Infinitesimals for Metaphysics: Consequences for the Ontologies of Space and Time

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Graduate School of The Ohio State University

By
Patrick Reeder, MA
Graduate Program in Philosophy

The Ohio State University
2012

Dissertation Committee:
Stewart Shapiro, Advisor
Ben Caplan
Neil Tennant
Abstract

In this dissertation, I defend unorthodox conceptions of continuity: I argue that they’re both conceptually viable and philosophically fruitful. After a brief introduction in the first chapter, I argue in the second chapter that the standard conception of continuity—which comes to us from Georg Cantor and Richard Dedekind, and which uses the real numbers as a model—doesn’t satisfy all of pretheoretic intuitions about continuity and indeed that no conception of continuity does. This opens up conceptual room for unorthodox conceptions of continuity. In the second chapter, I argue that an unorthodox conception of continuity based on infinitesimals—numbers as small as infinity is large—provides the basis for a novel account of contact: of when two material bodies touch. In the third chapter, I argue that two other unorthodox conceptions of continuity provide the basis for novel solutions to Zeno’s paradox of the arrow.
Dedication

For Faye Bartlett Reeder, Ph.D. (1892-1973)
Acknowledgments

I would like to thank all those who have read and commented on significant portions of this dissertation: Scott Brown, Steven Brown, Wesley Cray, Justin D’Arms, Matthew Davidson, Salvatore Florio, Peter Forrest, Timothy Fuller, Alison Kerr, Teresa Kouri, Nicholaos Jones, Lindsey Mason, James McGlothlin, Michael Miller, Cathy Muller, Bradley Rettler, Tony Roy, David Sanson, Kevin Scharp, Timothy Schroeder, Lisa Shabel, Nathan Smith, Declan Smithies, William Taschek, Gabriel Uzquiano and Daniel Wilkenfeld. For those whose names I have forgotten, I hope they will accept my apologies.

I wish to extend a special thanks to my advisors, Ben Caplan, Philip Ehrlich, Stewart Shapiro and Neil Tennant. Specifically, I want to thank Ben for being especially assiduous and patient with some of my drafts. He has also saved me from making especially embarrassing metaphysical claims. Any remaining embarrassments are completely due to my own philosophical clumsiness, not for lack of help. Thanks to Phil Ehrlich of Ohio University for exposing me to a vast amount of literature on the continuum that I would have never seen otherwise. I am especially thankful for his drawing my attention to the work done by Paolo Giordano. I also appreciate his willingness to work with me, a student outside of his university. I want to thank Neil Tennant for encouraging this project from very early on and especially for continually calling on me to finish my work in the mathematics department at Ohio State. Without the math background, there would be numerous papers that I would have been unable to read or use in this project. Together, my advisors have read several
rounds of these chapters and poured many hours into helping me sharpen a number of the ideas here. Without them, this dissertation would not have come to be. I take full responsibility for any remaining errors.

Finally, I wish to offer my warmest gratitude to the chair of my committee, Stewart Shapiro. First, this project has its origins in a reading group we did in the Summer of 2007 on John Bell’s book, The Continuous and the Infinitesimal, [8]. I would not have ever known of the exciting landscape of infinitesimals were it not for him. More personally, he has sacrificed hundreds of hours on me alone, patiently meeting with me and reading endless pages and drafts. I cannot express my appreciation for the way he consistently balanced encouragement with incisive critique.

I dedicate my dissertation to my great grandmother, Faye Bartlett Reeder (1892-1973). She received her Ph.D. in History from Ohio State in 1937—exactly seventy-five years ago. Her dissertation was entitled The Evolution of the Virginia Land Grant System in the Eighteenth Century. While completing her dissertation, she raised three boys as a single mother. The youngest of these is my grandfather, David W. Reeder (1927-). I’m sorry I never met her. She was the original Reeder Buckeye, the original Reeder scholar.
Vita

2004...................................................................................B.A., Ohio State University
2007...................................................................................M.A., Ohio State University

Publications

“A Scientific Enterprise? Penelope Maddy’s Second Philosophy” (with Stewart Shapiro),

“Parts of Singletons” (with Ben Caplan and Chris Tillman), Journal of Philosophy,

Fields of Study

Major Field: Philosophy
Contents

Abstract ii

Dedication iii

Acknowledgments iv

Vita vi

List of Tables xi

List of Figures xii

1 Prologue 1

1.1 The Thesis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
1.2 Philosophical Method . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
1.3 Summary of Chapters . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

2 The Labyrinth of Continuity 8

2.1 Infinite Divisibility . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2 Conceptual Terrain . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
2.2.1 Top-Down or Bottom-Up . . . . . . . . . . . . . . . . . . . . . . . 16
2.2.2 Points on points . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
2.3 Dedekind’s Orthodox Conception . . . . . . . . . . . . . . . . . . . . . 22
2.3.1 Motivations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
2.3.2 Reasons for Concern ........................................... 25
2.3.3 Fully Formalized ................................................. 26
2.4 Indecomposability and Excluded Middle ....................... 30
  2.4.1 Motivation ..................................................... 31
  2.4.2 Reasons for Concern ........................................... 34
  2.4.3 Fully Formalized ............................................... 36
2.5 Infinitesimal Magnitudes ......................................... 40
  2.5.1 Motivations ..................................................... 42
  2.5.2 Reasons for Concern ........................................... 46
  2.5.3 Infinitesimals in The Hyperreals ............................ 49
  2.5.4 Infinitesimals in Giordano’s Fermat Reals ................ 53
2.6 Aristotelian Intuitions: Non-Supervenience .................... 57
  2.6.1 Motivation ..................................................... 58
  2.6.2 Reasons for Concern ........................................... 62
  2.6.3 Fully Formalized ............................................... 63
2.7 Conclusion ......................................................... 67

3 Zeno’s Arrow .......................................................... 69
  3.1 Zeno’s Arrow ..................................................... 71
  3.2 Denying other Premises .......................................... 73
    3.2.1 The Calculus and Premise (1) ............................... 73
    3.2.2 Aristotle and Premise (2) ................................... 75
  3.3 White on Zeno .................................................... 80
    3.3.1 Robinson’s Hyperreals ...................................... 80
    3.3.2 White against Zeno ......................................... 82
    3.3.3 Looking Elsewhere .......................................... 85
  3.4 Outline for a Solution .......................................... 86
List of Tables

4.1. Comparisons of Theories of Contact on *Desiderata* ......................... 145

4.2. Comparisons of Theories of Contact on Refined *Desiderata* ............... 147
List of Figures

2.1 Equilateral Construction from Euclid’s Elements. ......................... 14
2.2 Sequence of Regular Polygons Converging to a Circle .................. 42
2.3 Inscribed Triangles in a Regular Octagon and Circle.................... 44
2.4 Connection (α ⊙ β) Due to Convergence ................................. 66
3.1 Hyperreals magnified by an infinite factor .............................. 81
4.1 A Mixed Pair: the closed ball (left) includes the shared boundary point ... 113
4.2 An Open Pair: the shared boundary point is not included in either ball ... 114
4.3 A Closed Pair: Red and Green are always separated .................. 115
4.4 Sequence of Regular Polygons Converging to a Circle ................. 117
4.5 Discrete 3-4-5 Triangle. .................................................. 122
Chapter 1

Prologue

Only Geometry can provide a thread for the Labyrinth of the Composition of the Continuum ... and no one will arrive at a truly solid metaphysics who has not passed through that labyrinth.

—Leibniz, De usu geometriae

For scientific inquiry to play a role in revealing the meaning of a term, it seems that either this term must be one that belongs entirely to scientific theory (like color in quantum chromodynamics), or else there must be some pre-theoretic meaning of the term that scientific inquiry can help clarify. “Continuous” seems to be of the latter type, for which pre-theoretic intuitions are essential to a complete understanding of the term. Perhaps it’s a more difficult case than many others, because we don’t have any clear instances of actual discontinuous motion, but this just suggests that empirical investigation of the actual may be even less useful than it is in explaining the meanings of other terms, like “heat.”

—M. Colyvan and K. Easwaran, “Mathematical and Physical Continuity”

Chapter 1.1

The Thesis

In this dissertation, I defend unorthodox conceptions of continuity: I argue that they’re both conceptually viable and philosophically fruitful. Most specifically, I want to take very seriously conceptions of continuity that contain infinitesimal lengths in some way or another. Infinitesimal lengths $\lambda$ are so small that any integral number of $\lambda$’s stood end-to-end will still fall short of any given unit of measure. After sufficiently

---

1Quoted in Arthur, [119], xxiii.
2See Colyvan and Easwaran, [40], 90-1.
developing the notion(s) of an infinitesimal, I proceed to apply it (them) to a variety of metaphysical puzzles, including Zeno’s paradox of the Arrow and the problem of contact—whether two material bodies can touch. Each of these puzzles depends in one way or another on how continuity is to be understood. By introducing infinitesimals, I make available novel solutions to each.

Chapter 1.2

Philosophical Method

Allow me to begin with some very brief methodological remarks. First, I take this study to be essentially conceptual. Some studies concern an investigation into nature and such studies require significant empirical work. Studies of nature begin with the question, *What is the world like, independent of the various tools used by humans?* On the other hand, studies of concepts begin with what I will call here the **Core Question**: *What are the various tools used by humans like, independent of how the world is?* There are two major ways in which one may answer this latter question. The first way involves analyzing how the best, most current natural or mathematical sciences deploy a concept. This way has a significantly empirical, *a posteriori* element to it. The second way to answer the **Core Question** proceeds completely *a priori* with special attention to historical puzzles associated with a concept.

In Chapter 2, when trying to make sense of the concept of continuity, I use the first style of answer to the **Core Question**. Specifically, I ask, how do practicing mathematicians use the concepts of *continuity*, or *the continuum*? This involves significant historical analysis and some cursory sociology of mathematics. The “sociology” is primarily based on my own experience with working mathematicians, not on any explicit studies where data are gathered and carefully analyzed. Either way, I take this route when it comes to *continuity* because my sense is that mathematicians have
a significant amount to say about this concept in a way that outstrips the neighboring philosophical literature. In this sense, the concept is largely a mathematical one. Below, the reader will see that within mathematics, rather than there being a univocal concept of continuity, there is a plurality of distinct concepts of continuity. Furthermore, each of them is perfectly consistent and therefore on equal footing, mathematically speaking. The applicability and value of these different notions of continuity is a matter of dispute. This dissertation is intended to broaden the appeal of those specific concepts of continuity that require the presence of infinitesimals.

In Chapters 3 and 4, my investigations are almost entirely *a priori*, following the second style of answering the **Core Question**. What does it look like to analyze a concept *a priori*? I will assume throughout that an acceptable way to dissect a concept is to examine individual (or in some cases communal) intuitions on how a concept is marshaled across a wide variety of real or imagined cases. Indeed, it is primarily through examining highly abstract possibilities that one gets a tighter grip on the relevant concepts.\(^3\) This technique has come under a large amount of scrutiny lately. Most likely, this is because of a complementarily large amount of usage of these techniques in contemporary philosophy.\(^4\) This is not the time to settle those issues. Between now and that time, some light is shed on this technique by John Hawthorne in an article dealing with nearby difficulties for the concept of causation:

Some may complain that [the world in question] is too distant to be worth being interested in. Actual walls do have extra repulsive forces, don’t get to be rigid and impenetrable at any thickness and so on. Such a reaction is far too hasty. Distant worlds can often be either revealing or

---

\(^3\)There is extended discussion of this technique in David Chalmers and Frank Jackson’s “Conceptual Analysis and Reductive Explanation,“ [37], specifically §3. See also Ned Block and Robert Stalnaker’s “Conceptual Analysis, Dualism, and the Explanatory Gap,” [16].

\(^4\)For literature in support of this kind of conceptual analysis, the reader should consult Dean Zimmerman’s “Prologue: Metaphysics after the Twentieth Century,” [227]. A direct treatment from the opposition is found in Mark Wilson’s “Beware of the Blob: Cautions for Would-Be Metaphysicians,” [222]; for a more exhaustive discussion of concepts and of related issues in the philosophy of language, see Wilson’s *Wandering Significance*, [221]. Illuminating remarks can also be found in the “Epilogue” to Tim Maudlin’s [132].
therapeutic with regard to our actual conceptual scheme. So let us press on. (Hawthorne, [86], 693)

The idea is that we may learn much about our concepts in the actual world by imagining and considering their application in other ones; to take a famous example, one learns about the concept of natural kinds by asking, Is there any water in the lakes of Twin Earth?\(^5\) Even if much of this study has been undertaken a priori, I do not take my judgements on these concepts to be incorrigible.

Chapter 1.3

Summary of Chapters

In Chapter 2, I examine the concept(s) of continuity. I focus specifically on four distinct requirements on continuity:

- **Dedekind Continuity**—An object is continuous just in case it is composed of a linear ordering of points and there is, for any two adjacent parts, exactly one point dividing them.

- **Indecomposability**—An object is continuous only if it is not identical to the sum of any of its disjoint proper parts.

- **Infinitesimal Magnitudes**—An object is continuous only if it is composed of infinitesimal magnitudes.

- **Non-Supervenience**—An object is continuous only if all of its parts have a proper part.

Although each of them has its origins in earlier stages of mathematics or philosophy, they are all exemplified within modern formalized mathematics. For each requirement, I provide philosophical motivations and problems for it, followed by a fully formalized presentation of it from the most current literature. The reader will note that a tension exists between these different requirements: some are more sensitive to philosophical ideals, others are more sensitive to mathematico-scientific needs. I

\(^5\)See Putnam’s, “The Meaning of ‘Meaning,’ ” [158].
do not actually express preference for any one of these over the others—all of them satisfy basic requirements on mathematical theories. The interesting question is, *Is such-and-such conception of continuity applicable in ways that help solve philosophical problems?* In some cases, I hope to answer this question in the affirmative.

I should take this moment to acknowledge that there are certainly concepts of continuity that I have neglected—some to my chagrin. John Conway’s Surreals and the Intuitionistic Continuum underneath Constructive Analysis are the two major concepts of the continuum that I was forced to neglect. The Surreals form an extraordinarily rich system of numbers that elegantly generalizes the idea behind Dedekind cuts—the latter is described below at length in §2.3.3. The Intuitionistic continuum of Constructive Analysis satisfies *Indecomposability* described above. In spite of its historical value and philosophical interest, I drew the line just short of including it. My selections were based on two things: (i) I had to include what I took to be essential notions of continuity from history (Aristotle’s and Dedekind’s) and (ii) any other notions of continuity that would be employed in the later metaphysical chapters.

In Chapter 3, I offer a novel solution to one of Zeno’s paradoxes of motion, The Arrow. Consider an arrow in flight. In the present, the arrow is moving in virtue of its being in flight. However, if the present is a single point in time, then the arrow is frozen in place during that time. Therefore, the arrow is both moving and at rest. One historical solution suggests conceiving of the present as an extended interval of time. In this way, one may deny that the arrow is at rest in the present. Unfortunately, since no length for the present is ever specified, this solution has a strong degree of arbitrariness. In “Zeno’s Arrow, Divisible Infinitesimals, and Chrysippus,” Michael J. White suggests using an infinitesimal value as the length of the present allowing him to temper some of this arbitrariness. I follow White in the basic outline of his solution, although his view still suffers from a measure of arbitrariness. I argue that the
arbitrariness can be further diminished if one uses a different theory of infinitesimals. In particular, I employ Paolo Giordano’s Fermat Reals and F.W. Lawvere’s Smooth Infinitesimal Analysis. Each of these contains a special kind of infinitesimal called a nilpotent. Neither solution is perfect: use of the former system eliminates only so much arbitrariness; the latter, while completely eliminating arbitrariness, involves substantial theoretical costs including requiring the use of intuitionistic logic.

In Chapter 4, I examine a much-discussed puzzle concerning how pairs of material bodies can be in contact. A receptacle for a material body $M$ is a spatial container in which $M$ fits exactly. The puzzle of contact is that if both bodies’ receptacles contain or lack all of their boundary points, our pretheoretic intuitions about contact are violated. If both receptacles lack their boundary points, then there is a spatial gap between the two bodies where the boundary would be; if both receptacles contain their boundary points, there is a spatial glut at the boundary where the two bodies’ matter must share a point in space. In the literature, some authors argue that neither of these cases counts as a case of contact.\footnote{Most famously, Bernard Bolzano held this view. I examine a version of Bolzano’s view developed by Hud Hudson in §4.5.4.} Other authors conceive of the structure of space in such a way as to eliminate these gaps and gluts.\footnote{Dean Zimmerman and William Kilborn both adopt this strategy.} Others still try to show why these gaps and gluts are not so bad after all.\footnote{This is most notably defended by Sheldon Smith in his excellent “Continuous Bodies, Impenetrability, and Contact Interactions: The View from the Applied Mathematics of Continuum Mechanics,” [194].} My own solution is a synthesis between these latter two. I count any bodies in contact when they have no more than an infinitesimal length between their receptacles (where being an infinitesimal is relative to the size of the receptacles). As such, the structure of space is conceived to contain infinitesimals, gaps are tolerated and gluts are not. The upshot is that neither the presence nor the absence of boundary points presents an obstacle to contact; furthermore, this solution allows me to avoid any difficulties associated with
Lastly, I will conclude things with some final methodological reflections in the Epilogue.
Chapter 2

The Labyrinth of Continuity

There are two labyrinths of the human mind: one concerns the composition of the continuum, and the other the nature of freedom, and both spring from the same source—the infinite.

—Leibniz, “On Freedom”

Of all conceptions Continuity is by far the most difficult for Philosophy to handle ... As to the proper method of reasoning about continuity, the dictate of good sense would seem to be that philosophy should in this matter follow the lead of geometry, the business of which it is to study continua.

—C.S. Peirce, “The Logic of Continuity”

The concepts of continuity and the continuum have many incarnations in contemporary mathematics. As a first gloss, something is continuous if it is represented naturally by a line, plane or solid. This gloss suggests that continuity is a geometric concept at its core. Although interesting questions regarding continuity arise within ordinary Euclidean geometry, the most perplexing difficulties emerge in an attempt to understand continuity within the formalistic and abstract setting of contemporary mathematics. Even more specifically, numerous attempts have been made to explicate the concept of continuity in a fashion consonant with the types of questions in which contemporary mathematicians are interested and with the types of tools they use. These attempts at explication have produced a plurality of conceptions of continuity that are distinct in mathematically tractable ways.

---

1See [117], pp. 406-407.
2See Peirce, [149], 2D.
3The term ‘conception’ is not meant to be distinct from ‘concept’ in any philosophically substantive way. The only reason for using this term is that it more strongly connotes that there is some history behind the concept.
In light of these things, the question I want to attempt to answer is, *What do these different conceptions of continuity have in common, if any? Is there some common concept of continuity to which all of these answer?* In what follows, I argue that at a certain level of generality, all of these conceptions satisfy the following principle:

1. *Infinite Divisibility*—An object $X$ is *infinitely divisible* just in case it is extended and any extended part of $X$ has at least two disjoint (non-overlapping) extended parts.

