*,* $i < m_0$*, and, for each* $i < m_0$*, disjoint* $C_{i,j} \subseteq B, j < n_0$*, such that:*

1. $|A - \cup_{i<m_0} C_i| < \rho m$*,* $|B - \cup_{j<n_0} C_{i,j}| < \sigma n$ *for any* $i < m_0$*;*

2. *For all* $i < m_0$*,* $j < n_0$*,* $S \subseteq C_i$*,* $T \subseteq C_{i,j}$*, with* $|S| > \epsilon_1 |C_i|$*,* $|T| > \epsilon_2 |C_{i,j}|$*, we have*

$$\beta(S, T) \geq \beta(C_i, C_{i,j}) - \delta.$$

3. *For all* $i < m_0$*,* $j < n_0$ *and* $x \in C_i$*,*

$$|k(x) \cap C_{i,j}| \leq (\beta(C_i, C_{i,j}) + \delta)|C_{i,j}|.$$

B The Metamathematics of Szemerédi's Theorem and Ergodic Theory

Here I provide a summary of what is known concerning the metamathematics of Szemerédi's theorem and various of its proofs.[109] It is important to note that most of these results should be understood as "confident claims," unless otherwise noted, based upon an examination of bounds and resources involved in the arguments. However, in order to rigorously justify that a theorem can be formalized in some weak theory T, one must carefully write down all definitions involved encoded in the language of second-order arithmetic. This task can be very tedious (and may be non-trivial in some cases), and it has not been done for the theorems considered here.[110]

Claim B.1. Szemerédi's original proof in [Szemerédi, 1975] can be formalized in RCA_0. Utilizing bounds by Gowers, this can likely be gotten down to elementary function arithmetic (EFA).

I have not discussed Timothy Gowers's proof of Szemerédi's theorem, but I cite its metamathematical data for completeness:

Claim B.2. Gowers's proof of Szemerédi's theorem in [Gowers, 2001] can be formalized in $I\Delta_0(\exp)$. It is likely that it can be formalized in RCA_0 as it has elementary bounds and is an entirely combinatorial argument.

Let us now turn to the ergodic theorems employed in Furstenberg's proof(s) ([Furstenberg, 1977] and [Furstenberg et al., 1982]) of Szemerédi's theorem. The axiomatic strength of these proofs will turn upon the strength of the Furstenberg Structure Theorem (Theorem 2.3.37) and whether the full power of this theorem is required. There is a rather complicated story to tell here. It was originally claimed in [Avigad, 2009] that the Structure Theorem was equivalent (over ACA_0) to Π^1_1-CA_0. However, a few years later, it was noted in [Montalbán, 2011] that Avigad and Towsner were comfortable asserting only the formalizability of the Structure Theorem in Π^1_1-CA_0. The reversal, i.e., that the Structure Theorem + ACA_0 implies Π^1_1-CA_0, was—and remains—open. This is because, though one might try to straightforwardly apply the techniques of [Beleznay and Foreman, 1996] to get the reversal, a more delicate approach is required to get this to go through ACA_0. Thus, we have

Claim B.3. The Furstenberg Structure Theorem (Theorem 2.3.37) can be formalized in Π^1_1-CA_0.

Conjecture B.4. The reversal holds.

Now let us assess what is known about the ergodic proofs of Szemerédi's theorem. The crucial question will be: How far into the countable ordinals need the construction of the maximal distal factor of arbitrary measure-preserving system X extend? That is, need

[109] This information was gathered from personal correspondence with Jeremy Avigad. He very generously fielded my questions and provided references, unpublished notes, and past email communications for me to utilize.

[110] I take it with the exception of the formalized of the Furstenberg Structure Theorem in Π^1_1-CA_0 as this was part of the work done by Avigad and Towsner. They are confident in this particular direction of the equivalence claimed in Theorem 5.3 of [Avigad, 2009].

it extend arbitrarily far and thus require the *full* Furstenberg Structure Theorem? It is relatively well known that Furstenberg's original proof did not. However, the presentation in [Furstenberg et al., 1982] does appear to use the full Structure Theorem. Nonetheless, Avigad and Towsner show in [Avigad and Towsner, 2010] that the maximal distal factor need only extend to the ω^{ω^ω}th level.[111] In short we can say:

Claim B.5. Furstenberg's original proof in [Furstenberg, 1977], when proved *for each* k, can be formalized in ACA_0. When proved for all k, the proof goes slightly beyond ACA_0. The same is true for the proof given in [Furstenberg et al., 1982].

The ergodic proof(s) in both [Furstenberg, 1977] and [Furstenberg et al., 1982] are then in principle much weaker than they appear because they do not *require* the full Furstenberg Structure Theorem. But this raises an interesting question: why didn't Furstenberg *et. al.* in [Furstenberg et al., 1982] simply avoid using this much power? As Avigad noted in personal correspondence, "I am sure they knew that it was possible. It would have changed the presentation only slightly: they could throw away the limit argument for the SZ property and the appeal to transfinite induction, and then they only needed to modify one of their calculations slightly." Thus, the answer is, probably, that they did not care to do so: the Structure Theorem is an incredibly interesting result and provides understanding as to why the ergodic analogue of Szemerédi's theorem holds. Why then dispense with it or try to whittle away at its logical strength if it provides a perspicuous proof?

Thus, I believe the following morals can be drawn. Even though the ergodic proofs are not as axiomatically strong as they appear, they are still much stronger than Szemerédi's original proof. We should emphasize the *relative distance*, which is nicely precisified by reverse mathematical analysis: Szemerédi's combinatorial proof is incredibly weak but also incredibly difficult to understand. On the other hand, Furstenberg's proof(s), especially that of [Furstenberg et al., 1982], takes us from RCA_0 (possibly even EFA) to just beyond ACA_0 and thus into the realm of the infinitary. In so doing, we get a perspicuous high-level proof of Szemerédi's theorem that emphasizes crucial structural features of the mathematics. Is this a consequence of the increase in axiomatic strength? The questions about logical strength may actually be the less interesting part of the story here. Perhaps what is more interesting and most germane to the explanatory advantages of the ergodic techniques is the way in which these techniques package and modulate the information about structure and randomness present in the combinatorial setting. We have seen that the Furstenberg Structure Theorem gives us a direct means of presenting the dichotomy, and this generates a significant gain in clarity over the combinatorial proof. (For instance, we can avoid the very delicate structuring of generalized arithmetic progressions sketched in Section 2.5.4.) In short, the passage to infinite spaces, factors, and extension maps allows us to avoid a good deal of complicated, highly non-linear, combinatorial work that obfuscates our understanding of Szemerédi's theorem. When we do pass to these resources, naturally we get an attendant increase in axiomatic strength; however, this strength is a by-product of the mathematical concepts we "needed" for an explanatory proof and not the cause.[112]

[111] This is actually quite curious as Furstenberg's original proof makes do with even less than this.

[112] Thanks very much to Jeremy Avigad for emphasizing this point so clearly to me.

C Reverse Mathematics

Here I define various formal systems that occur throughout the paper. I begin with the very weak (and first-order) *Primitive Recursive Arithmetic* (PRA) and proceed to subsystems of second-order arithmetic. I primarily follow the canonical presentation of [Simpson, 1999], but deviate at some points for ease of exposition, e.g., I try to give a more natural presentation of PRA below. At the beginning of each subsection, I provide a very brief motivation for the subsystem defined, note its philosophical background, and mention some of the mathematics that can be done in the subsystem.