As I will demonstrate more below, there are mathematical objects that satisfy *Infinite Divisibility* but are not continuous. Some further principle must be added. From this point forward, there is significant divergence in what counts as continuous.

To give a preview of coming attractions, I will briefly outline a quartet of distinct requirements on continuity beyond *Infinite Divisibility*.

2a. *Dedekind Continuity*—An object is continuous just in case it is composed of a linear ordering of points and there is, for any two adjacent parts, exactly one point dividing them.\(^5\)

2b. *Indecomposability*—An object is continuous only if it is not identical to sum of any of its disjoint proper parts.

2c. *Infinitesimal Magnitudes*—An object is continuous only if it is composed of infinitesimal magnitudes.

---

\(^4\)‘Extension’ historically means something nearly synonymous with ‘continuity.’ I take *extended* to apply also to a sum of points scattered through space. Also departing from certain historical uses, an extended object need not be a material body extended in physical space. Extension can apply also to purely mathematical entities. For example, a line or plane in Euclidean space would qualify as extended. Although many of my examples below will involve three-dimensional material bodies, other purportedly continuous phenomena are properly modeled using fewer than three dimensions; for example, the pathway of an electromagnetic ray through space.

\(^5\)First, this definition only applies to 1-dimensional objects. An $(n+1)$-dimensional version of continuity can be developed by requiring of any $(n+1)$-dimensional object $X$ that each $n$-dimensional cross-section of $X$ is continuous. Second, ‘adjacent’ is meant to be used in an ordinary way here; two adjacent parts are side by side. To be more precise, two parts are *adjacent* just in case they are disjoint and there is no more than one point between them, where betweenness can be captured either order-theoretically or topologically as the context dictates. Similarly, adjacent parts $X$ and $Y$ are *divided* by a point $p$ when that point is between adjacent parts $X \setminus \{p\}$ and $Y \setminus \{p\}$. For concreteness, the intervals on the real line, $(0,1)$ and $(1,2)$ are adjacent and divided by the point at 1, whereas $(0,1)$ and $(3,4)$ are not adjacent. Similarly, single points cannot be adjacent to one another. In effect, in a domain of only points, the concept of adjacency only applies to a linear array of points that contains accumulation, or limit, points. The latter is guaranteed by *Infinite Divisibility*. 
2d. **Non-Supervenience**—An object is continuous only if all of its parts have a proper part.\(^6\)

Although some of the above requirements are logically compatible with others, each of these requirements has been selected so as to highlight its uniquely associated historical conception of the continuum. A natural hope would be that all of these conceptions would converge on a single mathematical representation. As I’ve suggested already, this is exactly what does not happen. The conceptions of continuity associated with the above requirements each gives rise to a provably distinct formal system.

Here is a roadmap for the chapter. First, I will briefly examine the basic intuitions and ideas behind *Infinite Divisibility*, highlighting why it is necessary but insufficient for understanding the continuum. Second, I will discuss some methodological issues and at a fairly abstract level how these different views can be categorized. After this, I will unravel each of those further requirements on continuity listed above one by one. Throughout, a strong tension will emerge between fidelity to core philosophical presuppositions about the continuum and the practical needs of working mathematicians. At the limit of practicality, the conceptions under consideration must at least be able to prove basic Euclidean geometric theorems; whereas at the limit of philosophical fidelity, I will examine theories that surrender even the law of excluded middle to uphold intuitions about continuity. However, I do not consider any theories of the continuum that surrender the law of non-contradiction for conceptual purposes, in spite of there being examples.\(^7\) What will become clear below is that no conception satisfies both our philosophical and mathematico-scientific appetites.

---

\(^6\)A few words on this definition. First, the terminology comes from White’s *The Continuous and the Discrete*. I am simply borrowing his term, regardless of the otherwise confusing connotations for ‘supervenience’. Secondly, in the contemporary literature, this requirement necessitates that continuous things are gunky.

\(^7\)Examples of inconsistent continua are gestured at by Priest in his [156] and discussed in greater detail by Mortensen in [138]. Part of the reason for limiting myself here is that I find dialethias much more philosophically problematic than indeterminacies.
Infinite Divisibility

To begin, I want to understand why continuous phenomena are infinitely divisible. In order to see this, it would be helpful to compare continuity with discreteness. Aristotle writes:

Of quantities some are discrete, others continuous; and some are composed of parts which have position in relation to each other, others are not composed of parts which have position.

Discrete are number and language; continuous are lines, surfaces, bodies and also, besides these, time and place. (Categories VI)

Besides providing illuminating examples of each, Aristotle seems to suggest that discrete things are marked by their separateness. For Aristotle, some discrete things might together actually compose a whole, but even then the whole can be separated into natural, distinguishable parts.\(^8\) He offers the syllables of spoken language as such an example. For each word, it can be broken into its component syllables naturally. Although within a given word the syllables are in some sense together, each one is ultimately distinguishable.\(^9\) By comparison, continuous things form a kind of homogeneous unity. The parts of continuous things \textit{qua} continuous are in some sense artificial: a continuous thing could be divided \textit{any}where. The artificiality of a continuous thing's parts is responsible for its infinite divisibility.

To be more concrete, take a room filled with five chairs with only one painted red. The chairs in the room are discrete. The chairs as such can be separated out naturally into the five. Their separateness is bound up with the nature of \textit{chair} and is represented mathematically by the number five. In contrast, \textit{the surface of the seat}

\(^8\)For each example of discrete that he gives, he concludes, “[they] are separate” and again, “they are always separate.” (4b27, 4b31)

\(^9\)This example is actually disputable as a genuine case of discreteness. Regardless, this is Aristotle’s example and it gives a rough and ready idea of a case of unseparated but separable parts.
*on the red chair* is not discrete. It would be best represented mathematically as a square in a plane. Like a square, the surface of the red chair forms a unified whole. In the same way that the square can be bisected indefinitely, the surface of the chair is casually speaking infinitely divisible.\(^{10}\) This means that along a particular axis, one could cut the seat of the chair at an infinity of different places normal to that axis. Alternatively, to say the surface is infinitely divisible is as follows. If one cuts the surface of the seat in half, there is a left and right—call them \(\text{Left}_1\) and \(\text{Right}_1\). One may then cut \(\text{Left}_1\) in half, creating \(\text{Left}_2\) and \(\text{Right}_2\). Again, cut \(\text{Left}_2\) in half. One can imagine this process going on indefinitely. That is, for any positive integer \(n\), \(\text{Left}_n\) can be divided into \(\text{Left}_{n+1}\) and \(\text{Right}_{n+1}\). This is not true of the chairs in the room. The chairs altogether divide up into just the five chairs. One would have to start counting parts of chairs to go any further.\(^ {11}\)

The above description of infinite divisibility is equally compatible with either a potential or actual infinity. Although this issue is not essential to my aim in the chapter, a small amount of attention should be devoted to it, given its historical interest. Explicitly, a concept \(X\) is *potentially infinite* just in case there is a process for generating \(X\)’s; in other words, the \(X\)’s are not empty and if there are at least \(n\)-many \(X\)’s, then it is possible that there are at least \((n + 1)\)-many \(X\)’s. The above definition of potentially infinite allows for an actual infinity to also count as a potential infinity. A merely potential infinity is an essentially modal notion.\(^ {12}\) In that case, the \(X\)’s are finite but *could be* indefinitely more numerous. Concretely, the above

\(^ {10}\)For the moment, we are ignoring that the physical world is actually composed of certain basic units with some length or other. For this reason, I say “casually speaking.” We are thinking about this surface in a phenomenological manner; i.e. that there are no natural parts discernible to the naked eye into which the surface can be resolved.

\(^ {11}\)Even if one allows bunching the chairs into pairs, triples and quadruples, the number of distinct ways of dividing them up is \(\sum_{k=0}^{2} \binom{5}{k} = 16\).

\(^ {12}\)At this stage, I have no interest in embroiling myself in difficult questions about whether modal notions relate to possible worlds. I take these notions to be clear enough for one to use them without an explicit mention of possible worlds. I am neither hostile nor friendly towards possible worlds and prefer to stay that way for the time being.
example of dividing up the red chair’s seat involves at least a potentially infinity. Now, a concept \( X \) is \textit{actually infinite} just in case the \( X \)’s are not empty and if there are at least \( n \)-many \( X \)’s, then there are at least \( (n + 1) \)-many \( X \)’s. Note that in the actual case, the modal operator is dropped. The most straightforward example of an actual infinite is the classical set of natural numbers, \( \mathbb{N} = \{0, 1, 2, \ldots \} \).

For the sake of the present investigation, it is not relevant which notion is more legitimate or philosophically stable. I am interested in trying to make sense of the conceptions of the continuum developed by practicing mathematicians over the last century or so. This is not the place to settle the legitimacy of the practice. For the duration of this investigation, I will simply assume that the techniques and practices of mathematicians are not in need of any revision. Similarly, the products of these mathematicians’ labors—including the various incarnations of the continuum—are worthy of careful study in their own right. That being said, \textit{all} of the different conceptions investigated below in any detail assume that the continuum is actually infinitely divided.

As I have already suggested, infinite divisibility is at the conceptual core of what it is to be continuous, and there is little disagreement about how to understand infinite divisibility (modulo the potential-actual distinction). Unfortunately, as I have suggested already, understanding continuity is not so simple: infinite divisibility is a necessary but not sufficient criterion. One might think, “Infinite division is \textit{a lot} of divisions, what would be left after infinitely many divisions? Doesn’t \textit{Infinite Divisibility} tell us enough about the continuum?”

\footnote{The intuitionists view the natural numbers as potentially infinite, so one must be cautious to distinguish the guise under which she is discussing the natural numbers.}

\footnote{Since my primary interest is in understanding how these different conceptions of the continuum are expressed in modern mathematical terms, those mathematical representations considered below sensitive to Aristotle’s conception of the continuum are all actually infinite in size. Furthermore, Aristotle’s arguments against the possibility of an actual infinity can be distinguished from his views concerning the relationship of points and the continuum. Therefore, I will ignore those aspects of Aristotle’s work requiring merely potential infinity.}
I now turn to an example showing how mere infinite divisibility has mathematico-scientific consequences as well, independently of various purely conceptual concerns: Euclid’s construction of an equilateral triangle found in Book I, Proposition 1. The problem is that a very basic and straightforward geometric construction cannot be carried out in the rational plane\(^{15}\) even though such a plane is infinitely divisible. Let us take a closer look. Essential to this construction is the assumption that when two circles overlap one another, there are two points of intersection—one on top and one on bottom, as is Figure 2.1.\(^{16}\)

![Figure 2.1. Equilateral Construction from Euclid’s Elements](image)

Suppose that the two circles are of radius 2, and centered at (1,0) and (-1,0)—points B and C in Figure 2.1. To complete the construction of the equilateral triangle, draw lines from C and B to one of the intersection points of the circles. Following Euclid, I chose the top one, labelled A. The problem is that the points of intersection would be

---

\(^{15}\)The rational plane is a geometric plane composed only of points from the rational numbers, where the rational numbers are those values of the form \(\frac{m}{n}\) for integers \(m\) and \(n\). In familiar set-theoretic terminology, the rational plane is \(\mathbb{Q} \times \mathbb{Q}\).

\(^{16}\)The original way this example gets used is to show that Euclid’s own axioms are not sufficient for establishing the existence of these two points. Euclid’s system requires some kind of continuity assumption (whether by assuming continuity implicitly in the diagrams or by using a more axiomatic technique) to get this and other constructions to work. For discussion on this and many other fascinating issues, see for example Lisa Shabel’s [180] or Michael Friedman’s [73], especially Ch. 2.
at \((\pm\sqrt{3}, 0)\)—irrational values. It follows that in the rational plane, these two circles do not intersect. They are more like interlocking gears with punctiform teeth. All this is to show that something more is needed past \textit{Infinite Divisibility}. At this point, the different conceptions of continuity radically diverge.

\textbf{Chapter 2.2}

\textbf{Conceptual Terrain}

In this section, I will try to map out some terrain essential to the debate about continuity. In the space of concepts of continuity, there are four categories into which the various views fall, cross-cut by two axes. Although the two issues represented by these axes are related in certain ways, there are not clean entailments between them. The first axis is divided over what mereological relation receives metaphysical priority in conceiving the continuum: division or composition. Or, equivalently, are the parts of the continuum fundamental or is the whole continuum itself fundamental? The second axis is divided over the role of points in any given conception of the continuum: are they \textit{parts} of the continuum? Are they mere \textit{locations}?

It should be noted that the distinctions here are drawn in philosophical interpretations of the mathematics or in some cases mathematicians’ philosophical musings on their own work. The aim here is to establish a taxonomy or vocabulary for understanding these conceptions of the continuum and how they interrelate. Although I personally find these distinctions natural from examining the mathematics, I am not claiming that these are necessarily \textit{inherent} in the mathematics. Moreover, contemporary abstract or algebraic mathematics is especially notorious for enabling unusual interpretations and witnessing sometimes shocking structural identities in otherwise disparate domains. For these sorts of reasons, I am not offering these distinctions as the way things must be seen but rather as a suggested route for the tour of the
conceptual terrain.

2.2.1 Top-Down or Bottom-Up

One way of thinking about the continuum is to begin with something continuous and try to analyze its parts qua parts. On this view, the continuum is itself fundamental; or, at least, it is more fundamental than its parts, let alone points. Call this view top-down. On the other hand, one might view the parts of the continuum as fundamental. The continuum is then built up from its parts. Call this view bottom-up. This distinction is especially important in trying to understand the structure of the continuum in contemporary mathematics. Contemporary mathematical practice prioritizes a bottom-up approach in nearly every area of mathematical activity; whereas, traditional approaches to continuity have emphasized a top-down approach.\(^\text{17}\) Now for a closer look.

Top-Down Approaches

As I suggested above, top-down approaches have their origins in antiquity. Although Aristotle’s conception of the continuum has significant appeal in its own right, top-down approaches emerge further upstream conceptually. In particular, the top-down picture can been witnessed in Euclidean geometry. To some extent, Euclidean geometry is silent on this question: there is no mathematical discussion of the parts of the continuum. On the other hand, lines, planes and solids are all treated as existing unto themselves to be divided by various constructions (for example, by a linear bisection).

\(^\text{17}\)I will discuss this more below, but the category-theoretic approach, unlike the set-theoretic approach, is much more “top-down.” One may discount this claim about mathematical practice to the extent that certain branches of mathematics are moving in that direction. My own experience is that, although the top-down, category-theoretic picture is gaining ground, many practicing mathematicians still operate with an overall set-theoretic method or framework. For illuminating discussion, see Steve Awodey’s [3].
into parts.¹⁸ In that way, the top-down view is presupposed throughout the *Elements*.

One thing is for sure: the notion that points compose the line is neither explict nor implicit in the *Elements*. In Euclidean geometry, points and lines are treated as playing very distinct roles. No mereological relationship between points and lines is suggested in Euclid’s *Elements*. Points are boundaries of line segments, or divisions along the line. If points play any mereological role at all, it is in drawing one’s attention to the line’s parts. Points do not play a mereological role by themselves being parts of the line. The same goes for higher dimensions: lines only mark off the edges of polygons, they do not compose polygons; polygons mark off the edges of solids, they do not compose solids.

Along with the fact that the top-down approach is a natural outgrowth of Euclidean geometry, the top-down approach undergirds a long-standing intuition that the continuum is unified through and through. At this stage, the formulation of the intuition is deliberately inchoate. Only slightly more helpfully, recall the discussion above regarding the distinction between discreteness and continuity. Discrete quantities exhibit distinctness and plurality; continuous quantities exhibit coherence and unity. The top-down approach provides a natural place for this unity intuition. The fundamentality of the continuum over its parts, let alone over points, delivers an image of the continuum as a single unity primarily and as composed of parts secondarily. Top-down approaches tend to view the parts of the continuum as being especially “sticky”: the parts are not merely arranged in a neat order, but are in some sense “glued” together. The continuum is thus sometimes described as being viscous.¹⁹ The way these intuitions get captured formally will be drawn out at length below. For the time being, this should help set the stage for more detailed discussions.

---

¹⁸Euclid does actually use the language of part-whole, including using even ‘µέρη’—the word from which we get mereology—in the original. See for example, Book V, Proposition 15.

¹⁹To my knowledge, use of the expression ‘viscous’ comes in this way primarily from Carl Posy’s work on Brouwer’s continuum, e.g. [152].
of particular conceptions of, and requirements on, the continuum.

**Bottom-Up Approaches**

In contrast with top-down approaches, bottom-up approaches—where the priority is on the relation of composition—are much newer historically. Newer still is the idea of treating the continuum as composed of points within mathematics. Starting no later than the medieval period and continuing into the eighteenth century, mathematicians would treat the continuum as composed of infinitesimally extended line segments; i.e. magnitudes \( m \) so small that every integral multiple of \( m \) is still strictly less than any chosen unit \( u \). More formally, \( m \) is infinitesimal iff for any unit \( u > 0 \), integer \( n > 0 \), \( m \cdot n < u \).\(^{20}\) Even then, many of these mathematicians would treat these as convenient fictions. In their philosophical moments, they would take back what they gave in their mathematical ones. For example, Leibniz is known for viewing infinitesimals as fictions; to use infinitely small and large values is fruitful but they do not exist strictly speaking:

> Even though these [geometric figures composed of infinitesimals] are fictitious, geometry nevertheless exhibits real truths which can also be expressed in other ways without them. But these fictitious entities are excellent abbreviations for expressions and for this reason extremely useful. (Leibniz, [119], 89, 91)

In spite of Leibniz’s otherwise liberal usage in his mathematical work, his preferred metaphysical view of the continuum was very much top-down.\(^ {21}\) Therefore, even though technical developments in mathematics suggested bottom-up thinking, it was not until the late nineteenth century that a truly bottom-up picture emerged.\(^ {21}\)

---

\(^{20}\)One must be careful to note that in certain settings, the notion of an infinitesimal is a relative one. This is true of Abraham Robinson’s hyperreals. This relativity is not true universally. Below, I will examine Giordano’s Fermat reals. This number system contains an absolute notion of infinitesimal.

\(^{21}\)In a paper called “Infinite Numbers,” he writes, “In the continuum, the whole is prior to its parts; the absolute is prior to the limited; and so is the unbounded prior to that which has a bound.” (Leibniz, [119], 97) Leibniz is slightly equivocal, which makes him a rather mysterious figure. His somewhat elusive fictionalism allows him to maintain these two voices.
The first bottom-up conception of the continuum accepted by the vast majority of mathematicians comes from Richard Dedekind and is discussed at length below. More germane to the present issue is that the enormous mathematical success of his conception reversed a multimillenial trend of viewing the continuum from the top down. From that point forward, the most extreme bottom-up approach—that of treating the continuum as composed of points—became enshrined as the new orthodoxy. Even in spite of contravening such a long history of viewing the continuum in other ways, the bottom-up approach is now so entrenched that other conceptions of the continuum are viewed as intellectual curiosities at best, wastes of time at worst. The tide is turning back some, but not nearly with the same kind of fervor that inaugurated the reign of the original bottom-up approach.

Given the weight of tradition pushing heavily against this shift, what are some of the merits of the bottom-up approach? As I indicated, these merits are primarily bound up with technical developments internal to mathematics. Dedekind’s bottom-up approach fits very naturally into a contradiction-free rigorous analysis. In other words, Dedekind’s own bottom-up approach constrasts with the early concept of the infinitesimal found at the birth of the differential calculus—a discipline fraught with contradiction.\footnote{This rigorous analysis eschewed the infinitesimals of the early differential calculus. As I will argue below, infinitesimals can be resurrected within a fully rigorous framework.} Those portions of present day analysis relevant to the present discussion are in few ways—if any—different from how analysis was rigorized in the mid to late nineteenth century.\footnote{See Walter Rudin’s [172], Appendix to Chapter 1. There, Rudin constructs the real numbers using Dedekind’s techniques.} Furthermore, the overall set-theoretic worldview that dominates contemporary mathematics is also essentially bottom-up: everything in the mathematical universe is built up from the most primitive parts. One begins with only the empty set and forms new sets by iterating a variety of set-forming principles. From a purely historical point of view, the fact that Dedekind’s work
developed in tandem with set theory and its companion, point-set topology, only
further served to entrench the bottom-up approach. In other words, Dedekind’s
bottom-up conception of the continuum was consonant with an overwhelming package
of very fruitful mathematical tools.

In summary, such a radical reversal in the history of mathematical thought about
the continuum is nicely explained by the bottom-up approach’s harmony with set the-
ory and the standards of rigor in contemporary analysis. Compared with top-down
approaches, bottom-up approaches prioritize mathematico-scientific values over tra-
ditional philosophical ones. I suspect that such a tension between mathematical
fruitfulness and philosophical fidelity is what pushed philosophically sensitive math-
ematicians, like Leibniz, towards viewing infinitesimal magnitudes as merely fruitful
but unacceptable for the “official business” of philosophy.

2.2.2 Points on points

Every view of the continuum involves points. The more interesting issue is, what is
the exact relationship of points to the continuum? The most prominent relationship
is mereological: are points parts of the continuum? Many of the conceptions that
I will examine treat the points as parts. Call these conceptions of the continuum
punctate. For those that are not punctate, points must play some other role, to be
described in a moment.

Nearly every conception of the continuum below is punctate. Allow me to be more
precise about what counts as punctate. First, I will use set-theoretic inclusion as the
mark of when something counts as a part. Also, I will count singletons of members of
the continuum as points. More explicitly, a conception of the continuum \( \mathcal{C} \) is punctate

\[24\text{John Bell uses the expression }\text{punctiform }\text{ in }\text{The Continuum and the Infinitesimal, } [8].\text{ I do not prefer this term because ‘punctiform continuum’ connotes that the entire continuum is shaped like a point. The expression ‘punctate’ simply suggests that there are points all over the continuum.}\]
just in case for any part $\mathcal{P}$ of $\mathfrak{C}$ there is at least one point $p$ such that,

$$p \subseteq \mathcal{P} \quad \text{and} \quad |p \cap \mathcal{P}| = 1,$$

(2.2.1)

where $|A|$ is the standard cardinality notation.

In review of this formulation, one might think that it is unhelpful to define punctation in terms of set-theoretic inclusion. I should note that I do not take punctation to be necessarily captured in this set-theoretic way. That being said, most of the formal expressions of the requirements below are given in set-theoretic terms. Therefore, the set-theoretic formulation of these notions will be helpful rather than present a significant barrier. One might also worry that (2.2.1) is too weak—why have I put ‘at least one point’ rather than something stronger? This formulation is just a translation of the mereological thesis of Atomism—that all the way down, things have simples or atoms as parts. In this case, at the bottom of the continuum, the points are basic proper parts. Aside from the question of Atomism, a weaker formulation is ideal: there are (especially peculiar) conceptions of the continuum that contain points as parts but are not composed of them. Therefore, it is best not to capture punctation as simply requiring that the continuum be a sum of points. All told, (2.2.1) will be the rubric for assessing the punctation of those conceptions treated in detail below.

Another way of viewing the points of the continuum is that they are places, or markers. This view is most closely associated with an Aristotelian view, according to which points are not parts at all. That view will be described in greater detail below. In the mean time, consider Aristotle’s remarks:

That which is divisible in two dimensions is a plane, that which is divisible in one a line, that which is no way divisible in quantity is a point or a unit,—that which has not position a unit, that which has position a point. (Metaphysics Δ, 1016b28-30, emph. added)

As far as I know, every view of the continuum treats points as locations on the
continuum. Punctuation is not precluded by viewing points as locations. The concepts of part and location are not mutually exclusive.\textsuperscript{25} For the punctate conceptions, the points do double duty. The point of introducing this concept is to explain the role that points are to play for those who reject their mereological role.

Chapter 2.3

Dedekind’s Orthodox Conception

Within mathematics, the best known conception of continuity comes from the late nineteenth century. The combined work of Augustin-Louis Cauchy, Bernard Bolzano, Karl Weierstrass, Richard Dedekind and Georg Cantor is responsible for the way that calculus and basic analysis is learned by all mathematics students even into the twenty-first century. Given its wide use and application, this view may be properly called orthodox. The conception of continuity to be discussed here is the one satisfying the aforementioned:

2a. \textit{Dedekind Continuity}—An object is continuous just in case it is composed of a linear ordering of points and there is, for any two adjacent parts, exactly one point dividing them.

The underlying thought here is to ensure that there are no gaps between a suitably ordered assembly of points. The rational numbers, though infinitely divisible, do not succeed in composing a continuum. A number of gaps are left between the rational numbers. \textit{Dedekind Continuity} ensures that these gaps are filled. Dedekind’s orthodox conception is the archetype of a bottom-up, punctate conception of the continuum. In what follows, I will try to further unravel some of the motivations as well as examine a few reasons that philosophers and mathematicians have balked at Dedekind’s portrayal of the continuum since its introduction.

\textsuperscript{25}At present, the relation between parts and locations is receiving a lot of attention. For example, see Josh Parsons’ “Theories of Location,” [147], and Gabriel Uzquiano’s “Mereological Harmony,” [202].
2.3.1 Motivations

This conception plays a prominent role in part because it is among earliest portrayals of continuity with no recourse to primitively geometric objects or spatial intuition.\textsuperscript{26} The mathematicians responsible for this view are not motivated as much by supplying an especially philosophically sensitive portrayal of continuity. Rather, they want to provide a foundation for their activities that are consonant with their own standards of rigor. In previous times, explicit appeal to spatial intuition was acceptable, if not required. As the standards rose to requiring mathematicians to render one’s mathematical inferences perfectly explicit, intuition no longer cut it, except perhaps in pedagogical settings. On his motivation, Dedekind remarks:

\begin{quote}
Even now such resort to geometric intuition in a first presentation of the differential calculus, I regard as exceedingly useful, from the didactic standpoint, and indeed indispensable, if one does not wish to lose too much time. But that this form of introduction into the differential calculus can make no claim to being scientific, no one will deny ... For myself this feeling of dissatisfaction was so overpowering that I made the fixed resolve to keep meditating on the question till I should find a purely arithmetic and perfectly rigorous foundation for the principles of infinitesimal analysis. The statement is so frequently made that the differential calculus deals with continuous magnitude, and \textit{yet an explanation of this continuity is nowhere given}; even the most rigorous expositions of the differential calculus do not base their proofs upon continuity but, with more or less consciousness of the fact, they either appeal to geometric notions or those suggested by geometry, or depend upon theorems which are never established in a purely arithmetic manner. (Dedekind, [48], 1,2, emphasis added)
\end{quote}

As suggested in this passage, Dedekind’s own reason for developing this view of continuity was to offer a purely formal presentation of continuity. Up to that point, continuity was an informal notion upon which mathematicians leaned when the formal mathematics itself gave out. The standards of rigor up to that point did not

\textsuperscript{26}The \textit{intuition} discussed here is very much like what Kant means when he refers to space as a form of pure intuition.
require anything more explicit than what was already on hand. Casual references to spatial intuition or diagrams sufficed. As the standards of rigor changed through the nineteenth century, reliance on intuition became less and less acceptable.

So much for historical introduction—why should this help in explicating continuity? The best way to understand this requirement is to return to the contrast between discreteness and continuity from above. As a reminder, discrete things are marked by their separateness. Dedekind’s aim here is to countenance only the minimal resources to eliminate separateness.

I will take separation as the guiding feature of discreteness, and interpret a separation as a gap. So, on the Dedekind conception, separateness or gappiness marks discreteness, and togetherness or lack of gaps marks continuity.

To sharpen things a bit, recall the example above from Euclid’s *Elements*. When the two circles come together, there is supposed to be a point of intersection. The worry is, how can we be sure there is a point of intersection? Simply appealing to a sense that they should intersect does not suffice for rigorous demonstration. The rational numbers are a well-defined class of values that fall along the line—what about them? Unfortunately, the measure of sides and heights of mere triangles proliferates irrational values and thus proliferates gaps in the rational line; similarly for the rational plane. This is a weakness that cannot be ignored given that triangles are the simplest figures in plane geometry. Dedekind’s insight was to stop the gaps.

Dedekind remarks on a similarly revealing example:

> The above comparison of the domain \([\mathbb{Q}]\) of rational numbers with a straight line has led to the recognition of the existence of gaps, of a certain incompleteness or discontinuity of the former, while we ascribe to the straight line completeness, *absence of gaps*, or continuity. (Dedekind, [48], 10, emph. added)

\[27\] Notably, based on strict interpretations of the definitions above, Aristotle would categorize Dedekind’s conception of continuity as discrete. I will discuss Aristotle’s views more below and how his conception of continuity is unfriendly towards Dedekind’s.
The idea of filling in the gaps left by the rationals was completely revolutionary. Although Dedekind probably did not expect it at the time, his own conception is now the gold standard. As such, it has been hard for anyone to call it into question.

### 2.3.2 Reasons for Concern

Dedekind only required *Dedekind Continuity* along with *Infinite Divisibility*. Other conceptions are significantly more stringent, as anticipated by Dedekind himself:

> [T]he majority of my readers will be very much disappointed in learning that by this commonplace remark the secret of continuity is to be revealed. (Dedekind, [48], 11)

In this sense, the *Dedekind Continuity* criterion places very low requirements on what may qualify as continuous. As I go further into the requirements on continuity below, it will become clearer why people have a problem with this conception. That is, the concerns raised here will be further illuminated by seeing how others have viewed the continuum. Nevertheless, I will explore in a preliminary way here why people balk at the Dedekind conception of continuity. In effect, it is at odds with the more traditional top-down approach; i.e. its parts do not *cohere* enough—or at all.

Most generally, the task of developing a logically rigorous (or in traditional terminology, arithmetic) presentation of the continuum is a difficult one. The reason for this is that to be continuous is to be “sticky” or “viscous” to an extent that seems to be undermined by the precise and analytical nature of contemporary formal mathematics. For many, the unity of the continuum is its premier feature. In particular, to form the continuous (a line) from the non-continuous (points on the line) is to subvert or entirely ignore this feature. This is to view the continuum as a kind of hyper-plurality rather than a unity. Henri Poincaré affirms as much, only with approval:

> Of the celebrated formula, ‘the continuum is unity in multiplicity,’ only
the multiplicity remains, the unity has disappeared. The analysts are none the less right in defining their continuum as they do ... ([150], 43)

To those indoctrinated with the Dedekind picture, this is exactly how the continuum is viewed today: the continuum is best known for being uncountably large.

2.3.3 Fully Formalized

In this section, I will construct the real line using the familiar method of Dedekind cuts. Briefly, this method involves “dividing,” or cutting the rational line, \( \mathbb{Q} \), and then forming equivalence classes of these cuts of \( \mathbb{Q} \). These cuts satisfy Dedekind Continuity, articulated here in Dedekind’s own words:

If all points of the straight line fall into two classes such that every point of the first class lies to the left of every point of the second class, then there exists one and only one point which produces this division of all points into two classes, this severing of the straight line into two portions. (Dedekind, [48], 11)

In what follows, I will explicitly construct the real line, \( \mathbb{R} \) from these cuts of \( \mathbb{Q} \) and then demonstrate how they satisfy Dedekind Continuity.

Basic Presentation

To begin, I assume that \( \mathbb{Q} \) is a dense ordered field. Generally, a field is a set \( F \) with binary operations \((+,\cdot)\) satisfying the following axioms:

1. **Closure:** For any \( x, y \in F \), both \( x + y \in F \) and \( x \cdot y \in F \).
2. **Associativity:** For any \( x, y, z \in F \), \((x+y)+z = x+(y+z)\) and \((x\cdot y)\cdot z = x\cdot(y\cdot z)\).
3. **Commutativity:** For any \( x, y \in F \), \( x + y = y + x \) and \( x \cdot y = y \cdot x \).

\(^{28}\)The reader might be tempted to view this “dividing” or “cutting” as somehow suggesting that this view is top-down, prioritizing division, rather than composition. Those terms are in some sense misleading, given that the cutting is of the rational line of points. This process starts with points of rationals, themselves constructed from the integers, and constructs sets of sets of them. The notion of cutting is simply that the rational line is split in two and out of this splitting emerges the points of the real line.
4. **Distributivity:** For any \( x, y, z \in F \), \( x \cdot (y + z) = (x \cdot y) + (x \cdot z) \).

5. **Identities:** For any \( x \), there are unique \( e \) and \( o \) such that \( x + o = x \) and \( x \cdot e = x \) and \( e \neq o \).29 (Here \( e \) is like 1 and \( o \) is like 0.)

6. **Additive Inverses:** For any \( x \in F \), there is some unique \( y \in F \) such that \( x + y = o \). (Often this \( y \) is written \( -x \).)

7. **Multiplicative Inverses:** For any \( x \neq o \), there is some unique \( y \) such that \( x \cdot y = e \). (Often this \( y \) is written \( \frac{1}{x} \).)

To form a dense ordered field, one must adopt the order relation \(<\) and then add the following to the above criteria:

8. **Transitivity:** For all \( x, y, z \in F \), if \( x < y \) and \( y < z \), then \( x < z \).

9. **Totality:** For any \( x, y \in F \), one and only one holds among \( x < y \), \( y < x \) and \( x = y \).

10. **Additivity:** For any \( x, y, z \in F \) \( y < z \) just in case \( x + y < x + z \).30

11. **Positivity:** For any \( x, y \in F \), both \( x > 0 \) and \( y > 0 \) only if \( x \cdot y > 0 \).31

12. **Density:** For any \( x, y \in F \) where \( x < y \), there is some \( z \in F \) such that \( x < z < y \).

Generically, let a cut, \( x \), be a pair of subsets of \( \mathbb{Q} \), \( L_x \) and \( R_x \)—intuitively for left and right. The cut will explicitly be the ordered pair, \( x = (L_x, R_x) \). \( L_x \) and \( R_x \) satisfy the following properties:

1. \( L_x \cap R_x = \emptyset \) and \( L_x \cup R_x = \mathbb{Q} \).

2. For all \( y \in L_x \) and \( z \in R_x \), \( y < z \).

A few remarks are in order. First, some of these cuts divide the rational line at a rational point. These cuts are copies of the rational values among the cuts. Other cuts—the most interesting ones—are those that divide \( \mathbb{Q} \) so that \( L_x \) has no least

---

29 Sometimes authors define fields without this final clause. To omit this clause allows for the degenerate one-membered field.

30 Combining this with existence of the multiplicative identity (and its inverse) guarantees that the field is unbounded (on both sides).

31 That \( x \cdot y > 0 \) when \( x, y < 0 \) is provable from other features of ordered fields. It does not need to be stated as part of the axiom.
upper bound and $R_x$ has no greatest lower bound. In other words, for any $z \in L_x$, there is some $y \in L_x$ so that $z < y$; similarly for any $z \in R_x$, there is some $y \in R_x$ so that $y < z$. These latter cuts are to make up the difference between $\mathbb{Q}$ and $\mathbb{R}$: they represent the irrational numbers among the cuts.

Before any interesting work can be done, one must define an equivalence relation $\equiv$ on these cuts. Without loss of generality, suppose $L_x \subseteq L_y$. Now, let

$$(L_y, R_y) \equiv (L_x, R_x) \iff_{df} L_y = L_x \text{ or } L_y \cap R_x = \{z\} \text{ for some } z \in \mathbb{Q}.$$ 

Clearly identical cuts should count as equivalent; the equivalence relation’s main point is to count as identical those examples where the left has the same least upper bound as the right has greatest lower bound. For example, if $L_x = (-\infty, 1)$ and $L_y = (-\infty, 1]$ then $(L_y, R_y) \equiv (L_x, R_x)$.

As it turns out, one may prove that the set of these equivalence classes under $\equiv$ form an ordered field satisfying Dedekind Continuity.\footnote{The details for showing that it is an ordered field are slightly tedious. In what follows, I will speak of each value as having a Left or Right, uniquely, ignoring the equivalence relation for now. To get the idea, the order relation ($<$) is defined to be proper inclusion on each non-equivalent Left. For two cuts $\chi_1$ and $\chi_2$, where $L_i$ is the Left of $\chi_i$, define $\chi_1 + \chi_2$ so that its Left is the set all the members $x_1 + x_2$ whenever $x_1 \in L_1, x_2 \in L_2$. Multiplication is more delicate. One must start with the positive numbers. Define the Left of $\chi_1 \cdot \chi_2$ to be the set of all $y \leq x_1 \cdot x_2$ for some $x_1 \in L_1, x_2 \in L_2$ where $\chi_1, \chi_2 > 0$. Multiplication of negative numbers is just defined by stipulation to fit the usual rules. Lastly, set the multiplicative identity to be the (equivalence class of the) cut where $L_e = (-\infty, 1)$ and the additive identity to be the (equivalence class of the) cut where $L_o = (-\infty, 0]$. From these definitions, it is relatively straightforward—even if a bit pedantic—to prove that this set of equivalence classes is an ordered field.} In fact, in many sources, $\mathbb{R}$ is just defined to be the set of equivalence classes of cuts of $\mathbb{Q}$. Now, I will not go through and actually show that this is an ordered field, but I will give a sketch of how this new system of numbers satisfies Dedekind Continuity.