Let me first provide some basic notions for the reader less familiar with mathematical logic. A *formal system* is defined by specifying a formal language and axioms. We say that any formula of the given language deducible by logic from the axioms is a *theorem* of the given formal system. Here our logic is classical and either first-order (for PRA) or second-order (for the remaning systems). A subsystem T' of a formal system T is itself a formal system in the given language whose axioms are theorems of T. Finally, we say that T_2 is a conservative extension of T_1 iff for every sentence $\varphi \in \mathsf{L}_{\mathsf{T}_1}$ such that $\vdash_{\mathsf{T}_2} \varphi$ we have $\vdash_{\mathsf{T}_1} \varphi$. Informally, we say that T_2 proves no new theorems of T_1.

C.1 First-Order Arithmetic and PRA

Our language is that of first-order arithmetic written as L_1.

Definition C.1. (Language of First-Order Arithmetic; L_1)

1. *Variables*: there is one sort of variable called *number variables* which range over the natural numbers $\omega = \{0, 1, 2, \ldots\}$. These are written as lower-case Roman letters $i, j, k, \ldots$;
2. *Constant and Function Symbols*: the constant symbol 0, the successor symbol 1, binary operation symbols $+$ and $\cdot$ for addition and multiplication, respectively. We also have function symbols for primitive recursive functions,[113] all of which can be explicitly defined in terms of $\{0, 1, +, \cdot\}$;
3. *Numerical Terms*: these are constructed from number variables and 0 by closing under the successor and other primitive recursive functions;
4. *Atomic Formulas*: equations $t_1 = t_2$ between numerical terms t_1, t_2;
5. *Formulas*: these are constructed from atomic formulas using the usual propositional connectives $\{\wedge, \vee, \neg, \rightarrow, \leftrightarrow\}$ and number quantifiers $\forall n, \exists n$.

Now we can provide the axioms for first-order arithmetic. Our logic is the classical first-order predicate calculus with equality.

[113] I define the language (and below the axioms) of first order arithmetic (PA) in this way so that we can see PRA is straightforwardly included in PA. See [Feferman, 1998b] for a similar formulation.

Definition C.2. (First Order Arithmetic; Z_1; PA)

1. (Number-Theoretic + Primitive Recursive Axioms):
   (a) $n + 1 \neq 0$;
   (b) $m + 1 = n + 1 \to m = n$;
   (c) $m + 0 = m$;
   (d) $m + (n + 1) = (m + n) + 1$;
   (e) $m \cdot 0 = 0$;
   (f) $m \cdot (n + 1) = (m \cdot n) + m$;
   (g) defining equations for each additional primitive recursive function.
2. (Induction Scheme)

$$(\varphi(0) \wedge \forall n(\varphi(n) \to \varphi(n+1))) \to \forall n\, \varphi(n) \tag{C.1}$$

   for each formula φ in L_1.

Having defined PA in this way, we get PRA very easily:

Definition C.3. (Language of Primitive Recursive Arithmetic; $\mathsf{L}_{\mathsf{PRA}}$) This is simply the quantifier-free part of L_1.

Definition C.4. (Primitive Recursive Arithmetic; PRA)

1. Number-Theoretic + Primitive Recursive Axioms;
2. (Quantifier-Free Induction Rule; $\mathsf{IR}_{\mathsf{QF}}$):

$$\text{From } \varphi(0) \text{ and } \varphi(n) \to \varphi(n+1), \text{ deduce } \varphi(n) \tag{C.2}$$

   for each quantifier-free φ.

PRA plays an important role in the arguments concerning simplicity and impurity in Section 2.5. It is also of philosophical interest as William Tait has argued that PRA serves as a reasonable precisification of "finitary reasoning" in Hilbert's program.[114] See the subsection on WKL_0 below for a brief discussion of Hilbert's Program and its contemporary, "relativized" iterations.

C.2 Second-Order Arithmetic and Its Subsystems

Our language is that of second-order arithmetic written as L_2. This is formed by adding set variables and quantifiers (or function variables and quantifiers) to L_1 along with a new binary relation symbol $\in$ between numbers and sets. I explicitly record the axioms for the convenience of the reader (here I follow [Simpson, 1999] closely):

[114] See [Tait, 1981].

Definition C.5. (Language of Second-Order Arithmetic; L_2)

1. *Variables*: There are two sorts of variables in L_2:
   (a) Number variables ranging over the natural numbers $\omega = \{0, 1, 2, \ldots\}$. These are written as lower-case Roman letters $i, j, k, \ldots$;
   (b) Set variables ranging over the full powerset of ω. These are written as upper-case Roman letters $X, Y, Z, \ldots$.
2. *Constant and Function Symbols*: the constant symbol 0, the successor symbol 1, binary operation symbols $+$ and $\cdot$ for addition and multiplication, respectively.
3. *Numerical terms*: these are constructed from number variables and 0 by closing under the successor, addition, and multiplication;
4. *Atomic formulas*: $t_1 = t_2$, $t_1 < t_2$, and $t_1 \in X$ with t_1, t_2 numerical terms and X a set variable. These formulas have the obvious intended meaning.
5. *Formulas*: these are constructed from atomic formulas using the usual propositional connectives $\{\wedge, \vee, \neg, \rightarrow, \leftrightarrow\}$; number quantifiers $\forall n, \exists n$; and set quantifiers $\forall X, \exists X$.

A sentence is a formula with no free variables.

Now we can provide the axioms of second-order arithmetic. Our logic is now the classical two-sorted predicate calculus with equality in L_1.

Definition C.6. (Second-Order Arithmetic; Z_2)

1. (Number-theoretic axioms):
   (a) $n + 1 \neq 0$;
   (b) $m + 1 = n + 1 \rightarrow m = n$;
   (c) $m + 0 = m$;
   (d) $m + (n + 1) = (m + n) + 1$;
   (e) $m \cdot 0 = 0$;
   (f) $m \cdot (n + 1) = (m \cdot n) + m$;
   (g) $\neg m < 0$;
   (h) $m < n + 1 \leftrightarrow (m < n \vee m = n)$.
2. (Induction Scheme; $\mathsf{R\text{-}IND}$):

$$\forall X((0 \in X \wedge \forall n(n \in X \rightarrow n + 1 \in X)) \rightarrow \forall n(n \in X)). \tag{C.3}$$

3. (Comprehension):

$$\exists X \forall n (n \in X \leftrightarrow \varphi(n)) \tag{C.4}$$

where $\varphi(n)$ is any L_2-formula in which X does not occur freely. Note that $\varphi(n)$ may contain other free variables besides n. Intuitively, the comprehension scheme asserts that there is a set X defined as the set of all n such that $\varphi(n)$ holds.

Finally, before proceeding to particular subsystems, I define classes of formulas that commonly appear in conservation results:

Definition C.7. (Σ^0_1 and Π^0_1 formulas) We say that an L_2-formula[115] φ is Σ^0_1 if it can be written as $\exists n\, \theta$ where n is a number variable and θ is a bounded quantifier formula.[116] Similarly, a formula φ is Π^0_1 if it can be written as $\forall n\, \theta$, where n is a number variable and θ is a bounded quantifier formula. Finally, we say a formula is Δ^0_1 if it is both Σ^0_1 and Π^0_1.

This definition generalizes to:

Definition C.8. (Σ^0_k and Π^0_k formulas) For $0 \leq k \in \omega$, a formula φ is said to be Σ^0_k (respectively Π^0_k) if it is of the form $\exists n_1 \forall n_2 \exists n_3 \cdots n_k\, \theta$ (respectively $\forall n_1 \exists n_2 \forall n_3 \cdots n_k\, \theta$) with $n_1, \ldots, n_k$ numerical variables and θ a bounded quantifier formula. Finally, we say a formula is Δ^0_k if it is both Σ^0_k and Π^0_k.

Definition C.9. (Arithmetical Formula) A formula φ is said to be *arithmetical* if it contains no set quantifiers (all quantifiers appearing in φ are numerical).