\section*{Satisfying the Requirement}

Before diving in head-first, here is the requirement again:
2a. *Dedekind Continuity*—An object is continuous just in case it is composed of a linear ordering of points and there is, for any two adjacent parts, exactly one point dividing them.

The first clause regarding order is satisfied already by the fact that \( \mathbb{R} \), as defined above, is an ordered field. The second clause is where the action is. As a reminder, I explicitly defined adjacency above\(^{33}\) so that \( X \) is a adjacent to \( Y \) just in case that there is no more than one point between \( X \) and \( Y \).\(^{34}\) Suppose \( X \) and \( Y \) are adjacent parts along \( \mathbb{R} \) as defined above. Since there is a copy of \( \mathbb{Q} \) within \( \mathbb{R} \), \( X \cap \mathbb{Q} \) and \( Y \cap \mathbb{Q} \) are corresponding parts of the rational line. Let \( X_\mathbb{Q} \) and \( Y_\mathbb{Q} \) be \( X \cap \mathbb{Q} \) and \( Y \cap \mathbb{Q} \) respectively. First, I claim that

\[
X_\mathbb{Q} \text{ and } Y_\mathbb{Q} \text{ are adjacent parts of the rational line;} \tag{2.3.1}
\]

\( i.e. \) there is at most one point between them. For *reductio*, assume there are at least two points, \( a \) and \( b \), between \( X_\mathbb{Q} \) and \( Y_\mathbb{Q} \). It follows that there is a rational interval \( [a, b] \cap \mathbb{Q} \) between them. Therefore, \( [a, b] \) is between \( X \) and \( Y \). So, \( X \) and \( Y \) are not adjacent—contradiction.

Now, from claim (2.3.1), it follows that there is some cut \((L, R)\) such that \( X_\mathbb{Q} \subseteq L \) and \( Y_\mathbb{Q} \subseteq R \). Following convention, let \([L, R]_\equiv \) represent the equivalence class of \((L, R)\); call it \( p \). I claim that

\[
p \notin \mathbb{Q} \implies L < p < R, \tag{2.3.2}
\]

where \( L < p < R \) means that for all \( x \in L \) and \( y \in R \), \( x < p < y \). Suppose that \( p \notin \mathbb{Q} \). It follows that \( p \notin L \cup R \). Suppose for *reductio*, there is some \( x \in L \) such that \( p \leq x \). Hence, since \( L \subseteq L_x \) and since \( x \in L \), \( L_x \subseteq L \). Therefore, \( L = L_x \)

\(^{33}\)See footnote 2.

\(^{34}\)To be perfectly explicit, a point is between \( X \) and \( Y \) just in case for all \( x \in X, y \in Y, x < p < y \).
and so \( x \in \mathbb{Q} \)—contradiction. By parity of reasoning, for all \( y \in R \), \( p < y \). Hence, \( L < p < R \).

Since \( X \) and \( Y \) are adjacent then either (a) there is one point or (b) there are no points between them. Suppose (a), there is one point \( q \) between them. I claim that \( p = q \). If not, then without loss of generality, let \( p < q \). So, there is some \( x \in X_Q \) such that \( p < x \). It follows that there is some \( y \in L \) such that \( p < y \), contradicting (2.3.2). Therefore, \( p = q \), thus satisfying Dedekind Continuity. Alternatively, suppose (b), there are no points between them. In that case, \( p \in X \cup Y \). Without loss of generality, suppose \( p \in X \). It follows that \( X \setminus \{p\} \) and \( Y \) are adjacent. In that case, the point \( p \) divides \( X \) and \( Y \), satisfying Dedekind Continuity.

Chapter 2.4

Indecomposability and Excluded Middle

The Indecomposability requirement takes the top-down approach the most seriously, in spite of its also being punctate. Hermann Weyl favorably reports of Brouwer’s continuum, “A continuum cannot be put together out of parts,” (Weyl, [212], 135, emph. original). The parts are dependent on the whole so much that they cannot stand alone. As a reminder, this requirement is:

2b. Indecomposability—An object is continuous only if it is not identical to sum of any of its disjoint proper parts.

This view has its origins in the mathematical intuitionism(s) of L.E.J. Brouwer and Hermann Weyl, though my interest is primarily in Smooth Infinitesimal Analysis (sia), developed by Anders Kock and F.W. Lawvere.\(^{35}\) Here we consider those mo-

\(^{35}\)Many of the relevant mathematicians make reference to both Aristotle and Kant. From my own investigations, there does not seem to be ample evidence that Aristotle holds anything so strong. However, Kant is found saying the following regarding space, an archetype of continuity:

Space is not a discursive or, as is said, general concept of relations of things in general, but a pure intuition. For, first, one can only represent a single space, and if one
tivations developed by early Intuitionists because (i) these motivations also find ex-
pression in seria and (ii) these motivations are more fully developed philosophically
than what is exclusively available for seria. In fact, I will focus specifically on Weyl’s
philosophical writings given their comparative lucidity and given that he provides
ample philosophical commentary on Brouwer’s work as well as his own.

2.4.1 Motivation

This particular conception of the continuum is the most strict conceptually. Much of
the motivation for this strictness emerges from the Kantian features found in both
Weyl and Brouwer’s thinking. For the reader’s ease, I will write INTUITION and
CONCEPT to indicate the distinctively Kantian meaning of those terms and leave all
other uses to fit more contemporary usage. The only exception to this rule will occur
in quotations.

A starting point for both Weyl and Brouwer is their anxiety about Dedekind’s
approach. Dedekind explicitly avoids Kant’s insistence on the role of INTUITION in
mathematics:

In speaking of arithmetic (algebra, analysis) as a part of logic I mean
to imply that I consider the number-concept entirely independent of the
notions or intuitions of space and time, that I consider it an immediate
result from the laws of thought. (Dedekind, [49], 44)

The direct reference to space and time as INTUITIONS is no doubt intended to refer to
Kant’s grounding of the various branches of mathematics in these two forms of pure

speaks of many spaces, one understands by that only parts of one and the same unique
space. And these parts cannot as it were precede the single all-encompassing space
as its components (from which its composition would be possible), but rather are only
thought in it. It is essentially single, the manifold in it, thus also the general concept
of spaces in general, rests merely on limitations. (Kant, [102], B39, italic emphasis
added)

In some sense, then, the origins of Indecomposability are in Kant. Although many of the math-
ematicians discussed in the section are inspired by Kant, this particularly strong requirement does
not find mathematical fulfillment until the early twentieth century.
INTUITION, viz. space and time.\textsuperscript{36}

To see Weyl’s own assessment of the orthodox conception of Dedekind, consider the following picturesque description:

To represent the continuous connection of the points, traditional analysis, \textit{given its shattering of the continuum into isolated points}, had to have recourse to the concept of a \textit{neighborhood}. (Weyl, [211], 115; bold emph. added)

From a historical perspective, Weyl is correct. The concept of a neighborhood (effectively, a set of points around a point) gives rise to a whole topology \textit{on} the continuum, $\mathbb{R}$. As Weyl suggests, the introduction of a topology \textit{on} $\mathbb{R}$ is necessary since the orthodox conception of the continuum—called ‘traditional’ here by Weyl—itself is very minimal. A topology $\mathcal{T}$ on $\mathbb{R}$ is a family of sets (called \textit{open sets}) of $\mathbb{R}$ meant to be viewed (very loosely) as neighborhoods and sums of neighborhoods of $\mathbb{R}$.\textsuperscript{37} To use Weyl’s language, these open sets are principled collections of the shattered particles of the continuum. Weyl’s frustration with orthodoxy bespeaks his commitment to viewing the continuum from the top down. The orthodox conception has to impose \textit{additional} structure on the continuum in order to get it to “come together” properly. Weyl believes that the continuum should already be together. If structure is to be added, it should be in dividing the continuum. Either way, Weyl’s suggestion that the orthodox continuum is shattered echoes Poincaré’s words above: “only the multiplicity remains, the unity has disappeared.” (Poincaré, \textit{loc. cit.})

Weyl and Brouwer attempt a return to a Kantian conception of mathematical thought broadly, and to a more \textit{intuitive} continuum in particular. Weyl praises Brouwer’s mathematical program with the following Kantian sentiments:

With Brouwer, mathematics gains the highest intuitive clarity; his doctrine is idealism in mathematics thought to an end (Weyl, [212], 136).

\textsuperscript{36}For detailed discussion of this topic, see Shabel’s [180].
\textsuperscript{37}Explicitly, $\mathcal{T} \subseteq \mathcal{P}(\mathbb{R})$, where $\emptyset, \mathbb{R} \in \mathcal{T}$ and $\mathcal{T}$ is closed under arbitrary unions and finite intersections. Using mereological language, it is arbitrary sums and finite products of neighborhoods.
...It is greatly beneficial that Brouwer has strengthened again the sense in mathematics for the intuitively given. (ibid., 141)

By referring to the likes of ‘the intuitively given’, Weyl is no doubt invoking Kant, specifically his epistemology of mathematics and his insistence on the place of intuition therein.\footnote{38}

From this return to an intuition-sensitive epistemology of mathematics, a very robust picture of the continuum emerges. Weyl writes:

Within a continuum, one can very well generate subcontinua by introducing boundaries; yet it is irrational to claim that the total continuum is made up of the boundaries and subcontinua.

The point is, a genuine continuum is something connected in itself, and it cannot be divided into separate fragments; this conflicts with its nature. (Weyl, [211], 111; bold emphasis added)

The above quotations are somewhat interpretatively delicate, given that they seem to suggest opposite things about the continuum. The first suggests that one can generate subcontinua, thereby suggesting that one can divide the continuum; and the other suggests that one cannot divide the continuum at all. The most straightforward way of harmonizing these sits perfectly with Indecomposability. The first passage is almost a rephrasing of the statement of Indecomposability. Regarding the second passage, the appropriate reading of “cannot be divided into separate fragments,” is not that there are no parts of the continuum—a tempting way to read things. Rather, the parts cannot stand alone ontologically and still count as parts of the continuum. In other

\footnote{38The nature of these Kantian invocations is noteworthy because Kant’s own view of continuity is similar to Aristotle’s, though stronger. What makes these mathematicians’ interest in Kant so significant is their eagerness to adopt his much broader philosophical outlook, rather than specific features of his philosophy of mathematics. Brouwer in particular acknowledges some of the historical Kant’s difficulties:} 

However weak the position of intuitionism seemed to be after this period of mathematical development, it has recovered by abandoning Kants apriority of space but adhering the more resolutely to the apriority of time. (Brouwer, [26], 57)

As such, what is evident is that these mathematicians are mostly looking to Kant for philosophical rather than strictly mathematical guidance.
words, if the parts of the continuum are separated at all they are no longer parts of the continuum.\textsuperscript{39} In any case, Weyl clearly defends a picture of the continuum where the parts of the continuum take an ontological backseat while the continuum itself is up front in the driver’s position—a very top-down view, indeed.

\subsection*{2.4.2 Reasons for Concern}

The most striking problem for this viewpoint is that all extant theories of continua that satisfy this requirement require a logic no stronger than intuitionistic logic. The typical way of understanding parts in a set-theoretic universe is by treating set inclusion as the parthood relation and arbitrary unions as sums. That is, $A \subseteq B$ iff and only if $A$ is a part of $B$; and

$$D = \bigcup_{\alpha \in I} C_{\alpha},$$

just in case $D$ is the sum of the $C_{\alpha}$’s, where $I$ is some indexing set. Now, I will show how one instantly confronts a contradiction with this interpretation in combination with the assumption of the Law of Excluded Middle (\textsc{lem}). A natural choice for two disjoint proper parts would be the intervals $(-\infty, 0)$ and $[0, \infty)$. Obviously, these are disjoint, but does their sum compose the continuum? Denote the continuum with $\mathcal{C}$. Now, let $x$ be any point on the continuum, i.e. $x \in \mathcal{C}$. By \textsc{lem}, $x < 0$ or $x \geq 0$. Therefore, either $x \in (-\infty, 0)$ or $x \in [0, \infty)$. In other words, $x \in (-\infty, 0) \cup [0, \infty)$. Therefore, $\mathcal{C} \subseteq (-\infty, 0) \cup [0, \infty)$. It follows that $\mathcal{C} = (-\infty, 0) \cup [0, \infty)$. But by \textit{Indecomposability}, $\mathcal{C}$ is not identical to sum of any of its disjoint proper parts—contradiction.\textsuperscript{40}

What this shows is that in order to satisfy the \textit{Indecomposability} requirement, one

\textsuperscript{39}Thanks to Lisa Shabel for her help in making sense of these passages.

\textsuperscript{40}A very similar argument is given by Carl Posy on p. 346 of his [152].
must either give up this natural interpretation of parthood within the set-theoretic setting, or one must abandon LEM.\footnote{Strictly speaking, there is another option: deny the existence of points. This option will be discussed at length below while examining Aristotle’s requirements on continuity (§2.6). Carl Posy argues that that a punctate continuum that satisfies \textit{Indecomposability} is a significant philosophical contribution from the intuitionists. He writes, “Brouwer has shown us that we can have a viscous [indecomposable] continuum and make it out of points as well. This is no windmill-tilting: this is a profound mathematical insight.” (Posy, [152], 347)} The extant theories satisfying \textit{Indecomposability} all choose the latter option. In each case, there are other reasons to abandon LEM that are closer to the conceptual center of the theories. So, this is not the only reason why LEM is abandoned. In any event, the extreme consequences of adopting such strong requirements on the continuum are often savored—rather than choked down—by their adherents.

Now, allow me to examine more directly the requirement itself; that is, beyond the necessitation of intuitionistic logic, what might be a concern for \textit{Indecomposability}? Recall the example of Book I Proposition 1 (§2.2.1): the existence of the point of intersection of the two circles is just presupposed from within Euclidean geometry. The general idea here is that when two continuous things (whether lines or curves) cross one another, there is actually a point of intersection. Moving away from Euclid, a very central result in ordinary analysis ensures this holds for curves in space (i.e., continuous functions): the intermediate value theorem. According to the intermediate value theorem, if a function, \( f : [a, b] \rightarrow \mathbb{R} \) is continuous on \([a, b]\) and \( f(x_1) > c \) and \( f(x_2) < c \) for some \( x_1, x_2 \in [a, b] \), then there is some \( x_3 \in [a, b] \) such that \( f(x_3) = c \). For the specific formal system I will examine below, SIA, the intermediate value theorem is not merely unprovable, it is false. For simple quadratics, the typical tools of completing the square or using the quadratic formula suffice for determining where a curve crosses the \( x \)-axis. However, even for polynomials of degree 3, the intermediate value theorem is provably false.\footnote{Details of this can be found in Moerdijk and Reyes, [136], pp. 317-318.} This is a very negative result when it comes to the practical facility of SIA. Similar problems plague the other systems.
satisfying *Indecomposability*.\(^{43}\)

As a final note, with regard to the difficulties presented by this particularly strong requirement on continuity, I want to quote Weyl one last time. Below, Weyl expresses pluralistic sentiments with regard to the continuum:

> It is greatly beneficial that Brouwer has strengthened again the sense in mathematics for the intuitively given. His analysis expresses in a pure manner the content of the mathematical basic intuition and is therefore shone through by clarity without mystery. Yet beside Brouwer’s way, one will also have to pursue that of Hilbert; for it is undeniable that there is a theoretical need, simply incomprehensible from the merely phenomenal point of view, with a creative urge directed upon the symbolic representation of the transcendent, which demands to be satisfied. (Weyl, [212], 141)

Weyl seems to appreciate the overall limitations generated by views like his and Brouwer’s to the point where he clears the way for other views to be developed. In spite of the fact that Weyl’s reasoning for his pluralism about the continuum are quite distinct from my own, he appears to appreciate that more is better.

### 2.4.3 Fully Formalized

As I indicated above, I am looking most closely at *sia*, rather than Weyl’s own system. Since the philosophical literature from Weyl is strong, John Bell commends Weyl’s commentary on the continuum through regular, direct citation while discussing

---

\(^{43}\)In fact, if the intermediate value theorem were generally true, then the continuum would be decomposable. The idea is that the parts of \( f \geq c \) and \( f < c \) would compose the image of \( f \) on \([a, b]\), written \( f([a, b])\). But \( f([a, b])\) is an interval and even intervals are indecomposable. For more detail, see the next subsection. Of course, linear and quadratic functions do satisfy the theorem. To some extent, this is an accident and I want to suggest that the fact that quadratics satisfy the theorem shows something quite different. Since there is actually an algorithm for determining the roots, this allows one to divide the continuum at the roots by actually removing the roots. Once a determinate point is removed, the parts of the remainder do compose the remainder, which is not true (indeed, refutable) in constructive analysis. Generally, the indecomposability of *sia* is even less extreme than for other extant theories satisfying *Indecomposability*. For details of this, see Bell’s “The Continuum in Smooth Infinitesimal Analysis,” [7], compared with van Dalen’s “How Connected is the Intuitionistic Continuum?” [206].
Indecomposability and SIA.\textsuperscript{44} Here I will focus directly on SIA and then see how well it fits Indecomposability.

Basic Presentation

The origins of SIA are from work done within a relatively new branch of mathematics, category theory. Fortunately, even for one without a background in category theory, John Bell also provides an axiomatic theory of SIA. First, he specifies that the consequence relation is intuitionistic, the details of which I will omit here. The first requirement placed on SIA past the logic is that $\mathbb{R}$ is an ordered field.\textsuperscript{45} Due to some of the logically sensitive features of $\mathbb{R}$, order is defined slightly more carefully than it is above:

1. Irreflexivity: For all $x \in \mathbb{R}$, it is false that $x < x$.
2. Semi-antisymmetry: If $x \neq y$, then either $x < y$ or $y < x$.
3. Transitivity: For all $x, y, z \in \mathbb{R}$, if $x < y$ and $y < z$, then $x < z$.
4. Additivity: For any $x, y, z \in \mathbb{R}$, if $y < z$, then $x + y < x + z$.
5. Generalized Positivity: For any $x, y, z \in \mathbb{R}$, if $x < y$ and $z > 0$, then $x \cdot z < y \cdot z$.
6. Semi-completeness: For all $x \in \mathbb{R}$, either $x > 0$ or $x < 1$—notably, $0 < 1$.
7. Roots: For all $x \in \mathbb{R}$, $x > 0$ then there is some $y \in \mathbb{R}$ such that $x = y^2$.

Part of the reason that special care needs to be taken is that $\mathbb{R}$ cannot be totally ordered; i.e. it does not satisfy the requirement that for any $x, y \in \mathbb{R}$, exactly one of the following obtains: $x < y, y < x$ or $x = y$. This is in particular due to a very special subset of $\mathbb{R}$, $\Delta$ and its properties. $\Delta$ is the set of nilsquares, or those values

\textsuperscript{44}It is also worth noting that the mathematical literature on Weyl’s own views is somewhat limited. His book *The Continuum*, [210], uses an idiosyncratic notation, rendering it nearly impervious to an uninitiated reader. Solomon Feferman offers an excellent discussion of this text in his “Weyl Vindicated: *Das Kontinuum* seventy years later” in [65]. Feferman also takes care to translate Weyl’s notation into contemporary mathematical notation. See also Dirk van Dalen’s “Hermann Weyl’s Intuitionistic Mathematics,” [207].

\textsuperscript{45}See §2.3.3 for the field axioms, numbered 1.-7..
ε such that ε² = 0. The property that is at odds with a total order is sometimes referred to as the Kock-Lawvere Property (KLP):

\[(\forall f : \Delta \rightarrow \mathbb{R})(\exists! m \in \mathbb{R}) (\forall \varepsilon \in \Delta) \ f(\varepsilon) = f(0) + \varepsilon \cdot m.\]

The idea behind KLP is that for function f, there is a tangent line for f with slope m at point 0 spanning the nilsquares. Notice that KLP is the natural representation for differentiation and therefore velocity within SIA where \(m = f'(0)\), in conventional notation. What is most peculiar is that combining the field axioms with KLP yields a denial of totality. The proof of this is found below in the chapter on Zeno’s Arrow, so I will not repeat it here. Nevertheless, this lack of total order of SIA is among its most distinctive, but also among its most detrimental, features.

Finally, the last axiom is called the Constancy Principle (CP). One of the primary features of SIA is that all functions are smooth—hence, Smooth Infinitesimal Analysis. (CP) guarantees that constant functions—functions that have the same value everywhere—are well-behaved. Officially, CP states that for all functions, f:

\[(\forall x \in \mathbb{R})(\forall \varepsilon \in \Delta) f(x + \varepsilon) = f(x) \implies (\forall x, y \in \mathbb{R}) f(x) = f(y)\]

More intuitively, this states that if the derivative of f is everywhere zero, then f is constant. From a purely mathematical points of view, the major need for this principle is in application of the Fundamental Theorem of Calculus. It turns out

\[46\text{Allow me to note two things. First, this can be generalized by simple linear transformations to apply to any point, not just zero. Second, the reason KLP is written as an equation in product is that } 0 \in \Delta \text{ and so one cannot simply divide by all of the } \varepsilon \in \Delta \text{ as in the usual expression of the derivative as a quotient. That is, since } 0 \in \Delta \text{ and } \varepsilon \text{ ranges over } \Delta, \text{ the following expression is bogus in SIA:} \]

\[m = \frac{f(\varepsilon) - f(0)}{\varepsilon}\]

\[47\text{Very casually, the Fundamental Theorem of Calculus states that the area under a given curve between two points is equal to the difference between the values of the anti-derivative evaluated at}\]
that CP is also used in proving *Indecomposability*.

**Satisfying the Requirement**

In order to demonstrate how this requirement is satisfied, I must offer an interpretation of it in this setting. First, here is a reminder of *Indecomposability*:

2b. *Indecomposability*—An object is continuous only if it is not identical to sum of any of its disjoint proper parts.

Now, in effect, to prove that $\mathbb{R}$ is continuous, it suffices to show that for all disjoint $U, V \subseteq \mathbb{R}$ such that $U \cup V = \mathbb{R}$, either $U = \emptyset$ or $V = \emptyset$; i.e., $U$ or $V$ is all of $\mathbb{R}$.

To show that SIA satisfies *Indecomposability*, suppose for reductio that non-empty $U, V$ partition the continuum; that is, $U \cap V = \emptyset$, $U \cup V = \mathbb{R}$ and $U, V \subsetneq \mathbb{R}$. Define $f : \mathbb{R} \to \{0, 1\}$ as follows:

$$f(x) = \begin{cases} 0 & \text{if } x \in U \\ 1 & \text{if } x \in V \end{cases}$$

First, one must note that in SIA, all functions are smooth.\(^{48}\) A function is smooth just in case for any $c \in \mathbb{R}$ and any positive integer $n$, the $n$-th derivative of $f$ exists at $c$. Now, since all functions are smooth, $f$ is in particular. It follows that $f$ is continuous. In this setting, this means that $f(x) - f(y) \in \Delta$ whenever $x - y \in \Delta$.

Now, I claim that $f(x + \varepsilon) = f(x)$. Since $U$ and $V$ partition the continuum, then $f(x + \varepsilon) \in \{0, 1\}$ and $f(x) \in \{0, 1\}$. Since $f$ is continuous, either $f(x) = f(x + \varepsilon) = 0$ or $f(x) = f(x + \varepsilon) = 1$—in either case, $f(x) = f(x + \varepsilon)$. Generalize over $x$ and $\varepsilon$ and apply CP. It follows that $f$ is constant. If $f$ is constant, then either $U = \emptyset$ or $V = \emptyset$, contradicting that $U$ and $V$ are nonempty.

\(^{48}\)This is in fact guaranteed by the fact that KLP applies to all functions. By a simple linear transformation, one can “recenter” $f$ around the origin thus applying KLP to all points. That is, if one wants to apply KLP to $f$ at some point $c$, then consider KLP for the function $f(x - c)$. 

39
Chapter 2.5

Infinitesimal Magnitudes

Leibniz writes the following in the introduction to his *Theodicy*:

There are two famous labyrinths where our reason very often goes astray: one concerns the great question of the Free and the Necessary, above all in the production and origin of Evil; the other consists in the discussion of continuity and of the indivisibles which appear to be the elements thereof, and where the considerations of the infinite must enter in. (Leibniz, [116], 53)

Unfortunately for us, Leibniz never laid down a systematic *philosophical* presentation of his view of the continuum. This is a shame given the centrality with which he seemed to view the question of continuity and the extent to which he wrote on it. Leibniz gave continuity much less philosophical attention than he gave to other topics in which he had comparable interest. In what follows, I will try to examine to some extent some of Leibniz’s views on continuity, as found in his mathematical writings. Specifically, one finds him putting infinitesimals to great use in his own system of the calculus.

It should be noted that I focus specifically on Leibniz because of the central part that he played in giving birth to the calculus and because I prefer to focus on a single representative for the view listed above:

2c. *Infinitesimal Magnitudes*—An object is continuous only if it is composed of infinitesimal magnitudes.

To introduce other supporters of this view would not only be cumbersome but it would also fail to provide any significant increase in richness sometimes found through examining numerous figures on a single topic.\(^{49}\)

\(^{49}\)C.S. Peirce is another figure who took infinitesimals very seriously. However, his view is very distinctive and I would not actually count him as supporting *Infinitesimal Magnitudes*, strictly speaking. His view falls somewhere between *Indecomposability* and *Infinitesimal Magnitudes* by flooding the continuum with innumerably many potential point values. For deep discussion, see Philip Ehrlich’s “The Absolute Arithmetic Continuum and its Peircian Counterpart.”
Lastly, I do not intend to broach the especially thorny issues of Leibniz’s fictionalism about infinitesimals. Samuel Levey discusses these subtleties at length in his “Archimedes, Infinitesimals and the Law of Continuity: On Leibniz’s Fictionalism.”

My approach to Leibniz will take the following report from Levey for granted:

From an ontological point of view, the infinitesimals of his mathematics are taken merely to be fictions . . . From the point of view of mathematical practice, however, infinitesimals are not discarded but retained and actively promulgated. (Levey, [123], 121)

Given this fact, I take Leibniz’s infinitesimals seriously, in spite of his own “official” hesitations. From my perspective, the success of the practice speaks for itself. In mathematics, as in other sciences, the ends justify the means. The foundations of a discipline need not be in place for a practice to be received as mathematically legitimate—where legitimacy is meant to be slightly broader than straightforward truth. To some extent, this seems to be Leibniz’s own view:

Even algebraic calculation cannot avoid [infinitesimals] if it wishes to preserve its advantages, one of the most important of which is the universality which enables it to include all cases, even that where certain given lines disappear. It would be ridiculous not to accept this and so deprive ourselves of one of its greatest uses. (Leibniz, [117], 886)

Without entangling myself in the details of the cited example, Leibniz is claiming that to avail oneself of the full power of the standard algebraic formalism, one must accept—even if only in pretense—the existence of infinitesimal magnitudes. Or,

---

50See [123]. Similar discussion is also found throughout the collection in which Levey’s essay is found, *Infinitesimal Differences*, [80].

51Hilbert writes, “[I]n mathematics as elsewhere success is the supreme court to whose decisions everyone submits.” ( [91], 184)

52As further evidence suggesting Leibniz’s pragmatic approach, consider the following remark in his letter to Varignon:

Even if someone refuses to admit infinite and infinitesimal lines in a rigorous metaphysical sense and as real things, he can still use them with confidence as ideal concepts which shorten his reasoning, similar to what we call imaginary roots in the ordinary algebra, for example, $\sqrt{-2}$. Even though these are called imaginary, they continue to be useful and even necessary in expressing real magnitudes analytically. (Leibniz, [117], 882-883)
put positively, the advantage of infinitesimals is their ease of use in ordinary algebraic manipulations in comparison with the more arduous techniques of traditional geometry.

2.5.1 Motivations

A motivating picture for Leibniz is to think of a circle as an infinilateral regular polygon—an equilateral, equiangular polygon with infinitely-many infinitesimally extended sides.\(^{53}\) Observe the sequence of regular polygons in Figure 2.2.\(^{54}\) Notice that even at 12 sides (the fourth figure), a regular polygon already begins to resemble a circle. In a sense, the circle is composed of infinitesimally tiny lengths. In that way, this requirement is a bottom-up approach—regardless of Leibniz’s official word on the matter.\(^{55}\)

![Figure 2.2. Sequence of Regular Polygons Converging to a Circle](image)

In spite of the fact that Leibniz does not offer enough of a systematic philosophical presentation of the infinitesimals, he says a whole lot more than most of his mathematical contemporaries. In order to underwrite and license conceiving of circles and curves as composed of infinitesimals, Leibniz introduces his Law of Continuity,

\(^{53}\)Leibniz is not the first person to think of a circle this way. Several sources place the origination of this notion with Democritus. Minimally, Democritus grappled with questions in the neighborhood. Many others in the immediate historical vicinity of Leibniz treated circles and other similar curves—e.g., cycloids—as infinilateral polygons, including Descartes, Kepler and the Bernoullis. For more details, see Bos’ “Differentials, Higher-Order Differentials and the Derivative in the Leibnizian Calculus,” [21]; Boyer’s The History of the Calculus and its Conceptual Development, [22]; Edwards’ The Historical Development of the Calculus, [55]; Katz and Sherry’s “Leibniz’s Infinitesimals: Their Fictionality, Their Modern Implementations, and Their Foes from Berkeley to Russell and Beyond,” [103].

\(^{54}\)For each, I inscribed them in a black circle to reveal that the “error” decreases as the number of sides increases.

\(^{55}\)Let the reader note that two contemporary expressions of this requirement are both punctate.
described in a variety of ways. In a letter to the adversarial Nieuwentijt, Leibniz writes:

I take for granted the following postulate: 

In any supposed transition, ending in any terminus, it is permissible to institute a general reasoning, in which the final terminus may also be included. (Leibniz, [115], 147, emph. original)

To get a sense for how strong this statement is meant to be, consider how Leibniz puts his Law of Continuity in another (more casual) way in a letter to Varignon, “the rules of the finite are found to succeed in the infinite.” (Leibniz, [117], 884) One might put this as follows: if a sequence \((a_n)\) converges to some value \(T\) and a property \(\varphi\) is true of each \(a_i\), then \(\varphi\) is true of \(T\).\(^{57}\) This model can be applied to infinilateral polygons. On this model, each \(a_i\) is regular polygon with \(i \geq 3\) many sides and \(T\) is a circle. Since geometrical properties of lines and polygons are fairly well behaved, they can be more easily studied than those of curves and circles. So, on this model, the \(\varphi\) will be an algebraic representation of some geometrical property of a polygon. Then, since \(\varphi\) applies to each \(a_i\), Leibniz’s Law of Continuity permits one also to apply \(\varphi\) to \(T\).

Consider the following concrete case. One can discover the area of a regular polygon by summing all those isosceles triangles whose bases are the sides of the polygon and whose vertices are the center of the polygon, as in Figure 2.3. I will use features of triangles to derive the area of a circle.

\(^{56}\)Another version of this is found in a reply to Malebranche:

When the difference between two instances in a given series or that which is presupposed can be diminished until it becomes smaller than any given quantity whatever, the corresponding difference in what is sought or in their results must of necessity also be diminished or become less than any given quantity whatever. (Leibniz, [117], 539)

Leibniz sometimes translates his talk the infinite into language that strongly resembles the contemporary epsilon-delta formulation.

\(^{57}\)I do not actually take this to be how Leibniz would describe his principle. Neither would I say that Leibniz anticipated in any serious sense a contemporary theory of sequences. I take the above to be a helpful way of describing things for contemporary readers who are familiar with sequences, and more generally modern logico-mathematical language. In this way, one might describe the above description as forming a kind of “fusion of horizons” between Leibniz and us.
The geometric property is codified by the principle that for any number of sides, $n$, size of the side $b$ and the height $h$ of the inner triangle:

$$A = n \left( \frac{1}{2}bh \right)$$

Suppose $h$ remains constant. The base will then be dependent upon the number of sides, $n$ and $\theta$ is half the vertex angle of the isosceles. Note that $\tan \theta = \frac{1}{2}b$. So, $b = 2h \tan \theta$. Generally, $\theta = \frac{2\pi}{2n} = \frac{\pi}{n}$, for any $n$. If $n$ is infinite, then $\theta$ is infinitesimal. Now, whenever $\theta$ is infinitesimal, $\tan \theta = \theta$. It follows that when $n$ is infinite,

$$A = n \left( \frac{1}{2}bh \right) = n \left( \frac{1}{2} \cdot (2h \tan \theta)h \right) = nh^2\theta = nh^2 \cdot \frac{\pi}{n} = \pi h^2.$$

So, by using Leibniz’s Law of Continuity, one may derive the well-known area of a circle, as I just demonstrated.

Leibniz’s Law of Continuity is meant to apply even to examples that one would not expect. There are several other cases beyond the infinilateral polygon to which he applies his Law. For example, he uses his Law to analyze the behavior of parallel lines in such a way that they behave like other pairs of lines. In so doing, he treats all lines as intersecting, where parallel lines intersect “at infinity,” and the angle between

\footnote{This is a lesser known fact exploited by people in this time period. One can show similar results using contemporary techniques.}
them is infinitesimal. Again, using my gloss from above, let \((a_n)\) be a sequence of distinct pairs of intersecting lines each with an angle of \(\frac{1}{n}\) radians between them. The terminus \(T\) of \((a_n)\) is a pair of lines with zero angle between them. Furthermore, each member \(a_i\) of sequence \((a_n)\) satisfies the property, \(a_i\) comprises two distinct lines. Hence, \(T\) comprises two distinct lines; that is, the lines of \(T\) are not coincident but parallel. Such examples might make others balk, but this strange Law ended up paying serious mathematical dividends.

Again, although I do not want to get too deep into the nature of Leibniz’s fictionalism, one should note that Leibniz does view the Law as normatively binding in some respect. In a letter to Varignon, he attaches a small discussion, “Justification of the Infinitesimal Calculus by that of Ordinary Algebra.” There he writes:

Although it is not at all rigorously true that . . . a circle is a kind of regular polygon, it can be said, nevertheless, that . . . the circle terminate[s] . . . the regular polygons which arrive at [it] by a continuous change and vanish in [it] . . . And although these terminations are excluded, that is, are not included in any rigorous sense in the variables which they limit, they nevertheless have the same properties as if they were included in the series, in accordance with the language of infinites and infinitesimals. Otherwise the law of continuity would be violated, namely, that since we can move from polygons to a circle by a continuous change and without making a leap, it is also necessary not to make a leap in passing form the properties of polygons to those of a circle. (Leibniz, [117], 887, emph. added)

Here Leibniz reasons that one must apply the properties of the various members of the sequence to the infinite case and that to fail to do so would be to violate the Law of Continuity. Leibniz takes the Law seriously enough that he cites it here without qualification in defense of his infinilateral polygon. So, whatever nature Leibniz’s fictionalism has, one commits some sort of normative failure in refusing to extend the properties of some sequence to its limit.
2.5.2 Reasons for Concern

There are a few concerns for *Infinitesimal Magnitudes*. For many centuries, there was a lingering anxiety about, if not hostility towards, infinitesimals. One of the biggest concerns was that infinitesimals are inconsistent. As sections 5.3 and 5.4 will suggest, the problem for *Infinitesimal Magnitudes* has much less to do with its mathematical consistency. The outstanding concern for *Infinitesimal Magnitudes* is its violation of the Archimedean property. Informally, the Archimedean property suggests that all magnitudes are commensurable. More formally, for any values $x$ and $y$ where $x < y$, there is some positive integer $n$ such that $nx > y$. The idea is that if there is some difference between two numbers, the difference can be made up by integral multiples. In contrast, infinitesimals are incommensurable by their very nature. In some number systems, the infinitesimals are explicitly defined to be violations of the Archimedean property: $\varepsilon$ is an infinitesimal just in case $\varepsilon \neq 0$ and for any positive integer $n$, $\varepsilon < \frac{1}{n}$; i.e. $n\varepsilon < 1$. I personally am not moved by these concerns, though people like Cantor viewed the Archimedean property to be constitutive of the concept of linear magnitude—or, in effect, basic measurable parts of the continuum.