Definition C.10. (Σ^1_1 and Π^1_1 formulas) We say that a formula φ is Σ^1_1 if it can be written as $\exists X\, \theta$ where X is a set variable and θ is an arithmetical formula. Similarly, a formula φ is said to be Π^1_1 when it is of the form $\forall X \theta$ for X a set variable and θ an arithmetical formula. Finally, a formula is Δ^1_1 if it is both Σ^1_1 and Π^1_1.

This generalizes to:

Definition C.11. (Σ^1_k and Π^1_k formulas) For $0 \leq k \in \omega$, a formula φ is said to be Σ^1_k if it can be written as $\exists X_1 \forall X_2 \exists X_3 \cdots X_k\, \theta$ for $X_1, \ldots, X_k$ set variables and θ arithmetical. Similarly, a formula φ is Π^1_k if it can be written as $\forall X_1 \exists X_2 \forall X_3 \cdots X_k\, \theta$ for $X_1, \ldots, X_k$ set variables and θ arithmetical. Finally, a formula is Δ^1_k if it is both Σ^1_k and Π^1_k.

Unless otherwise noted, the subsystems considered below will all consist of the number-theoretic axioms, a restricted induction scheme,[117] and a specific comprehension scheme.

[115] Moving forward, I will simply assume that all formulas are L_2-formulas and drop the qualification. I will explicitly remark upon any formula that is not L_2.

[116] A bounded quantifier formula is a formula φ such that all quantifiers occurring in φ are bounded quantifiers. A bounded (numerical) quantifier is one of the following: $\forall n < t$, $\forall n \leq t$, $\exists n < t$, $\exists n \leq t$.

[117] More precisely, Σ^0_1 induction, which is defined below. This is stronger than the induction scheme, $\mathsf{R\text{-}IND}$, given in the definition of Z_2, but weaker than full second order induction. [Simpson, 1999] moves back and forth between various induction schemes, which is unnecessary. Fixing our induction scheme as Σ^0_1 from the start does not affect the following results in any meaningful way. Also, please see the concluding remark of this section.

That is, only the set existence axiom[118] will differ in each subsystem. This is, however, very significant as the goal of the reverse mathematical program is to ascertain the set existence axioms necessary to prove theorems of "ordinary" (or non-set-theoretic) mathematics. It turns out that, in some cases, if a theorem of ordinary mathematics can be proved from an appropriate set existence principle, then the theorem will be equivalent (over some weaker subsystem) to the set existence principle. More precisely, consider some theorem of "ordinary" mathematics τ. One then formalizes this theorem as $\widetilde{\tau}$ (in Z_2) and shows that $\widetilde{\tau}$ is not provable in some weak base theory T_1. Then one shows that some stronger extension T_2 of T_1 proves $\widetilde{\tau}$. This gives one direction of the equivalence. Next take $\widetilde{\tau} + \mathsf{T}_1$ and show that this proves the axioms of T_2 (this direction is often called a "reversal"). Thus, conclude the axioms of T_2 and $\widetilde{\tau}$ are equivalent over T_1.

Of course, there are infinitely many subsystems of Z_2; however, it is quite interesting that five subsystems ($\mathsf{RCA}_0, \mathsf{WKL}_0, \mathsf{ACA}_0, \mathsf{ATR}_0, \Pi^1_1\text{-}\mathsf{CA}_0$) occur repeatedly in the context of such "reversals." I define these subsystems (along with a few others) in the order of increasing axiomatic strength. Indeed, from the model theory of these subsystems we have an ascending chain of proper subsystems:

$$\mathsf{RCA}_0 \subset \mathsf{WKL}_0 \subset \mathsf{ACA}_0 \subset \mathsf{ATR}_0 \subset \Pi^1_1\text{-}\mathsf{CA}_0\,. \tag{C.5}$$

Remark C.12. (Induction Schemes) I have followed [Simpson, 1999] in recording R-IND as the induction scheme for Z_2. Note that this is weaker than the full second order induction scheme (IND):

$$(\varphi(0) \wedge \forall n(\varphi(n) \to \varphi(n+1))) \to \forall n\, \varphi(n) \tag{C.6}$$

for φ any L_2-formula. However, from the unrestricted comprehension axiom (given in the definition of Z_2) and R-IND we can derive IND.

Whenever arbitrary subsystem X is written with a 0-subscript, i.e., X_0, this indicates the use of some form of restricted induction. The most commonly discussed subsystems of Z_2 use an induction scheme called Σ^0_1-IND, which is stronger than R-IND, but much weaker than IND. The reader may assume, for the sake of uniformity, that all subsystems written with a 0-subscript have Σ^0_1-IND as their induction scheme. However, whenever a subsystem is written without such a subscript, it will employ full induction IND. This only becomes relevant for our discussion of Δ^1_1-CA below, which features in my discussion of [Feferman, 1987] in Section 2.6 of the main paper. The full induction scheme in this system is of particular importance because it allows for $\Delta^1_1\text{-}\mathsf{CA} \vdash \mathsf{Con}(\mathsf{PA})$. This is all quite subtle, and most readers need not concern themselves with this.

[118] I identify the comprehension schemes as set existence principles, but this identification would not be accepted by all. See [Dean and Walsh, 2017], Section 6 and [Eastaugh, 2019]. For our purposes, at least, very little turns upon this.

RCA_0

The first subsystem of Z_2 we consider is *Recursive Comprehension* written as RCA_0. This is a rather weak system; it may be thought to correspond in a suitably loose sense[119] to Bishop's development of constructive analysis[120]. RCA_0 is especially important in the reverse mathematical program as it usually serves as the appropriate base theory in which equivalences (in the sense above) are proved. Simply put, many results of reverse mathematics are theorems of RCA_0.

Informally, RCA_0 contains the following axioms: the number theoretic axioms given above; Σ^0_1-induction; a set existence axiom asserting the existence of Δ^0_1 (recursive/computable) sets. We require the following definition:

Definition C.13. (Σ^0_1-induction) The Σ^0_1-induction scheme, Σ^0_1-IND, is the restriction of the full second-order induction scheme IND to formulas $\varphi(n)$ which are Σ^0_1. That is, we take the universal closure of

$$(\varphi(0) \wedge \forall n(\varphi(n) \rightarrow \varphi(n+1))) \rightarrow \forall n(\varphi(n)) \tag{C.7}$$

where $\varphi(n)$ is any Σ^0_1 formula.

We now turn to the relevant set existence axiom:

Definition C.14. (Δ^0_1 Comprehension) The Δ^0_1 comprehension scheme consists of the universal closures of all formulas of the form

$$\forall n(\varphi(n) \leftrightarrow \psi(n)) \rightarrow \exists X \forall n(n \in X \leftrightarrow \varphi(n)), \tag{C.8}$$

where $\varphi(n)$ is any Σ^0_1 formula, $\psi(n)$ is any Π^0_1 formula, n is a numerical variable, and X is a set variable not freely occurring in $\varphi(n)$.

Definition C.15. (RCA_0) RCA_0 is the subsystem of Z_2 consisting of the number-theoretic axioms of Z_2, Σ^0_1-IND, and Δ^0_1 comprehension.

An important conservation result for RCA_0 is the following:

Theorem C.16. RCA_0 *is conservative over* PRA *for* Π^0_2 *sentences.*

This follows from Parsons's result[121] that $I\Sigma_1$ is conservative over PRA for Π^0_2 sentences and the fact that RCA_0 and $I\Sigma_1$ prove the same first-order sentences. See [Simpson, 1999], pp. 369 for the latter result.