Cantor had a vitriolic antipathy for infinitesimals, calling them the “cholera bacillus of mathematics.”\(^{59}\) J.W. Dauben suggests that such acrimony was motivated by fear that other theories concerning both infinite and infinitesimal values alike would dethrone Cantor’s own theory of transfinite numbers. History has shown that all of these elegant mathematical concepts can exist harmoniously, if not symbiotically. What then is the nature of Cantor’s complaints? A full treatment is provided in §8 of Philip Ehrlich’s, “The Emergence of non-Archimedean Systems of Magnitudes.” Below, I will extract what I take to be the core concern from Ehrlich’s translation of

---

\(^{59}\)Cantor put it thus in a letter to Vivanti found in [135] and is quoted in both Philip Ehrlich’s “The Emergence of non-Archimedean Systems of Magnitudes,” [59], and also Joseph W. Dauben’s chapter 5, “The Development of Cantorian Set Theory,” in [81].
The basic idea is that Cantor believed the Archimedean property was not merely an axiom to be held or not, but rather it is essential to the concept of a linear magnitude. By analogy, this would be comparable to arguing that Euclid’s Parallel Postulate is constitutive of space: it follows from a more fundamental representation of space. What then does Cantor say about these linear magnitudes? There is a pair of places where one finds Cantor describing linear magnitudes. Consider first Cantor’s description of a linear number: numbers that “may be regarded as bounded, continuous lengths of straight lines.” (Cantor, translated in Ehrlich [59], 41, emphasis added) Presumably, a linear number is just the arithmetic companion to the geometric linear magnitude. So, one may take a linear magnitude to be the length of a bounded, continuous straight line. Notably, this characterization does not immediately preclude infinitesimals. This next passage sharpens one’s sense for the concept of a linear magnitude:

Thus the assumption we made [that an infinitesimal magnitude exists (for reductio)] contradicts the concept of a linear magnitude, which is of the sort that according to it each linear magnitude must be thought of as an integral part of another, in particular of finite linear magnitude. (Cantor, loc. cit., 42, emphasis added)

Here Cantor describes something suspiciously similar to the Archimedean property by requiring that “each linear magnitude . . . be thought of as an integral part of another.” This requires that any linear magnitude, \( X \), with length \( x \), when summed with an identical magnitude yields another magnitude of length \( 2x \), and this process can be iterated indefinitely to \( 3x, 4x, 5x, \ etc. \). Furthermore, the above passage implies that such an \( X \) be an integral part of a finite linear magnitude. In other words, for some positive integer \( n \), \( n \)-many \( X \)’s will compose a finite linear magnitude—impossible for an infinitesimal. One might begin to suspect that Cantor is begging the question.

\(^{60}\)The relevant translation is found on pp. 41, 2 of this paper, with the German text appearing as Appendix IV.
here. His description of a linear magnitude seems to take for granted a version of the Archimedean property. Is he begging the question?

In the paper mentioned above, Ehrlich does an excellent job of charitably interpreting Cantor by providing a very careful algebraic presentation of linear numbers (and by extension, magnitudes). From this, Ehrlich actually derives the impossibility of infintesimals. Ehrlich’s work is cogent and to complement it, I want to offer a distinct argument for the same conclusion that Cantor is not begging the question. Based on some textual evidence in a passage in Cantor’s Grundlagen,\textsuperscript{61} I believe that Cantor has an epistemology of mathematics with striking Kantian features, and in particular that Cantor takes the Archimedean property to be given in intuition. As evidence of this warmth towards Kantian intuition, consider his remarks on another mathematical concept:

Here I confine myself to the demonstration that from the concept of well-ordered set the fundamental operations for the integers . . . arise in the simplest manner and that the laws governing them can be derived from immediate inner intuition [German, Anschauung, i.e. Kant’s intuition] with apodictic certainty. (Cantor, [32], 886, bold emphasis added)

The phrase “immediate inner intuition with apodictic certainty” is unmistakably Kantian.\textsuperscript{62} That Cantor would offer intuition as providing the grounds for “the laws” or constitutive principles for any mathematical concept gives me strong reason

\textsuperscript{61}This is short for Grundlagen einer allgemeinen Mannigfaltigkeitslehre, or Foundations of a General Theory of Manifolds.

\textsuperscript{62}I should caution, however, that Cantor is not nearly as Kantian as Weyl, and even expresses some negativity within the Grundlagen towards the specific features of Kant’s characterization of intuition, \textit{viz.} the role of space and time. Indeed, §8 from the Grundlagen is often viewed as a manifesto for a revolution against intuition. Here is one especially famous comment: “the essence of mathematics lies precisely in its freedom.” (Cantor, loc. cit., 896)

I was surprised to discover from a brief review of the literature on Cantor’s Grundlagen that the tension between the revolutionary passages and his positive remarks on intuition has not received much attention. Oddly enough, Dauben in his [44] cites the above passage, makes a promissory note that he will return to the thorny passage in the subsequent chapter and then never does. The only direct mention of this tension that I have discovered is found in a footnote in W.W. Tait’s “Cantors Grundlagen and the Paradoxes of Set Theory,” [197]. Since it is mostly irrelevant to his paper, Tait’s remarks are extremely brief. A related but distinct tension within the Grundlagen is discussed in Ignasi Jané’s “Idealist and Realist elements in Cantor’s Approach to Set Theory,” [99].
to suspect that he treats linear magnitudes in a similar fashion. What is clear is that Cantor does not have a strictly formalistic, axiomatic understanding of the concept of linear magnitude. Such a perspective puts him at odds with those eager to completely eschew intuition.

Laying aside the details of the historical Cantor's response to infinitesimals, I do not accept this perspective on magnitudes. I take a broadly Hilbertian stance on mathematical concepts. To require more of a concept of linear magnitude would be comparable to insisting that all spaces satisfy the parallel postulate. In the following subsections, two distinct consistent formal systems of infinitesimal magnitudes will be provided.

2.5.3 Infinitesimals in The Hyperreals

Although Abraham Robinson's development of the hyperreals in nonstandard analysis is not the first rigorous development of infinitesimals, it is probably the best known. Here I will present the hyperreals in a fashion similar to the one found in Abraham Robinson's original monograph, Non-standard Analysis, [168].

\[63\] Even if Cantor does not follow Kant to the conclusion that space and time constitute the substance of intuition, my interpretation offers some sense for why he would declare infinitesimals inconsistent. He seems to have something more than mere consistency guiding why he would in the first place insist on such a strong notion of linear magnitude, viz. infinitesimals conflict with the nature of linear magnitudes. The idea that Cantor saw them as inconsistent simpliciter cannot be maintained because numerous contemporaries of Cantor's had already developed non-Archimedean systems, including Paul du Bois-Reymond, Tullio Levi-Civita and Otto Stolz. The issue for Cantor seemed to be not the models but the numeric or geometric interpretation of these models (e.g., viewing classes of functions—themselves consistent—as magnitudes.) This work is discussed in exhaustive (but not exhausting!) detail in Ehrlich's [59].

\[64\] A great historical irony: not only is the parallel postulate unnecessary, it is physically false.

\[65\] Extended discussion can be found in both Ehrlich's [59] and Detlef Laugwitz's [112]. Comparing the hyperreals to a more simple system of infinitesimals, Laugwitz writes, "Abraham Robinson's Nonstandard Analysis ... was much more powerful when it came to contemporary research," (Laugwitz, [112], 128)

\[66\] I mostly follow Robert Goldblatt's Lectures on the Hyperreals, [79]. Like most mathematics, ground-breaking working rarely has the cleanest exposition. For this reason, I will not be as closely consulting Robinson's text, even though Goldblatt's techniques are fairly similar.
Basic Presentation

The most formidable and central feature of the hyperreals is ultrafilters. Here, I will examine ultrafilters on the set of natural numbers, \( \mathbb{N} \). But, first things first. What is an ultrafilter? An ultrafilter \( \mathcal{F} \) on a set \( X \) satisfies the following requirements:

1. \( \mathcal{F} \subseteq \mathcal{P}(X) \) and \( X \in \mathcal{F} \).
2. For any \( A, B \in \mathcal{F} \), \( A \cap B \in \mathcal{F} \).
3. For any \( A \in \mathcal{F} \) and \( B \supseteq A \), \( B \in \mathcal{F} \).
4. For any \( A \subseteq X \), either \( A \in \mathcal{F} \) or \( X \setminus A \in \mathcal{F} \).

Generically, an ultrafilter is too weak for the present purposes. More specifically, I must specify that the ultrafilters are nonprincipal: none of the \( A \in \mathcal{F} \) is finite. Below, one will see that this feature is essential to defining the hyperreals.

When it comes to defining the hyperreals (\( ^*\mathbb{R} \)), one must begin with sequences of real numbers, \( (a_n) : \mathbb{N} \to \mathbb{R} \). The set of sequences is also written \( \mathbb{R}^\mathbb{N} \). The \( ^*\mathbb{R} \) is an extension of \( \mathbb{R} \). The simplest way of representing a given member \( c \in \mathbb{R} \) within \( ^*\mathbb{R} \) is as the sequence \( (a_n) \) with \( a_n = c \) for all \( n \in \mathbb{N} \). One creates variety within \( ^*\mathbb{R} \) and extends \( \mathbb{R} \) with sequences that are not constant. These sequences on their own are not field extensions though. In order for these sequences to be better behaved, one must introduce an equivalence relation on them. Here is where the ultrafilters come in: agreement between sequences is determined by nonprincipal ultrafilters over \( \mathbb{N} \).

More specifically, let \( \mathcal{F} \) be such a filter.\(^{67}\) Define the equivalence relation \( \cong \) between sequences \( (a_n) \) and \( (b_n) \) as follows:

\[
(a_n) \cong (b_n) \iff \{ n \in \mathbb{N} : a_n = b_n \} \in \mathcal{F}
\]

\(^{67}\)It turns out that the hyperreal construction is dependent on the choice of \( \mathcal{F} \). Different hyperreals are isomorphic if one supposes the continuum hypothesis (CH). This is a rather surprising consequence given that CH is independent of ZFC. That is to say, the isomorphism between hyperreal constructions is independent of ZFC.
The idea is that \((a_n)\) and \((b_n)\) are identical “almost everywhere,” where “almost everywhere” is understood with respect to any of the members of \(\mathcal{F}\), all of which are infinite. In fact, \(\equiv\) is an equivalence relation.\(^{68}\) Officially, then, members of \(*\mathbb{R}\) are equivalence classes of sequences of reals modulo \(\mathcal{F}\). A construction of this sort where one defines equivalences classes on \(\mathbb{R}^\mathbb{N}\) using ultrafilters is called the ultraproduct construction.\(^{69}\)

**Satisfying the Requirement**

So far, I have not indicated what an infinitesimal looks like in this number system. As a matter of fact, the hyperreals are incredibly rich, including infinite and infinitesimal values. In order to capture this notion, I must introduce an ordering on \(*\mathbb{R}\). Again, I will begin with sequences and define an order on them. Let \(a, b \in *\mathbb{R}\), where \(a = [(a_n)]_\mathcal{F}\) and \(b = [(b_n)]_\mathcal{F}\). Now, let

\[
    a < b \iff \{ n \in \mathbb{N} : a_n < b_n \} \in \mathcal{F}
\]

It is not difficult to show that this is a total order, as described in the field properties in §2.3.3. In what follows, I will treat the sequences themselves as numbers, except in cases where it is necessary to be sensitive to the equivalence classes of sequences.

Now, to show that there is an infinite number and an infinitesimal, consider the following sequences:

\[
    \eta = \langle 1, 2, 3, 4 \ldots \rangle \quad \text{and} \quad \zeta = \langle 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots \rangle
\]

\(^{68}\)It is obvious that \(\equiv\) is symmetric and reflexive. To see that it is transitive, suppose \((a_n) \equiv (b_n)\) and \((b_n) \equiv (c_n)\). Since \(\mathcal{F}\) is an ultrafilter, it is closed under intersections, in particular, \(\{ n \in \mathbb{N} : a_n = b_n \} \cap \{ n \in \mathbb{N} : b_n = c_n \} \in \mathcal{F}\). This intersection is included in \(\{ n \in \mathbb{N} : a_n = c_n \}\). Since \(\mathcal{F}\) is closed upward under inclusion, then \(\{ n \in \mathbb{N} : a_n = c_n \} \in \mathcal{F}\). So, \((a_n) \equiv (c_n)\).

\(^{69}\)All foregoing details, including those in footnotes, are found in one form or another in [79].
Take any ordinary real number, \( c > 0 \), represented by \( \langle c, c, c, \ldots \rangle \). I will not prove it here, but I will use the fact that any cofinite set is a member of any nonprincipal ultrafilter. Now, the sets \( \{ n \in \mathbb{N} : n > c \} \) and \( \{ n \in \mathbb{N} : \frac{1}{n} < c \} \) are both cofinite and therefore \( c < \eta \) and \( c > \zeta \). This generalizes (since \( c \) is arbitrary), so \( \eta \) is greater than any real number—infinitive—and \( \zeta \) is less than any real number—infinitesimal.

Now, one can capture the requirement listed above by considering a string of infinitesimal intervals. In order to get these infinitesimal lengths to go to work, I will need integral multiples of them; for example,

\[
\cdots - 3\zeta, -2\zeta, -\zeta, 0, \zeta, 2\zeta, 3\zeta \ldots
\]

However, the “integers” here cannot only be the standard integers, \( \mathbb{Z} \). Those alone would not suffice to fill out a standard region. I need the hyperintegers \( \ast \mathbb{Z} \), a subset of \( \ast \mathbb{R} \) that includes \( \mathbb{Z} \) but extends them into infinite territory. For example, the number defined above, \( \eta \), is an infinite hyperinteger.

Now that all the resources of the hyperreals are available, the basic interval I want to consider is \([0, \zeta)\). In the series I am considering, the next interval would be \([\zeta, 2\zeta)\) and so on; the previous interval would be \([-\zeta, 0)\) and so on. More generally, the intervals I have in mind are of the form, \( A_n = [n\zeta, (n+1)\zeta) \) for all \( n \in \ast \mathbb{Z} \). The idea behind this approach is comparable to simply taking an interval, like \([0, \frac{1}{2})\), and multiplying it by all standard integers. The union of such intervals will compose \( \mathbb{R} \).

Similarly,

\[
\bigcup_{n \in \ast \mathbb{Z}} A_n = \ast \mathbb{R}^{70}
\]

This latter point is exactly what it takes to satisfy *Infinitesimal Magnitudes*; that is, the \( A_n \) composed \( \ast \mathbb{R} \). Of course, another size of infinitesimal would have done just

---

\(^{70}\)To see this, one must simply note that intervals \( B_n = [n, n+1) \) for \( n \in \ast \mathbb{Z} \) will cover \( \ast \mathbb{R} \). Now, suppose \( k \in \ast \mathbb{Z} \). \([k, k+1)\) is covered by \( A_{k\eta} \) through \( A_{(k+1)\eta - 1} \).
as well. The hyperreals have a plethora of infinitesimals from which to choose. The value \( \zeta \) is one of the more straightforward infinitesimals to define.

### 2.5.4 Infinitesimals in Giordano’s Fermat Reals

In this section, I will examine a number system recently developed by Paolo Giordano, the Fermat reals (\( \bullet \mathbb{R} \)). The system is similar to \( \mathbf{SLA} \) both because it contains nilpotents but also because it satisfies a version of the Kock-Lawvere property, \( \text{viz.} \) for open \( A \subseteq R^{71} \), and \( f : A \rightarrow R \) is smooth, then

\[
(\exists! m \in R) \ (\forall \delta \in \Delta) \ f(a + \delta) = f(a) + \delta \cdot m
\]

One crucial difference is that the consequence relation is classical.

#### Basic Presentation

The basic idea is that one treats any constant function \( f(t) = c \), for some \( c \in \mathbb{R} \) as playing the part of \( c \) in the new larger set of Fermat Reals, \( \bullet \mathbb{R} \). In this way, the Fermat reals are a nice generalization from \( \mathbb{N} \) to \( \mathbb{R} \): instead of constant sequences, Fermat reals use constant functions. One obtains greater diversity by allowing for functions over the reals to play the role of infinitesimals. To begin the construction, I must begin with a strict subset of the real-valued functions defined on the non-negative reals (\( \mathbb{R}^+ \)). Call these Fermat polynomials, \( \mathbb{R}_o[t] \).\(^{72}\) For notational purposes, it will sometimes be convenient to write \( x_t \) rather than \( x(t) \). I will use them interchangably.

To simplify the notation, I must introduce another set of functions, little-oh functions,\(^{72}\)

---

\(^{71}\)Here I use a generic letter \( R \) since the property could be stated either for either \( \mathbb{R} \), \( \mathbb{R} \) or in this case, the Fermat reals, Of course, it is false for \( \mathbb{R} \) on account of the non-uniqueness of \( m \) in \( \mathbb{R} \).

\(^{72}\)Giordano calls these little-oh polynomials. This is slightly confusing given some other terms he uses.
where \( o : \mathbb{R}^+ \to \mathbb{R} \) and
\[
\lim_{t \to 0^+} \frac{o(t)}{t} = 0.
\]

Little-oh functions are at the center of the following equivalence relation. Let \( x_t, y_t \in \mathbb{R}_o[t] \) and define \( x_t \sim y_t \) to mean that
\[
\lim_{t \to 0^+} \frac{x_t - y_t}{t} = 0. \tag{2.5.1}
\]

Another way of writing this (which will avoid gratuitous use of fractions) is
\[
x_t = y_t + o(t) \text{ as } t \to 0^+. \tag{2.5.2}
\]

The little-ohs enable both notational simplicity and technical maneuverability in the construction before washing them out with the equivalence relation.

Now, the official definition of a member of the Fermat polynomials is as follows. Each \( x_t \in \mathbb{R}_o[t] \) is such that one may write,
\[
x_t = \sum_{k=0}^{n} \alpha_k t^{b_k} + o(t) \text{ as } t \to 0^+ \text{ where } n \in \mathbb{N}, \alpha_k \in \mathbb{R}, b_k \in \mathbb{R}^+ \tag{2.5.3}
\]

These have the shape of ordinary polynomials but instead of positive integral powers, any positive power is eligible (hence, the \( b_k \)'s).\textsuperscript{73} The Fermat polynomials are not of significant interest in themselves: the goal is the Fermat reals, \( \bullet \mathbb{R} \). Loosely stated, \( \bullet \mathbb{R} \) comprises the Fermat polynomials where the \( o(t) \) is scrubbed out via the equivalence

\textsuperscript{73}Before using little-oh notation exclusively from now on, I remind the reader that the \( o(t) \) in (2.5.3) means that
\[
\lim_{t \to 0^+} \frac{x_t - y_t}{t} = 0
\]
where
\[
y_t = \sum_{k=0}^{n} \alpha_k t^{b_k}
\]
In other words, if \( y_t \) is the sum portion only, then one may add a little-oh function to it and still be within \( \mathbb{R}_o[t] \).
relation. Officially, $\bullet \mathbb{R}$ is defined to be the set of equivalence classes from the Fermat polynomials, $\mathbb{R}_o[t]$, under the relation $\sim$ from (2.5.1).

**Satisfying the Requirement**

Where are the infinitesimals? The get a sense for how infinitesimals look in $\bullet \mathbb{R}$, it will help to understand what a “standard part” is in $\bullet \mathbb{R}$. For the sake of analogy, let $x \in \mathbb{R}$ with $m \leq x < m + 1$ for integer $m$. From this, the *integral part* of $x$, written $[x]$, is the greatest integer less than or equal to $x$; here, $[x] = m$. In finding the integral part of a real value, I am effectively clipping off the decimal portion of the real value. Similarly, the “standard part” of a Fermat real (hyperreal) is when one clips off the infinitesimal portion of the Fermat real (hyperreal). The Hyperreals’ standard parts are defined exactly like the integral parts only one replaces greatest integer with greatest real. In the case of a Fermat real $y(t) \in \bullet \mathbb{R}$, the standard part is $y(0)$. The beauty of this definition is that the infinitesimal “fringe” is determined by the polynomial part with nonzero degree—the part which vanishes when $t = 0$—the standard part is the remainder when $t = 0$. In that case, one may define the infinitesimals of the Fermat reals to be those values,

$$D_\infty = \{ x \in \bullet \mathbb{R} : x(0) = 0 \}.$$

In other words, the infinitesimals are those values where the standard part is zero.

In order to have a concrete example, consider, $m \in \bullet \mathbb{R}$ where $m = [m_t]_\sim$ and $m_t = t \in \mathbb{R}_o[t]$. First, in line with the above definition, $m(0) = 0$. Moreover, $m$ is a

---

74 One might wonder, why all this talk of little-ohs in the first place? As I mentioned, it allows technical ease. Put another way, it is often much easier to work within $\mathbb{R}_o[t]$ than directly in $\bullet \mathbb{R}$. The little-ohs enable this, especially when the $b_k > 1$ in (2.5.3).
nontrivial nilsquare; i.e. \( m \neq 0 \) but \( m^2 = 0 \), since

\[
\lim_{t \to 0^+} \frac{t}{t} \neq 0,
\]

but also

\[
\lim_{t \to 0^+} \frac{m^2_t}{t} = \lim_{t \to 0^+} t = \lim_{t \to 0^+} t = 0.
\]

Now that the details are in place, the real question is, can the Fermat Reals satisfy \textit{Infinitesimal Magnitudes}? Although defining sets that cover \( \mathbb{R}^\dagger \) is not nearly as natural as it was for the case of the hyperreals, it can be done. Define for each \( r \in \mathbb{R} \),

\[
F_r = \{ m_t \in \mathbb{R}^\dagger : m(0) = r \}
\]

Such a set contains all the infinitesimals around \( r \in \mathbb{R} \). It follows that

\[
\bigcup_{r \in \mathbb{R}} F_r = \mathbb{R}^\dagger
\]

In that case, \( \mathbb{R}^\dagger \) satisfies \textit{Infinitesimal Magnitudes}.

Now, one might contest that the \( F_r \) are not properly speaking magnitudes since they do not have measure—a seemingly reasonable requirement on a magnitude. The reason no \( F_r \) has measure is that there is no lowest upper bound or greatest lower bound. No \( F_r \) is an interval, nor are they finitely additive—both attractive qualities for measurable sets to have. As it turns out, each \( F_r \) is countably additive, but that requires passing through the following alternative. This alternative way of covering \( \mathbb{R}^\dagger \) requires adding the concentric rings around \( r \in \mathbb{R} \) as follows. To begin, I take for granted that \( \mathbb{R}^\dagger \) is ordered and if \( m_t = t^{\frac{1}{m}} \) and \( n_t = t^{\frac{1}{n}} \) for positive integers \( m > n \), then \( m_t > n_t \).

Now, I define inductively for each \( r \in \mathbb{R} \) the necessary interval parts

\[75\text{This fact can be demonstrated straightforwardly using details from Giordano’s “Infinitesimals Without Logic,” [77], §§10, 11.}\]
as follows:

\[
F_{r,0} = [r - t, r + t]
\]

\[
F_{r,n+1} = [r - t \frac{1}{n+1}, r + t \frac{1}{n+1}] \setminus F_{r,n}
\]

First, note that each \( F_{r,n} \) is a magnitude in the sense above. In particular, \( F_{r,0} \) has measure 2\( t \) and \( F_{r,n+1} \) has measure 2\( t \frac{1}{n+1} - 2t \frac{1}{n} \). Furthermore, note that

\[
\bigcup_{r \in \mathbb{R}, n \in \mathbb{N}} F_{r,n} = \bullet \mathbb{R}.
\]

Again, in which case, \( \bullet \mathbb{R} \), satisfies *Infinitesimal Magntitudes*.

### Chapter 2.6

**Aristotelian Intuitions: Non-Supervenience**

Aristotle’s conception of the continuum is the archetype of top-down approach. Michael J. White says the following, regarding Aristotle’s view:

[The existence of points presupposes the existence of magnitudes (or potentially divided ‘submagnitudes’) of which the points are the (potential or actual) limits . . . Points have to be points of things that are not themselves points. (White, [215], 16-17)]

This characterization underscores the top-down picture described above. It also suggests that points are ontologically dependent upon continuous things (referred to here as ‘magnitudes’). Orthodoxy has it the other way around: the continuum is dependent upon the points.

Aristotle’s conception is also the *sole* view that is not punctate. On Aristotle’s view, there are points, but they are not *parts* of the continuum. Indeed, if points were parts on this view, there would be little structural difference between orthodoxy and Aristotle’s view. The requirement on continuity discussed here is that, in
dividing up the continuum, the parts will always be themselves continuous. Stated contrapositively, non-continuous things cannot compose continuous things. Although the top-down aspect of this requirement is essential to it, so too is the notion that dividing up the continuum does not end in just points. Officially, the requirement discussed in this section reads:

2d. Non-Supervenience—An object is continuous only if all of its parts have a proper part.

The title, Non-Supervenience is derived from the idea that the property of continuity does not supervene on the non-continuous.\textsuperscript{76} Aristotle writes:

\ldots nothing that is continuous can be composed of indivisibles; e.g. a line cannot be composed of points, the line being continuous and the point being indivisible. (\textit{Physics}, 231a23-25)

Below, I will examine the motivation behind this view, followed by some concerns.

\subsection{Motivation}

This requirement on continuity emerges first in Aristotle’s picture of continuity. It did not stop with him, though: viewing the continuum as so composed carried on through the middle ages and arguably up through a significant part of the nineteenth century. This feature of the continuum was not deliberately surrendered until the work of Dedekind.\textsuperscript{77} Even in the twentieth century, this conception has been revived in response to Dedekind’s orthodoxy, perhaps most famously by Whitehead and Tarski. Indeed, the very recent literature about atomless gunk primarily draws inspiration from the work of Whitehead.

\textsuperscript{76}Whether this is an appropriate title is contentious. I am simply following White’s lead.

\textsuperscript{77}I will argue below that the intuitions underneath (a consistent) infinitesimal analysis violate Dedekind Continuity. Most of those working at that time were openly using inconsistent tools. Infinitesimals were at some times viewed as zeros (or points) and at other times as non-zero (extended very slightly). The latter of these two incompatible theses about infinitesimals is in the spirit of the Aristotelian Non-Supervenience.
Given its roots in Aristotle, there is a nice characterization of this requirement in White’s book, *The Continuous and the Discrete: Ancient Physical Theories from a Contemporary Perspective*. Before looking at this particular formulation, I need to introduce a distinctive component of Aristotelian continuity, *viz.* that continuity is treated as a dyadic relation (and can be generalized to a multi-sorted one).\(^78\) Aristotle distinguishes two closely related concepts: one translated as ‘continuity’ and the other ‘contiguity’—on the latter, notice the ‘g’ rather than the ‘n.’ It will turn out for our purposes that the difference is irrelevant. Nevertheless, I will draw out the difference first so that I may lay it aside. Aristotle himself was not entirely explicit that the distinction is geometrically irrelevant.

First, Aristotle says that two things are *contiguous* just in case they are touching. By touching, Aristotle has some kind of casual topological relation in mind; their extremities (parts at the boundary) are “together.” For the moment, these notions need not be especially mathematically sophisticated.\(^79\) Given his examples, togetherness or touching can be sufficiently understood by togetherness or touching discernible with the naked eye.

Now, two things are *continuous* just in case they are contiguous and form a unified whole. In some places, he says that the difference between continuity and mere contiguity is that the pairs’ extremities are “one” rather than simply “together.” Consider the following concrete example to illustrate this difference between mere contiguity and continuity. Take a bottle and fill it with half water and half oil. For Aristotle, the oil and water in the bottle count as contiguous. They cannot be

---

\(^78\)This difference is not meant to be especially deep. It is more a matter of presentation. To speak of continuity as being one way or the other is simply a matter of whether we chose to speak of the whole of some things as being continuous or as the things together being continuous. Furthermore, it is worth noting that Aristotle speaks both ways. In the *Categories*, Aristotle says, “A line, on the other hand, is a continuous quantity.” (*Categories*, 4b36-5a1) In any case, in those sections relevant to the present discussion, Aristotle primarily uses the dyadic conception of continuous.

\(^79\)In a later chapter, this concept is discussed in detail. There, mathematics is playing a very central role.
continuous, because they do not form a unified whole. If one were to replace the oil with all water, the pair consisting of the original water and the latter water would qualify as continuous.\textsuperscript{80} This example illustrates that continuity beyond contiguity is really a matter of the contiguous pair being constituted of the same “material”. In physical applications, this distinction is significant, whereas there is no such thing as geometric material. Indeed, it is most likely for this reason that there are no examples of Aristotle using ‘contiguous’ in reference to geometrical examples.\textsuperscript{81}

Finally, we come to White’s statement of this principle; he calls it the non-supervenience of continuity (N-SC):

Each partition of a continuous magnitude into proper parts yields parts each of which is pairwise continuous with at least one other part. (White, [215], 29)

What is the relationship between N-SC and Non-Supervenience? Is a magnitude continuous that is composed of pairwise continuous parts? In order to see how Aristotle gets from N-SC to Non-Supervenience, I will present a pair of arguments that Aristotle gives to that effect.

Call the first argument Aristotle gives, The Mereological Argument for Non-Supervenience (MANS). Aristotle’s argument there depends in large part on the thesis that only extended things have extremities; that is, mereological simples do not have extremities. Suppose for \textit{reductio} that an indivisible point, \( p \), is part of the continuum. According to N-SC, then it would have to be continuous with another part of the continuum, \( q \). From the definition of continuous, if \( p \) and \( q \) are continuous, then an extremity of \( p \) and an extremity of \( q \) are together. That \( p \) has an extremity implies that \( p \) has a proper part. But, points have no proper parts, let alone extremities of...

\textsuperscript{80}I borrowed this example with some modifications from [215].
\textsuperscript{81}White notes, “There is for Aristotle, I suspect, no \textit{topological} distinction between parts that are contiguous and parts that are continuous. Rather continuity pertains to what is homeomerous, while contiguity pertains to parts which are spatially joined but essentially different.” (White, [215], 27)
any kind—contradiction. Hence, points are not parts of the continuum. Therefore, the continuum has only continuous proper parts. Or as Aristotle epigrammatically puts it, “It is plain that everything continuous is divisible [only] into divisibles that are always divisible . . . ” (*Physics*, 231b16)

I will now examine another exemplary argument that follows MANS in the *Physics*. Call this argument, *The Density Argument for Non-Supervenience* (DANS). This argument requires introducing another relation that was left implicit above. Aristotle also uses the concept of *succession*. As expected, some things are in succession just in case they can be placed in some kind of order and nothing of the same kind is between them. In ordinary parlance, two things are in succession if one follows the other. Above, I discussed Aristotle’s notions of contiguity and continuity. Succession is necessary for each, but sufficient for neither. There is a strict ordering of increasing strength between these relations from succession to contiguity to continuity. It follows that continuous things are in succession but not all things in succession are continuous. In DANS, Aristotle shows that points cannot even be in succession, thence not continuous.

For *reductio*, suppose that the continuum is composed of indivisible points. It follows from N-SC that there are at least two points *p* and *q* that are continuous with one another. If *p* and *q* are continuous, then they are in succession. From this, it follows that nothing of the same kind is between them. However, between any two points, there is a line segment. This line segment can be bisected by a point *r*. Therefore, there is something of the same kind between *p* and *q*—contradiction. Therefore, not everything is composed of points. Note that this conclusion is compatible with the continuum being composed of a mix of points and lines. This would be not only unmotivated, but it does not fit the Aristotelian view that points are only

---

*82* Although Aristotle does not spell this out explicitly, one may safely assume that the relation of succession is a linear order: irreflexive, anti-symmetric and transitive and for any *x* and *y*, exactly one of the following holds: *x* < *y*, *x* = *y* or *x* > *y.*
points of lines. For that reason, I take DANS to deliver the desired result: “everything continuous is divisible [only] into divisibles that are always divisible . . .” (ibid.)

2.6.2 Reasons for Concern

The main reason anyone balks at this approach is due to its being a bit mathematically quixotic. The longstanding mathematical success of orthodoxy places a rather heavy burden on alternatives, especially if those alternatives involve costs for practice. White summarizes the essence of this concern:

Although . . . supervenience principles hold considerable intuitive appeal, they are squarely at odds with standard contemporary analysis. (White, [215], 14)

Few people would object that Non-Supervenience fails to capture our intuitions about the continuum; the problem with this requirement is that it is incompatible with an extremely fruitful mathematical paradigm.

To go a bit deeper, the problem is at least partly related to a much wider outlook within modern mathematics. As I mentioned above, the contemporary mathematical worldview is a necessarily bottom-up one: everything in the mathematical universe is built up from the most primitive parts. Again, contravening the dominant mathematical foundation counts as a cost.

Before moving on, I want to comment on set-theoretic foundations. Set theory is an undeniably fruitful way of thinking about the mathematical universe and models therein, but it is hardly required. It is a very recent development in the history and philosophy of mathematics. Mathematicians modeled continuous phenomena for centuries without the essentially atomistic framework of Zermelo-Fraenkel set theory. As we will see below, models are available that satisfy Non-Supervenience. An even more recent development within mathematics, Category Theory, has a special home for the Aristotelian mindset. Even still, the set-theoretic approach to analysis is the
most familiar to the typical mathematician. As a result, pragmatic pressure on the Aristotelian view remains even when moving away from the universe of sets.

2.6.3 Fully Formalized

Over the last century, several formal presentations of Aristotelian continuity have been developed, most notably Tarski’s formalization of Whitehead’s ideas in “Foundations of the Geometry of Solids,” [198]. In spite of this article’s value, many others have followed in Tarski’s wake, improving upon some of the details. In particular, Tarski’s original article is admirably simple but at the cost of conceptual transparency. I favor a version developed by Peter Roeper in “Region-Based Topology.” Although many other successful point-free programs exist, Roeper’s begins with Boolean algebras and it therefore generalizes naturally across dimensions. Below, I will present his region-based theory of continua with only minor notational and presentational differences.

Basic Presentation

First, I must make explicit in what a Boolean algebra consists. In what follows, I will use $\alpha \leq \beta$ to mean $\alpha$ is a part of $\beta$, $\alpha \lor \beta$ and $\alpha \land \beta$ to represent the sum and product of $\alpha$ and $\beta$ respectively. Finally, $-\alpha$ is the complement of $\alpha$.\(^{83}\) A Boolean algebra has a universal and null element, written 1 and 0 respectively. The following restrictions are placed on sums and products.

**Associativity** $\alpha \lor (\beta \land \gamma) = (\alpha \lor \beta) \land \gamma$ and $\alpha \land (\beta \land \gamma) = (\alpha \land \beta) \land \gamma$

**Commutativity** $\alpha \lor \beta = \beta \lor \alpha$ and $\alpha \land \beta = \beta \land \alpha$

**Distributivity** $\alpha \lor (\beta \land \gamma) = (\alpha \lor \beta) \land (\alpha \lor \gamma)$ and $\alpha \land (\beta \lor \gamma) = (\alpha \land \beta) \lor (\alpha \land \gamma)$

**Top and Bottom** $\alpha \lor 0 = \alpha$ and $\alpha \land 1 = \alpha$

\(^{83}\)Sum, product and complement can be defined in the usual way using just $\leq$. In particular, $\alpha \lor \beta$ is the unique object $\chi$ such that $\gamma \leq \alpha$ or $\gamma \leq \beta$ just in case $\gamma \leq \chi$; whereas, $\alpha \land \beta$ is the unique object $\xi$ such that $\gamma \leq \alpha$ and $\gamma \leq \beta$ just in case $\gamma \leq \xi$. Finally, $-\alpha$ is the unique object $\zeta$ such that all and only $\gamma \not\leq \alpha$, $\gamma \leq \zeta$.  

63
Complements $\alpha \lor -\alpha = 1$ and $\alpha \land -\alpha = 0$.

When it comes to describing a topological structure, Roeper introduces two fundamental notions: limited and connected. These notions are meant to be taken in their intuitive sense. For a quick gloss, to be limited is to be bounded and two regions are connected just in case they are “together” in Aristotle’s sense above. For notation, I will write, $\mathcal{L}(\alpha)$ to mean the region $\alpha$ is limited and $\alpha \Join \beta$ for the region $\alpha$ is connected to region $\beta$. Lastly, following Roeper, I use $\alpha \ll \beta$ to mean the region $\alpha$ is an interior part of region $\beta$. That is, $\alpha \ll \beta$ is defined so that it is not the case that $\alpha \Join -\beta$. I use Roeper’s numbering system:

$\textbf{A1}$ If $\alpha \Join \beta$, then $\beta \Join \alpha$.

$\textbf{A2}$ If $\alpha \not= 0$, $\alpha \Join \alpha$.

$\textbf{A3}$ For no $\alpha$, $0 \Join \alpha$.

$\textbf{A4}$ If $\alpha \Join \beta$ and $\beta \leq \gamma$, $\alpha \Join \gamma$.

$\textbf{A5}$ If $\alpha \Join (\beta \lor \gamma)$, then $\alpha \Join \beta$ or $\alpha \Join \gamma$.

$\textbf{A6}$ $\mathcal{L}(0)$

$\textbf{A7}$ If $\mathcal{L}(\alpha)$ and $\beta \leq \alpha$, then $\mathcal{L}(\beta)$.

$\textbf{A8}$ If $\mathcal{L}(\alpha)$ and $\mathcal{L}(\beta)$, then $\mathcal{L}(\alpha \lor \beta)$.

$\textbf{A9}$ If $\alpha \Join \beta$, then there is some $\beta' \leq \beta$ with $\mathcal{L}(\beta')$ and $\alpha \Join \beta'$.

$\textbf{A10}$ If $\mathcal{L}(\alpha)$, $\beta \not= 0$ and $\alpha \ll \beta$, then there is some non-null $\mathcal{L}(\gamma)$ such that $\alpha \ll \gamma \ll \beta$.