[119] For instance, mathematics done in RCA_0 utilizes the law of excluded middle and the meanings of propositional connectives and quantifiers are classical. The constructivists would not countenance either of these.

[120] See [Bishop and Bridges, 1985].

[121] See [Parsons, 1970].

WKL_0

Next define WKL_0, the subsystem of Z_2 consisting of RCA_0 and a set existence principle called *Weak König's Lemma*. It is important to note that WKL_0 is much stronger than RCA_0 and is able to accommodate many theorems of "ordinary" mathematics.[122] WKL_0 is of non-trivial philosophical significance as it enables what Simpson has called a "partial realization" of Hilbert's program. Broadly speaking, Hilbert sought to justify all classical, infinitary mathematics in terms of epistemically privileged finitary reasoning, where "finitary" means appealing to nothing but "extralogical concrete objects that are intuitively present as immediate experience prior to all thought" ([Hilbert, 1967], 376). Hilbert did not make what he meant by "finitary" mathematically precise, but it is commonly accepted that the system PRA is quite close to his intentions.[123] The program was to be carried out by formalizing the whole of classical mathematics and then providing a finitary consistency proof of this formal system. Unfortunately, Gödel's Second Incompleteness Theorem dashed the hopes of a complete realization of the program.[124] However, two lessons of the reverse mathematical literature are: (i) many infinitary theorems of ordinary mathematics are equivalent to WKL_0 over RCA_0; (ii) there are conservation results for WKL_0 over PRA (see below). Thus, we get a partial finitary reduction in Hilbert's sense.[125]

Lemma C.17. *(Weak König's Lemma) Write $2^{<\mathbb{N}}$ for the full binary tree, i.e., the set of codes for finite sequences of 0's and 1's (in RCA_0). Then every infinite subtree T of $2^{<\mathbb{N}}$ has an infinite path.*[126]

Definition C.18. (WKL_0) The subsystem WKL_0 consists of RCA_0 and Weak König's Lemma.

An important conservation result is the following:

Theorem C.19. *(Harrington) WKL_0 is conservative over RCA_0 for Π^1_1 sentences.*

This is attributed to Harrington, but was first published in [Simpson, 1999], pp. 369-72. This theorem along with Parsons's result yields:

Theorem C.20. *(Friedman; unpublished) WKL_0 is conservative over PRA for Π^0_2 sentences.*

[122] See [Simpson, 1999], Chapters IV and IX.

[123] Following the argument in [Tait, 1981].

[124] In particular, infinitary mathematics is not conservative over PRA for Π^0_1 sentences. Interestingly, however, Gödel himself was unsure of the significance of his incompleteness results for Hilbert's Program (see his [Gödel, 1933]). He later concluded that his findings were in fact deadly to the original enterprise ([Gödel, 1958]).

[125] This interpretation of "relativizing" Hilbert's Program follows [Simpson, 1999]. However, another sort of relativization became available after the publication of [Gödel, 1933]. Here Gödel proved that one can translate PA into Heyting arithmetic (HA), which differs from PA only insofar as it does not employ the Law of Excluded Middle, i.e., the logic of HA is intuitionistic. This result explicitly satisfied one aim of Hilbert's Program: justify classical mathematics (or at least number theory) on constructive principles alone. Implicitly, because one could argue that HA utilizes merely a potential infinity of numbers, Gödel's translation also serves to eliminate the actual infinite from number theory. Thus, Hilbert's Program might be "relativized" by employing only constructive methods of justification for infinitary theorems.

[126] More explicitly, we write $T \subseteq \mathbb{N}$ (via coding finite sequences as natural numbers) and let a path be a function $f : \mathbb{N} \to \{0, 1\}$ such that for all $n \in \mathbb{N}$ we have $\langle f(0), \ldots, f(n-1)\rangle \in T$.

This is due to Friedman, but appears in [Simpson, 1999], pp. 381. It indicates that, from the perspective of mathematical logic, WKL_0 is weak; however, as discussed above, it accommodates a good deal of ordinary mathematical practice. Thus, this result is crucial for the partial realization of Hilbert's program.

Finally, I define an extension of WKL_0 known as WKL_0^+.[127]

Definition C.21. The subsystem WKL_0^+ is defined as WKL_0 along with the following axiom scheme

$$\forall n\, \forall \sigma\, \exists \tau(\tau \to \sigma \wedge \varphi(n, \tau)) \to \exists X\, \forall n\, \exists k(\varphi(n, X[k])) \tag{C.9}$$

where n is a numerical variable, σ and τ range over $2^{<\mathbb{N}}$, X ranges over $2^{\mathbb{N}}$, and φ is an arithmetical formula. Essentially, this scheme asserts that, given some sequence of arithmetically defined dense subsets of $2^{<\mathbb{N}}$, the sequence will have non-empty intersection.

Theorem C.22. *(Brown and Simpson; [Brown and Simpson, 1993]) The subsystem* WKL_0^+ *is conservative over* RCA_0 *for* Π^1_1 *sentences.*

This conservativity result along with the result of Parsons yields:

Theorem C.23. *(Brown and Simpson; [Brown and Simpson, 1993]) The subsystem* WKL_0^+ *is conservative over* PRA *for* Π^0_2 *sentences.*

ACA_0

The subsystem ACA_0 consists of RCA_0 along with a comprehension scheme asserting the existence of arithmetically definable sets. A good deal more mathematics can be developed in ACA_0 than in the preceding subsystems. In particular, we can formalize many central concepts of analysis, topology, and countable algebra. Note that, even though one can define the number systems $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ in RCA_0, there are crucial results of analysis that only become available in ACA_0, e.g., the Bolzano-Weierstrass theorem and the monotone convergence theorem.[128] ACA_0 also has an impressive philosophical pedigree. Weyl's *Das Kontinuum* (1918) seeks to develop a *predicative*[129] foundation for classical mathematics by using a system somewhat like ACA_0. In Feferman's words, "Weyl's main step, then, was to see what could be accomplished in analysis if one worked...only with the principle of arithmetical definition" ([Feferman, 1987], 173). This approach was further expanded upon and refined in the work of Kondô, Kriesel, and, of course, Feferman.[130]

[127] This was first studied in [Brown and Simpson, 1993].

[128] Indeed, Bolzano-Weierstrass and monotone convergence are equivalent to ACA_0 over RCA_0.

[129] Poincaré calls a definition *predicative*, "only if it excludes all objects that are dependent upon the notion defined" (quoted in [Mancosu, 1998], 68). The distinction between predicative and impredicative definitions (see below for the latter) can be quite slippery as there are many different formulations of it (some incompatible) in the mathematical and philosophical literature. Poincaré's gloss should suffice here.

[130] See [Mancosu, 1998], Part II, for an introduction to the historical and philosophical context of Weyl's program. [Dean and Walsh, 2017], Section 2, also contains a discussion of predicativity in the context of reverse mathematics. See this latter paper for many references concerning the development of predicativity in the 20th century.

Definition C.24. (Arithmetical Comprehension Scheme) The arithmetical comprehension scheme is the universal closure of

$$\exists X \forall n(n \in X \leftrightarrow \varphi(n)) \tag{C.10}$$

for $\varphi(n)$ an arithmetical formula in which X does not occur free.

Definition C.25. (ACA_0) ACA_0 consists of RCA_0 and arithmetical comprehension.

It should be noted that there is a very close relationship between ACA_0 and first order Peano arithmetic (PA). Indeed, PA is the first order part of ACA_0. Put differently

Theorem C.26. *(Folklore)* ACA_0 *is a conservative extension of* PA *for all arithmetical formulas.*

Δ^1_1-CA_0

This subsystem immediately follows ACA_0 in axiomatic strength. I dispense with it quickly, since it serves as a stepping stone to Δ^1_1-CA.

Definition C.27. (Δ^1_1 Comprehension) The Δ^1_1 comprehension scheme is the universal closure of

$$\forall n(\varphi(n) \leftrightarrow \psi(n)) \rightarrow \exists X \forall n(n \in X \leftrightarrow \varphi(n)) \tag{C.11}$$

where $\varphi(n)$ is any Σ^1_1 formula, $\psi(n)$ is any Π^1_1 formula, n is a number variable, and X is a set variable with X freely occurring in neither $\varphi(n)$ nor $\psi(n)$.

Definition C.28. (Δ^1_1-CA_0) Δ^1_1-CA_0 consists of ACA_0 and Δ^1_1 Comprehension.

It is quite surprising that, just as with ACA_0, we have the following conservation result:

Theorem C.29. *(Barwise and Schlipf)* Δ^1_1-CA_0 *is conservative over* PA *for all arithmetical formulas.*

This was first proved in [Barwise and Schlipf, 1975].

Δ^1_1-CA

Here I define Δ^1_1-CA. This occurs briefly at the beginning of Section 2.6 as there it is remarked that Feferman shows that Δ^1_1-$\mathsf{CA} \vdash \mathsf{Con}(\mathsf{PA})$ ([Feferman, 1987], 191-192). This system is much stronger than PA and stronger than Δ^1_1-CA_0 because its axioms include the full second order induction scheme IND.

Definition C.30. (Δ^1_1-CA) Δ^1_1-CA consists of Δ^1_1-CA_0, but using the full induction axiom IND in place of R-IND.

ATR_0

I define this subsystem for the sake of completeness; however, it does not feature in the discussion of the paper unlike the other systems presented here.[131] The reader may skip the details, if they wish.

ATR_0 consists of ACA_0 and a comprehension scheme called *arithmetical transfinite recursion*, which asserts the existence of sets defined by iterating arithmetical comprehension along countable wellorderings. This set existence axiom registers a significant increase in axiomatic strength as one can begin to prove theorems concerning ordinals, descriptive set theory, and infinitary combinatorics in ATR_0. From a philosophical perspective, ATR_0 can be thought of as contributing to the program of *predicative reductionism* (cf. Hilbert's finitistic reductionism). This is because ATR_0 is conservative over Feferman's system IR of predicative analysis[132] for Π^1_1 sentences, just as WKL_0 is conservative over PRA for Π^0_2 sentences.

We require the following definitions in order to get the comprehension scheme:

Definition C.31. A *countable linear ordering* is a structure $(A, <_A)$ with $A \subseteq \mathbb{N}$ and $<_A \subseteq \mathbb{N} \times \mathbb{N}$ an irreflexive linear ordering of A. That is, $<_A$ is transitive and only one of the following may hold for $a, b \in A$: $a = b \vee a <_A b \vee b <_A a$. A countable linear ordering $(A, <_A)$ is a *countable wellordering* if there is no sequence $\langle a_n : n \in \mathbb{N} \rangle$ of elements of A such that $a_{n+1} <_A a_n$ for all $n \in \mathbb{N}$.

Definition C.32. (Arithmetical Transfinite Recursion) Let $\theta(n, X)$ be an arithmetical formula with n and X free variables (note that $\theta(n, X)$ may contain additional free variables). Fixing any additional free variables, θ may be considered as an operator:

$$\Theta : \mathscr{P}(\mathbb{N}) \to \mathscr{P}(\mathbb{N}), \quad \Theta(X) = \{n \in \mathbb{N} : \theta(n, X)\}. \tag{C.12}$$

For any countable wellordering $(A, <_A)$, let Y be the set obtained by transfinitely iterating Θ along $(A, <_A)$. More precisely, Y is defined by the conditions: $Y \subseteq \mathbb{N} \times A$ and for each $a \in A$, $Y_a = \Theta(Y^a)$, where $Y_a = \{m : (m, a) \in Y\}$ and $Y^a = \{(n, b) : n \in Y_b \wedge b <_A a\}$. That is, for each $a \in A$, Y^a is formed by iterating Θ along an initial segment of $(A, <_A)$ up to and not including a, and Y_a is formed by applying Θ one final time.

The comprehension scheme *arithmetical transfinite recursion* asserts that such a set Y exists for every operator Θ and every $(A, <_A)$.

Definition C.33. (ATR_0) ATR_0 consists of ACA_0 and the comprehension scheme of arithmetical transfinite recursion.

Π^1_1-CA_0

Finally, we arrive at the strongest of the most commonly studied subsystems of second order arithmetic: Π^1_1-CA_0. This system is of particular interest for us as its axioms are equivalent,

[131] See subsection I.11, Chapter V, and Chapter IX.5 of [Simpson, 1999] for more information.
[132] See [Feferman, 1964] and [Feferman, 1968] for the formal development of predicative analysis.

over ACA_0, to the Furstenberg Structure Theorem.[133] As Simpson notes, subsystems below Π^1_1-CA_0 suffice to prove most of ordinary mathematics, but a few "exceptional theorems" require the full strength of Π^1_1-CA_0. His examples come from diverse subfields of mathematics (countable algebra, topology, countable combinatorics, descriptive set theory), but they all "directly or indirectly involve countable ordinal numbers" ([Simpson, 1999], 18). The Furstenberg Structure Theorem is no exception to the rule.[134]

Π^1_1-CA_0, unlike ATR_0 and ACA_0, cannot be justified on predicative grounds,[135] although, as noted above in Section 2.6, this system does have a constructive justification via $\mathsf{ID}^{\mathrm{i}}_{<\omega}$. Thus, Π^1_1-CA_0 represents an important juncture in the reverse mathematical hierarchy. Indeed, Π^1_1-CA_0 is expressly impredicative since we quantify over the full powerset of $\mathbb{N}$ when defining some particular subset of $\mathbb{N}$ (note the set quantifier in Π^1_1 formulas). Despite the fact that Π^1_1-CA_0 sits at a "conceptual fault line," it has received less philosophical attention than the forgoing subsystems, in part because it does not correspond to a classical foundational program. One question I would like to pursue is whether Π^1_1-CA_0 can play a distinctive epistemological role in the philosophy of mathematics, perhaps via an association of its axiomatic strength with structural features involving countable ordinals and explanatory power.[136]

Definition C.34. (Π^1_1 Comprehension) The Π^1_1 comprehension scheme is the universal closure of

$$\exists X \forall n (n \in X \leftrightarrow \varphi(n)) \tag{C.13}$$

for all Π^1_1 formulas $\varphi(n)$ in which X does not occur free.

Thus we have:

Definition C.35. (Π^1_1-CA_0) Π^1_1-CA_0 consists of RCA_0 and Π^1_1 comprehension.

[133] See [Avigad, 2009], Theorem 5.3.

[134] Indeed, Beleznay and Foreman show that the length of the shortest tower satisfying the Furstenberg Structure Theorem exhausts the set of countable ordinals. See [Beleznay and Foreman, 1996].

[135] Indeed, even Δ^1_1-CA can be predicatively justified. See [Feferman and Sieg, 1981].

[136] Moving forward, one case to consider might be Kondô's uniformization theorem for coanalytic sets. This is equivalent to Π^1_1-CA_0 over ATR_0 ([Simpson, 1999], 225). Other interesting cases to investigate deal with the structure theory of abelian groups.

D Inflationary Theory

Let me first give an extremely coarse overview of the substance of inflationary theory, after which I show that the infinitude of the universe does not follow so easily from it (for a full discussion see [Ellis and Stoeger, 2009]). The most basic (and least ontologically committed) form of inflationary theory proposes that *prior to* the formation of our universe at the time of the Big Bang, there was extant matter called a "false vacuum." It is supposed that this false vacuum is extremely high energy and is characterized by strong repulsive gravitational effects. Because of these gravitational effects, an inflationary period of accelerating expansion would then precede the Big Bang. The inflationary period would terminate when the unstable false vacuum decays into a "genuine" vacuum, i.e., at the time of the Big Bang. Though this is, of course, speculative, inflationary theory has gained rather wide acceptance due to its explanatory power. A popular version is the so-called "multiverse" or "chaotic" inflationary theory in which it is supposed that inflation ends at different times in different places in the ambient "super-universe." The inflationary process generates many (perhaps infinitely many) "bubble universes" (of which our universe would be one) embedded in the continuously expanding "super-universe." This proposal in turn generates scores of philosophical puzzles and difficulties (indeed, I would be inclined to handle multiverse talk with extreme caution), but we only require this extremely general overview for our purposes.

The substance of my dispute with the claims in [Knobe et al., 2006] can be seen by considering Section II, "Bubble Geometry," of [Vilenkin and Winitzki, 1997]. Here the authors compute the three-volume of a hypersurface (a "bubble universe") produced by the inflationary process. They first consider the false vacuum metric, given by

$$ds^2 = -dt^2 + \exp(2H_0 t)[dr^2 + r^2 d\Omega^2] \tag{D.1}$$

with $d\Omega^2 := d\theta^2 + \sin^2\theta\, d\phi^2$, $H_0 = \sqrt{8\pi V_0/3}$, and V_0 the potential of the false vacuum. At the "moment of nucleation," i.e., when the false vacuum decays into a genuine vacuum, particular "bubble universes" form and evolve according to the equations of the FLRW model. In particular, we get the metric

$$ds^2 = -d\tau^2 + a^2(\tau)[d\xi^2 + \sinh^2\xi d\Omega^2] \tag{D.2}$$

The scale factor $a^2(\tau)$ will not matter for our purposes, so I ignore it. It is argued that, at the time of nucleation, we can relate the metrics of the "super-universe" and a particular FLRW universe by setting the time variables $t = \tau = 0$. But now what are the values of the spatial coordinates r and ξ? This is the crucial point because, depending on what we say about the value of r at the time of nucleation, the FLRW universe generated by inflation will be either actually infinite or merely tending to infinity. [Vilenkin and Winitzki, 1997] set $r = r_0$ when the false vacuum nucleates into a true vacuum, i.e., when the super-universe produces an FLRW universe. This means that the radial position is assumed to be a *point* and thus is assumed to have no spatial extent. The effect on Equation D.1 is easy to see: all spatial components simply vanish and so $ds^2 = -dt^2$. Thus, at the time of nucleation, we simply have $s = t = 0$ and $ds^2 = 0$. This then defines a future light cone that extends to infinity "in one direction." [Vilenkin and Winitzki, 1997] describe this result as "the surface

$\tau = 0$ is the future light cone of that center..." and so "the boundary of the bubble can be approximated by this light cone" (549). The essential point is that all of space, infinitely much of it, comes into existence instantaneously at the time of nucleation.

I wish to emphasize that this will only be true provided that the initial radial position r of the universe is assumed to be *point-like*, i.e., without extension. And this seems to be an artifact of the model in question; what is currently known of quantum cosmology indicates that the initial radial position is very small but still extended.[137] Thus, this assumption and its consequences seem somewhat dubious pending further evidence. Thus, it is not true that inflationary theory immediately implies that the universe is infinite.

[137] See [Freivogel et al., 2006] and discussion in [Ellis and Stoeger, 2009].

E Alpha-Theory and Numerosities

E.1 Properties of Counting Systems

Definition E.1. A counting system $(\mathfrak{U}, \mathfrak{N}, \mathfrak{n})$ satisfies the following properties:

1. $\mathfrak{U}$ is closed under set inclusion, disjoint union, and Cartesian products.
2. If $A \in \mathfrak{U}$ is finite, then $\mathfrak{n}(A) = |A|$ is the number of elements of A. ($\mathfrak{n}$ agrees with cardinalities in finite sets.)
3. If $A \subseteq B$, then $\mathfrak{n}(A) \leq \mathfrak{n}(B)$ (monotonicity).
4. If $\mathfrak{n}(A) = \mathfrak{n}(A')$ and $\mathfrak{n}(B) = \mathfrak{n}(B')$, then $\mathfrak{n}(A \sqcup B) = \mathfrak{n}(A' \sqcup B')$.
5. If $\mathfrak{n}(A) = \mathfrak{n}(A')$ and $\mathfrak{n}(B) = \mathfrak{n}(B')$, then $\mathfrak{n}(A \times B) = \mathfrak{n}(A' \times B')$.
6. $\mathfrak{n}(\{P\} \times A) = \mathfrak{n}(A \times \{P\}) = \mathfrak{n}(A)$ for every singleton P.

E.2 The Axioms of Alpha-Calculus Theory

Axiom E.2. Every real-valued sequence $\varphi(n)$ has a unique alpha-limit, $\lim_{n\uparrow\alpha} \varphi(n)$.

Axiom E.3. If $c_r(n) = r$ is the constant sequence with value $r \in \mathbb{R}$, then $\lim_{n\uparrow\alpha} c_r(n) = r$.

Axiom E.4. The alpha-limit of the identity sequence is the "new" infinite number α, i.e., $\lim_{n\uparrow\alpha} n = \alpha \notin \mathbb{N}$.

Axiom E.5. The set of all alpha-limits of real-valued sequences

$$\mathbb{R}^* = \left\{ \lim_{n\uparrow\alpha} \varphi(n) : \varphi : \mathbb{N} \to \mathbb{R} \right\} \tag{E.1}$$

satisfies the field axioms. Furthermore, sums and products of alpha-limits are compatible with pointwise sums and products of sequences.

Finally, there is the *Qualified Set Axiom*, which, given some choice of infinite $Q \subseteq \mathbb{N}$, can be consistently added to the above axioms.[138] This axiom merits its own discussion; see below.

E.3 Labelled Sets and Alpha-Limits

Proof of Proposition 4.2.6: Assume that $\gamma_{\mathbf{A}} = \gamma_{\mathbf{B}}$, i.e., the functions coincide for all $n \in \mathbb{N}_0$. Now we establish that $\mathbf{A} \cong \mathbf{B}$. First we show the existence of a bijection. Let $\overline{A_n} = \{a \in A : \ell_A(a) = n\}$ and $\overline{B_n} = \{b \in B : \ell_B(b) = n\}$ for all n. Then

$$\left|\overline{A_0}\right| = \gamma_{\mathbf{A}}(0) = \gamma_{\mathbf{B}}(0) = \left|\overline{B_0}\right|. \tag{E.2}$$

[138] This is proved in Theorem 2.56 of [Benci and Nasso, 2019].

Then, for every subsequent n, we can obtain the cardinality of $\overline{A_n}$ and $\overline{B_n}$ by writing

$$\left|\overline{A_n}\right| = \gamma_{\mathbf{A}}(n) - \gamma_{\mathbf{A}}(n-1) = \gamma_{\mathbf{B}}(n) - \gamma_{\mathbf{B}}(n-1) = \left|\overline{B_n}\right|. \tag{E.3}$$

Thus, for every n, we have a bijection $\varphi_n : \overline{A_n} \to \overline{B_n}$. Taking the union of these, we get a bijection $\varphi : A \to B$ where $A = \bigcup_{n\geq 0} \overline{A_n}$ and $B = \bigcup_{n\geq 0} \overline{B_n}$. Now we must show that the labellings for $\mathbf{A}$ and $\mathbf{B}$ coincide, i.e., $\ell_B \circ \varphi = \ell_A$. But this is immediate from our construction, since $\ell_A(a) = n$ iff $a \in \overline{A_n}$. Applying our bijection φ to a, we see

$$\varphi(a) = \varphi_n(a) \in \overline{B_n} \iff \ell_B(\varphi(a)) = n. \tag{E.4}$$

Thus, by this chain of equivalences, $\ell_A(a) = n = \ell_B(\varphi(a))$.
Now assume that $\mathbf{A} \cong \mathbf{B}$, i.e., we have a bijection $\varphi : A \to B$ such that $\ell_B \circ \varphi = \ell_A$. Then, as we saw, $\overline{B_n} = \{b \in B : \ell_B(b) = n\} = \left\{\varphi(a) : a \in \overline{A_n}\right\}$ with $\overline{A_n}$ as above. Thus,

$$\left|\overline{A_n}\right| = \left|\overline{B_n}\right| \implies \gamma_{\mathbf{A}}(n) = \gamma_{\mathbf{B}}(n) \tag{E.5}$$

for all n.

Proof of Proposition 4.2.9: First let us prove: if $f(n) \neq g(n)$ for all n, then $\lim_{n\uparrow\alpha} f(n) \neq \lim_{n\uparrow\alpha} g(n)$ (Claim 1). All we require are the axioms of Alpha-calculus and basic properties of fields. For all n, define

$$h(n) = \frac{1}{f(n) - g(n)} \tag{E.6}$$

with $f(n) - g(n) \neq 0$. Taking α-limits, we obtain $\lim_{n\uparrow\alpha} h(n) \cdot (\lim_{n\uparrow\alpha} f(n) - \lim_{n\uparrow\alpha} g(n)) = 1$. Hence, $\lim_{n\uparrow\alpha} f(n) \neq \lim_{n\uparrow\alpha} g(n)$.

Now we prove: if $f(n) = g(n)$ for all but finitely many n, then $\lim_{n\uparrow\alpha} f(n) = \lim_{n\uparrow\alpha} g(n)$ (Claim 2). Consider the set on which $f(n)$ and $g(n)$ disagree, written as $\{n_1, \ldots, n_k\}$. Now note that

$$(f(n) - g(n)) \cdot (n - n_1) \cdots (n - n_k) = 0. \tag{E.7}$$

Taking alpha-limits yields

$$\left(\lim_{n\uparrow\alpha} f(n) - \lim_{n\uparrow\alpha} g(n)\right) \cdot (\alpha - n_1) \cdots (\alpha - n_k) = 0. \tag{E.8}$$

Since $\alpha \notin \mathbb{N}$, it must be the case that $\alpha - n_i \neq 0$ for all i. Thus, in order to get Equation E.8, we must have $\lim_{n\uparrow\alpha} f(n) = \lim_{n\uparrow\alpha} g(n)$, as desired.

Finally, we can prove Proposition 4.2.9. Consider the following sequence

$$g'(n) = \begin{cases} g(n) & f(n) \neq g(n) \\ g(n)+1 & \text{else.} \end{cases} \tag{E.9}$$

Then $f(n) \neq g'(n)$ for all n, so $\lim_{n\uparrow\alpha} f(n) \neq \lim_{n\uparrow\alpha} g'(n)$ by the proof of Claim 1. Furthermore, we see that $\lim_{n\uparrow\alpha} g'(n) = \lim_{n\uparrow\alpha} g(n)$ because $g(n) = g'(n)$ for all but finitely many n by the proof of Claim 2. Therefore, $\lim_{n\uparrow\alpha} f(n) \neq \lim_{n\uparrow\alpha} g(n)$, as desired.

E.4 Qualified Sets

In order to make sense of the discussion in Subsection 4.2.1 and Theorem 4.3.2 above, we require the notions of an "Alpha-measure" and "qualified set."

Definition E.6. The Alpha-measure is the function

$$\mu_\alpha : \mathscr{P}(\mathbb{N}) \to \{0,1\} \tag{E.10}$$

defined as

$$\mu_\alpha(A) = \begin{cases} 1 & \alpha \in A^* \\ 0 & \alpha \notin A^* \end{cases} \tag{E.11}$$

where $A^* \subseteq \mathbb{R}^*$ is the hyper-extension of set $A \subseteq \mathbb{R}$, i.e., the set of all alpha-limits of real sequences taking values in A.

We then have the following:

Definition E.7. A set $A \subseteq \mathbb{N}$ is qualified if $\mu_\alpha(A) = 1$. Write $\mathfrak{Q}$ for the family of qualified sets.

The importance of qualified sets is justified by the following theorem

Theorem E.8. *Let φ, ψ be real-valued sequences. Then the set $\{n \in \mathbb{N} : \varphi(n) = \psi(n)\}$ is qualified (or: $\varphi(n) = \psi(n)$ holds almost everywhere) iff*

$$\lim_{n\uparrow\alpha} \varphi(n) = \lim_{n\uparrow\alpha} \psi(n). \tag{E.12}$$

That is, the alpha-limit of a sequence depends only on the values of the sequence taken on a qualified set.

Proof. See [Benci and Nasso, 2019], p. 25-26. □

Using this notion of qualified sets, we can see how to extend properties enjoyed by natural numbers to properties of the corresponding hypernatural numbers.

Definition E.9. For property P of natural numbers, the property P^* of hypernatural numbers is satisfied by $\nu \in \mathbb{N}^*$ if $\nu \in \{n \in \mathbb{N} : P(n)\}^*$.

Thus, if we are interested in examining properties of α, we note that α satisfies P^* iff $\alpha \in \{n \in \mathbb{N} : P(n)\}^*$ iff $\{n \in \mathbb{N} : P(n)\}$ is qualified. This provides us with the flexibility that proved so useful in Theorem 4.3.2. That is, in order to validate fundamental intuitions about infinite subsets of $\mathbb{N}$, we wanted α to be a multiple of some $k \in \mathbb{N}$, our property P. This is achieved using the *Qualified Set Axiom* for appropriate choice of Q:

Axiom E.10 (Qualified Set Axiom (QSA))**.** The set Q is qualified, i.e., $\alpha \in Q^*$.

Thus, by choosing $Q = \{m \in \mathbb{N} : m!\}$ to be qualified by invoking QSA, we guarantee that $\alpha \in Q^*$, and thus that α satisfies P^*, viz., $\alpha/k \in \mathbb{N}^*$ for all $k \in \mathbb{N}$. Finally, it should be noted that, for any choice of infinite $Q \subseteq \mathbb{N}$, one can consistently add QSA for Q to the other axioms of Alpha-Calculus. Again, this is proved in Theorem 2.56 of [Benci and Nasso, 2019].

F Kolmogorov's Axioms

Classical probability is based upon the work of Kolmogorov. A *Kolmogorov probability theory* is a triple $(\Omega, \mathscr{F}, P)$ where Ω is a non-empty set of elementary events (the sample space), $\mathscr{F}$ is a σ-algebra[139] over Ω (the event space), and $P : \mathscr{F} \to \mathbb{R}$ is a probability function. That is, for some event $E \in \mathscr{F}$, P assigns to E the real-valued probability that E occurs. It is required that $(\Omega, \mathscr{F}, P)$ satisfy the following axioms:

Axiom F.1. (Positivity). For all $E \in \mathscr{F}$, $P(E) \geq 0$.

Axiom F.2. (Normalization). $P(\Omega) = 1$.

Axiom F.3. (Finite Additivity). For all $E_1, E_2 \in \mathscr{F}$ such that $E_1 \cap E_2 = \emptyset$,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2).$$

Axiom F.4. (Continuity). Let $A = \bigcup_{n \in \mathbb{N}} A_n$ where $\forall n \in \mathbb{N}, A_n \subseteq A_{n+1} \in \mathscr{F}$. Then

$$P(A) = \sup_{n \in \mathbb{N}} P(A_n).$$

Remark F.5. Requiring the Continuity Axiom of one's probability space is equivalent to requiring *Countable Additivity*: for all $E_1, E_2, \ldots \in \mathscr{F}$ such that $E_1 \cap E_2 \cap \cdots = \emptyset$, we have

$$P(E_1 \cup E_2 \cup \cdots) = P(E_1) + P(E_2) + \cdots \tag{F.1}$$

[139] $\mathscr{F} \subseteq \mathscr{P}(\Omega)$ is a σ-algebra over Ω if $\mathscr{F}$ is closed under complements, intersections, and countable unions.

G Non-Archimedean Probability (NAP) Axioms

The most obvious difference between a Kolmogorov probability theory and a non-Archimedean theory is that our probability function, now written as P_N, will have different a domain and range. In particular, one "problem" with Kolmogorov's theory is that the domain of P is not the full powerset of Ω, but rather a restricted family of subsets, the σ-algebra $\mathscr{F}$. Intuitively, however, one would expect a probability assignment to be given for every element of $\mathscr{P}(\Omega)$, including singletons. This intuition requires that we change the range of our probability function to a non-Archimedean field, viz., a field with infinitesimals. A triple (Ω, P_N, J) is called an NAP probability theory, where Ω is a non-empty set of elementary events, $P_N : \mathscr{P}(\Omega) \to \mathbb{F}^*$ is a function with $\mathbb{F}^*$ a non-Archimedean field, and J is an algebra homomorphism (to be defined). It is required that (Ω, P_N, J) satisfy the following axioms:

Axiom G.1. (Positivity). For all $E \in \mathscr{P}(\Omega)$, $P_N(E) \geq 0$.

Axiom G.2. (Normalization). For all $E \in \mathscr{P}(\Omega)$, $P_N(E) = 1 \Leftrightarrow E = \Omega$.

Axiom G.3. (Finite Additivity). For all $E_1, E_2 \in \mathscr{P}(\Omega)$ such that $E_1 \cap E_2 = \emptyset$,

$$P_N(E_1 \cup E_2) = P_N(E_1) + P_N(E_2).$$

Since in NAP theory the range of our probability function P_N is no longer $\mathbb{R}$ but rather some non-Archimedean field, we must replace Axiom F.4 with a version suitable for non-Archimedean fields. (This is because in a non-Archimedean field the existence of a supremum is not guaranteed.[140]) Thus, we also require that (Ω, P_N, J) satisfy:

Axiom G.4 (Non-Archimedean Continuity)**.** For all $E_1, E_2 \in \mathscr{P}(\Omega)$ with $E_2 \neq \emptyset$, let the conditional probability of E_1 be given by

$$P_N(E_1|E_2) = \frac{P_N(E_1 \cap E_2)}{P_N(E_2)}. \tag{G.1}$$

Then the following hold:

1. For all $\lambda \in \mathscr{P}^0_{\text{fin}}(\Omega)$, $P_N(E_1|\lambda) \in \mathbb{R}^+$, where $\mathscr{P}^0_{\text{fin}}(\Omega)$ is the set of finite subsets of Ω excluding the empty set.

2. There is an algebra homomorphism

$$J : \mathscr{F}(\mathscr{P}^0_{\text{fin}}(\Omega), \mathbb{R}) \to \mathbb{F}^* \tag{G.2}$$

where $\mathscr{F}(\mathscr{P}^0_{\text{fin}}(\Omega), \mathbb{R})$ is the algebra of all real functions on $\mathscr{P}^0_{\text{fin}}(\Omega)$ such that for all $E \in \mathscr{P}(\Omega)$

$$P_N(E) = J(\varphi_E) \tag{G.3}$$

with $\varphi_E(\lambda) = P_N(E|\lambda)$ for any $\lambda \in \mathscr{P}^0_{\text{fin}}(\Omega)$.

[140] Up to isomorphism, the only complete ordered field is $\mathbb{R}$.

As the authors of [Benci et al., 2013] realize, the meaning of the above axiom is far from evident. Thus, they prove a theorem that helps to establish the following intuition: the knowledge of the conditional probability of event E, relative to a suitably chosen family of sets (here $\lambda_n \in \mathscr{P}^0_{\text{fin}}(\Omega)$) provides knowledge of the probability of E.[141] In the classical setting:

$$P(E) = \lim_{n\to\infty} P(E|\Omega_n) \tag{G.4}$$

where $\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$. In the non-Archimedean setting:

$$P_N(E) = J(\varphi_E(\lambda)) = J(P_N(E|\lambda)), \tag{G.5}$$

and so the homomorphism J can be thought of as a sort of limit. We can develop this idea by writing

$$P_N(E) = J(\varphi_E(\lambda)) = \lim_{\substack{\lambda\in\mathscr{P}^0_{\text{fin}}(\Omega)\\ \lambda\uparrow\Omega}} \varphi_E(\lambda) = \lim_{\substack{\lambda\in\mathscr{P}^0_{\text{fin}}(\Omega)\\ \lambda\uparrow\Omega}} P_N(E|\lambda). \tag{G.6}$$

The limit is itself determined by the choice of an ideal $I_\Lambda \subset \mathscr{F}(\mathscr{P}^0_{\text{fin}}(\Omega), \mathbb{R})$ and so depends on an appropriate choice of a family of sets Λ.[142] Thus, we can express this limit more compactly as

$$J(\varphi) = \lim_{\lambda\in\Lambda} \varphi(\lambda), \tag{G.7}$$

and call this the Λ-limit.

[141] See Section 3.2 of [Benci et al., 2013]. See also [Benci et al., 2018], Sections 3.2 and 3.3. I follow the notation of the former paper.

[142] More precisely: the values φ assumes on Λ. See [Benci et al., 2013], p. 135, Theorem 16.

H Translation of §5 of Bertrand's *Calcul des Probabilités*

Procedure I:

> One can say: if one of the ends of the chord is known, this information does not change the probability; the symmetry of the circle does not allow any other influence, favorable or unfavorable, to be attached to the occurrence of the requested event.
>
> One of the ends of the chord being known, the direction [of the chord] must be determined at random. If one draws the two sides of the equilateral triangle, having as their vertex the given point, they form, between themselves and the tangent, three angles of 60 degrees. The chord, in order to be longer than the side of the equilateral triangle, must be found in the one of the three angles that is included between the other two. The probability that, of the three equal angles that can receive it, the direction is in that one [i.e., the middle angle] seems, by definition, equal to $\frac{1}{3}$.

Procedure II:

> One can also say: if one knows the direction of the chord, this information does not change the probability. The symmetry of the circle does not allow any other influence, favorable or unfavorable, to be attached to the occurrence of the requested event.
>
> The direction of the chord being given, it [the chord], in order to be longer than the side of the equilateral triangle, must intersect one or the other of the radii that comprise the perpendicular diameter in the half closest to the center. The probability that this is so seems, by definition, equal to $\frac{1}{2}$.

Procedure III:

> One can say yet again: to choose a chord at random is to choose its midpoint at random. In order that the chord be longer than the side of the equilateral triangle, it is necessary and sufficient that the midpoint be at a shorter distance from the center than the midpoint of the radius, that is, inside a circle four times smaller in area. The number of points located in the interior of an area four times fewer is four times fewer. The probability that the chord whose midpoint is chosen at random be greater than the side of the equilateral triangle seems, by definition, equal to $\frac{1}{4}$.

Conclusion:

> Among these three answers, which is the true one? None of the three is false, nor correct, but the question is ill-posed.