Many of the above are natural requirements for connectivity ($\Join$) and limitation ($\mathcal{L}$). $\textbf{A1-5}$ deal exclusively with connectivity. In particular, these axioms assert that connectivity is symmetric ($\textbf{A1}$), all and only non-null regions are self-connected ($\textbf{A2}$, $\textbf{A3}$), the wholes of which two connected regions are parts are connected ($\textbf{A4}$) and if a region is connected to a sum of regions, then it is connected at least one of those regions ($\textbf{A5}$). The axioms $\textbf{A6-A8}$ are exclusively on the concept of limitation. The null element is limited ($\textbf{A6}$) and limitation is closed under parthood and products ($\textbf{A7-A8}$). The last two principles combine them. $\textbf{A9}$ is similar to (but not dual to) $\textbf{A4}$: in the same way that connection can be extended, connection can also be
pared down to a limited region. Finally, A10 is comparable to the Infinite Divisibility requirement discussed at length in §2.1.

The following additional axioms are much more specific for the purpose of forming a continuous system. A1-A10 is a fairly general system meant to mimic features of point-set topology (itself a fairly general framework) only with regions instead of points. To add the following principles sharpens our concepts in such a way as to capture continuity in this system. There is one last definition needed for the following axioms. Casually, a region $\alpha$ is convex just in case if any one were traveling in $\alpha$, no boundaries or gaps would have to be crossed to go from any one place in it to another in it. Officially, $\alpha$ is convex just in case for any $\beta$ and $\gamma$ with $\beta \lor \gamma = \alpha$, there is some $\alpha' \ll \alpha$ such that $(\alpha' \land \beta) \bowtie (\alpha' \land \gamma)$.

- **B2** There is some $\alpha$ such that $\alpha \notin \{0, 1\}$.
- **B3** For any non-null $\alpha, \beta$ where $\alpha \lor \beta = 1$, then $\alpha \bowtie \beta$.
- **B4** If $\mathcal{L}(\beta)$ and it is not the case that $\alpha \bowtie \beta$, then there is some finite sum $\gamma = \gamma_1 \lor \cdots \lor \gamma_n$ for convex $\gamma_i$ such that $\beta \leq \gamma \ll -\alpha$.
- **B6** There is a countable set $\Gamma$ of regions such that for some regions $\alpha$ and $\beta$ with $\mathcal{L}(\alpha)$ and $\alpha \ll \beta$, there is some region $\gamma \in \Gamma$ with $\alpha \ll \gamma \leq \beta$.

**B2** simply ensures that the algebra is non-degenerate; that is, it is more than just the universal and null elements. **B3** ensures that the entire space lacks gaps: the members of any binary partition of the space are connected. **B4** and **B6** are slightly less obvious and more technical in nature. They are better described by understanding their consequences. **B4** ensures that for any sum of a sequence of regions $\alpha$, if $\alpha$ “converges” towards some region $\beta$, then they are connected ($\alpha \bowtie \beta$). For an illustration, see Figure 2.4.

---

84Below, B5 is missing because it can be proved from B6 and the other principles. I left it out to simplify exposition as much as possible.
Figure 2.4. Connection ($\alpha \sqsupseteq \beta$) Due to Convergence.

It is stated somewhat negatively here; *viz.* that disconnected regions can be fit within a finite number of convex regions. Lastly, $\mathbf{B6}$ is meant to capture what is ordinarily called *second-countability*. Second-countability guarantees that the regions “lump” together in nice “even chunks.” To flesh this out, the real numbers with the Euclidean metric satisfy this feature; on the other hand, the real numbers with the discrete topology (where every singleton counts as an open region) do not satisfy this. In the latter case, the continuum splinters into every single point and does not modestly “lump together.” To make a rather crude analogy, individual water molecules cohere. Although the water *could be* separated into every individual water molecule, water naturally forms into distinct puddles and pools. The discrete topology is like the individual molecules by themselves and the topologies satisfying second-countability are like the coherent pools.

**Satisfying the Requirement**

How does this formal system square with the *Non-Supervenience* requirement and other Aristotelian intuitions? Before going into any of the details, allow me to explain the reasons I favor Roeper’s presentation. First, Boolean algebras are structures that do not necessarily begin with points in any way. Some Boolean algebras have points, but they need not. Secondly, Roeper defines points to be equivalence classes of ultrafilters (recall the definition from §2.5.3) on the Boolean algebra. Specifically,
these ultrafilters—thence points—are defined \textit{in terms of} the regions. To have the points be defined in terms of the regions gets the Aristotelian order of fundamentality correct. Along with that, ultrafilters are ready-made mathematical entities that are being put to valuable mathematico-philosophical work, rather than having to develop a completely new apparatus for representing points over regions.

When it comes to satisfying the requirement of \textit{Non-Supervenience}, it is worth noting that Roeper proves in his Theorem 6.3 that the above principles \textbf{A1-A10}, \textbf{B2-B4}, \textbf{B6} are sufficient for the following principle:

\textbf{B1} If region $\alpha \neq 0$ then there is a non-null $\beta \neq \alpha$ with $\beta \leq \alpha$.

\textbf{B1} expresses that the world is gunky, or in Aristotle’s words from above, all regions are “divisible [only] into divisibles that are always divisible . . .” (Aristotle, \textit{op. cit.}) Using the terminology from \textit{Non-Supervenience}, everything has a proper part—no points. This is enough to satisfy Aristotle, and as a matter of fact, Roeper’s system is even more topologically sensitive than Aristotle would require. This is not a problem for Roeper. I just want to register that principles like second-countability, \textbf{B6}, place much more subtle expectations on the continuum than Aristotle had ever envisioned.

\textbf{Chapter 2.7}

\textbf{Conclusion}

Throughout I have examined four major requirements one might place on continuity and the continuum. When expressed in contemporary mathematics, each of them yields a radically different image of the nature of continuity and the continuum. Nevertheless, each of them stands up to mathematical scrutiny in spite of their philosophically disparate motivations and mathematically tractable differences. Such a wide variety of incarnations suggest that the best posture towards the concept of
continuity is one of tolerance.
Chapter 3

Zeno’s Arrow

The more the difficulty is meditated,  
the more real it becomes.  
—Bertrand Russell, on Zeno’s Arrow

Of [Zeno’s paradoxes], I think perhaps the most  
profound is that of the arrow.  
—Graham Priest

Zeno of Elea’s paradoxes of motion have received attention both from Aristotle and from contemporary authors who use modern mathematical techniques. One of them is the Paradox of the Arrow, put concisely here by William James:

If a flying arrow occupies at each point of time a determinate point of space, its motion becomes nothing but a sum of rests, for it exists not out of any point; and in the point, it doesn’t move. (James, [98], 157; emphasis original)

Effectively, in any given moment, if Zeno’s flying arrow is frozen in place, how could it be flying? Later I will elaborate these premises more, but the full argument from Jonathan Lear goes like this:

(1) Anything that is occupying a space just its own size is at rest.

(2) A moving arrow, while it is moving, is moving in the present.

(3) But in the present the arrow is occupying a space just its own size.

(4) Therefore, in the present the arrow is at rest.

1See Russell, “The Problem of Infinity Considered Historically,” [175], p. 51. Pagination for [175] throughout comes from Wesley Salmon’s Zeno’s Paradoxes, [176].

2See Priest’s “Inconsistencies in Motion,” [154], p. 340.
(5) Therefore a moving arrow, while it is moving, is at rest. (Lear, [114], 91)

Michael J. White provides a solution to Zeno’s Arrow using Abraham Robinson’s Nonstandard Analysis in his “Zeno’s Arrow, Divisible Infinitesimals, and Chrysippus.” White denies premise (3) by claiming that the present has an infinitesimal length—lengths as small as the infinite is large. It follows that in the present the arrow occupies a space larger than its own size, though only by an infinitesimal amount. Unfortunately, as the reader will find below, White’s solution suffers from a measure of arbitrariness. I will follow the same basic strategy of using infinitesimals to deny premise (3). To avoid the arbitrariness, I will use a special kind of infinitesimal called a nilpotent, unlike White. A nilpotent $\varepsilon$ is any value where $\varepsilon^k = 0$, for positive integer $k$. Over the last fifty years, two distinct non-equivalent extensions of the real numbers have been developed that contain nilpotents. One of them, developed by Paolo Giordano, uses classical logic, but his extension fails to be a full field; otherwise, all nilpotents are identical to zero. The other one is known as Smooth Infinitesimal Analysis (sia, henceforth). This latter system uses—indeed requires—intuitionistic logic: if excluded middle is assumed, absurdity follows. Issues like these will demonstrate that the arbitrariness can be avoided, but at no small theoretical cost.

I will begin the chapter by providing a version of the Paradox of the Arrow. In §3.2, I will consider previous attempts to solve the Arrow that involve denying premises other than (3)—the one I deny. In §§3.4-5, I will present sia and Giordano’s nilpotents, exploring the application of each to the Arrow. Ultimately, I think a genuine advantage is had by using nilpotents. I leave it to the reader to choose between Giordano and sia.

---

3Below I will use the term ‘analysis’ numerous times: Nonstandard Analysis, Smooth Infinitesimal Analysis and there is everyday Real Analysis too. Analysis is a branch of mathematics that has at its historical core Newton’s and Leibniz’s calculus. Notions like limits, derivatives and integrals are at the center of this discipline.

4The smooth reals of SIA only extend $\mathbb{R}$ in the sense that for any $c \in \mathbb{R}$, there is some corresponding $c$ in the smooth reals. It is definitely not a model-theoretic extension.
Chapter 3.1

Zeno’s Arrow

Here I will try to elaborate and motivate as much as possible the premises of Zeno’s argument. As a simple book-keeping matter, I will avoid the locutions of ‘at a time’ or ‘in a time’ unless directly quoting the argument’s premises. This is a distinction that has appeared in the literature that relies mostly on how those prepositions are used. I will attempt to avoid this worry entirely by using the general expression, ‘during a time’ as a way to be deliberately amiguous.

Allow me to begin with (1): *anything that is occupying a space just its own size is at rest.* This premise is not exactly obvious. One must be careful when analyzing this premise in particular not to bring one’s sense from modern mechanics to bear too quickly. Indeed, people who have never lived a day without high speed photography are used to seeing something in motion (a galloping horse, for example) at least appearing to occupy a space just its size. Considering the contrapositive, one can begin to imagine the idea: anything not at rest does not occupy a space its own size. An object in motion “blurs” its way across the space through which it moves: a hummingbird’s wings occupy two small arcs as it flaps them; a squirrel occupies a helical space as it scampers up the tree. More directly, this premise gains its plausibility when considering a given body over some time.\(^5\) Perhaps instead (1) should be:

\[(1') \text{ Anything that is occupying some same space just its own size for an extended period of time is at rest during that time.}\]

The reason one might prefer (1’) is that premise (1) is hard to believe without inter-

\(^5\)It should be clear that the contemporary metaphysical notion of exact occupation is not being used here. Intuitively, $M$ *exactly occupies* region $R$ when $R$ hugs the very contours of $M$: (i) every subregion is filled and (ii) no part of $M$ is outside of $R$. If one were to interpret occupation in this way, then everything would always be at rest.
preting it to mean \((1')\).

Premise (2) is much easier to grasp: *a moving arrow, while it is moving, is moving in the present.* To one who readily accepts the reality of motion, a natural thought is that if something is moving steadily during a period of time, it is moving during any part of that period of time; in particular, the present is such a part. There is good reason to think that this is how Aristotle saw this premise. Another way of seeing this premise is arguably closer to Zeno’s view: if the arrow is moving *at all*, then it is moving during the present.\(^6\) On that view, the present has special temporal priority; indeed, on that view, the present is all that exists—much like the contemporary thesis of presentism. Typical presentists hold two theses. First, all things can be placed into an A-series; that is, all things are partitioned into the categories of past, present and future. Second, presentists believe that past things no longer exist and that future things do not yet exist—only present things exist.\(^7\) From the presentist point of view, Aristotle’s talk of the present as just another part of time would not be well received. The presentist interpretation of premise (2) will become very important in §3.2.2 and is treated in greater detail there.

Lastly, premise (3): *in the present the arrow is occupying a space just its own size.* Recall that this is the premise that White and I challenge. (3) seems obvious from a contemporary perspective, especially when one considers the present to be a point in time. Of course, this will not do. To interpret (3) this way will rule out the motivation provided for (1) above; in other words, if the arrow is occupying a space just its own size during the present and it is moving, then (1) is false. Of course, if I take (1) to mean \((1')\), then much more work will need to be done to establish a contradiction.\(^8\)

---

\(^6\)This is discussed more below in §3.2.1.

\(^7\)For further discussion of presentism, see Dean Zimmerman’s “The Priveleged Present,” [228].

\(^8\)White suggests that this argument (from Lear’s reading of Aristotle’s presentation of Zeno) equivocates on ‘occupying a space just its own size.’ On his reading, for premise (1) to make sense at all, it must be interpreted with the temporal qualification added in \((1')\); whereas, in premise (3), ‘occupying a space just its own size’ must be read unqualified. I am not going to pursue the issue of whether Zeno has equivocated. I am simply going to assume that this argument has the form of
Since in the final estimation, I am not trying to vindicate this premise, I will simply return to the high speed photograph of a galloping horse to see why one might believe this. The present is captured by the photograph; upon casual inspection, the horse carves out a crisp and clear horse-shape in the air. Later, I will argue along with White that if the present moment is extended, then a moving object will not occupy a space just its own size during the present.

Chapter 3.2

Denying other Premises

To understand some of the history of this paradox, I will examine previous attempts to refute it by denying either premise (1) or (2). Although I deny premise (3), I will borrow some insights from each of the following perspectives, viz. an answer that uses the tools of the calculus as well as Aristotle’s answer.

3.2.1 The Calculus and Premise (1)

A natural thought for the reader would be, “The calculus contains the concept of instantaneous velocity. Doesn’t that contradict Zeno’s premise (1)? I agree that the arrow is occupying a space just its own size during the present, but the calculus gives a perfectly good story about how it could nevertheless be moving during the present.” After briefly analyzing the limit-based motion during a point-instant, I will argue that it begs the question against Zeno.9

The core idea is to begin with the notion of average velocity and allow the amount of time to shrink indefinitely. For simplicity, consider the path of an object in one the BARBARA syllogism as it appears. I will deny that it is sound by arguing that one can plausibly deny premise (3).

9Since I think Zeno is wrong, not everything he believes can be accommodated. Nevertheless, I side with Lear that Zeno should be given a running start—pun intended!
dimension modeled by a function of time, \( f(t) \). Let \( t_1 < t_2 \) be any two points in time with \( \Delta t = t_2 - t_1 \). The average velocity between \( t_1 \) and \( t_2 \) is,

\[
\frac{f(t_2) - f(t_1)}{t_2 - t_1} = \frac{f(t_1 + \Delta t) - f(t_1)}{\Delta t} .
\]

Now, fix \( t_1 \) and let \( t_2 \) vary. Allowing \( t_2 \) to close in on \( t_1 \), one can find better and better approximations of the instantaneous velocity of the object modeled by \( f \) at time \( t_1 \). Equivalently, as \( t_2 \) approaches \( t_1 \), \( \Delta t \) shrinks towards zero and average velocity closes in on instantaneous velocity. Instantaneous velocity at \( t_1 \) is also referred to as the \textit{derivative} of \( f \) at \( t_1 \); written, \( f'(t_1) \). Explicitly, the \textit{derivative} at \( t_1 \) of \( f \) is:

\[
\lim_{\Delta t \to 0} \frac{f(t_1 + \Delta t) - f(t_1)}{\Delta t} .
\]

With this background, the thought is, “If a time-tested mathematical theory can lend credence to the concept of motion during an instant, then one may dismiss Zeno and his Arrow on the grounds that premise (1) is false: during the present, the arrow occupies a space just its own size and yet it is moving.”

\textbf{Looking Elsewhere}

The calculus provides a fine response to Zeno, provided that one is not concerned to engage the historical Zeno. If one wants an answer to the paradox that is consonant with basic contemporary mechanics, look no further. Nevertheless, I agree with Lear that on a certain historical understanding of Zeno, this argument begs the question. Lear even says, “The calculus is impotent to solve Zeno’s arrow.” (Lear [114], 100) Lear’s startling remark is intended to shock those who have taken Zeno’s defeat for granted.

\begin{footnote}{The official “epsilon-delta” definition of a limit is that a function \( f \) has limit \( L \) as \( x \) approaches \( c \) iff, for any \( \epsilon > 0 \), there is some \( \delta > 0 \) such that for any \( x \), if \( |x - c| < \delta \), then \( |f(x) - L| < \epsilon \).}
\end{footnote}
How does using the limit-based calculus beg the question against Zeno? Lear suggests that Zeno is a committed student of Parmenides, the latter of whom fits the profile of a contemporary presentist.\textsuperscript{11} To review, the essence of this paradox is, \textit{If the arrow is not moving in the present, how could it move in any larger part of time?} Advocates of the limit-based approach take motion during intervals of time for granted and attempt to provide an explanation of motion-during-the-present. This answer reverses the order of priority, switching motion-during-the-present from being an \textit{explanans} to an \textit{explanandum}. Supposing that the present is a point-instant, to provide an answer that starts with intervals and then tells a story about how points are dependent on the intervals is to beg the question against Zeno. If one wants to move Zeno at all, one’s conception of motion must give pride of place to the present.\textsuperscript{12}

3.2.2 Aristotle and Premise (2)

It is primarily in Aristotle’s text that one finds Zeno’s arguments. Aristotle’s response to Zeno’s arrow is therefore a natural place to look for hints in responding to Zeno. I will rely heavily on the insights provided by Jonathan Lear. According to Lear, one sees Aristotle denying Zeno two needed assumptions both involving the nature of instants in time.

\textsuperscript{11}This idea is found in Lear [114], p. 104, n. 19. Specifically, he cites the passage:

\begin{quote}
No was nor will: all past and future null;  
Since being subsists in one ubiquitous  
Now—unitary and continuous. (Henn, [90], 26)
\end{quote}

I recognize that there are issues of interpretation here for the historical Zeno. Even if Zeno does not think this, the question remains interesting, given that presentism has many living advocates.

\textsuperscript{12}One might wonder, how would an ersatz B-theorist respond to this? An ersatz B-theorist is a presentist who uses abstract objects when quantifying over past or future times. For the ersatz B-theorist, the real numbers in the mathematical theory of the calculus function as models for both concrete and abstract objects. Every member of $\mathbb{R}$ with one exception represents an abstract ersatz point in time. Indeed, motion-during-the-present (the only concrete part of the theory) is \textit{defined} completely in terms of the “motion” over non-concrete intervals of time surrounding the present. Given the order of construction, one would predict that the concrete objects would be the intervals of time and the point-instants would be the abstractions. The calculus has it completely backwards. The mathematical theories that I consider below do not have this peculiar feature. Thanks to Matthew Davidson and Tony Roy for pressing me here.


Instants and Composition

The primary point of contention between Aristotle and Zeno involves the composition of time: is it a collection of instants and if so, then what follows? For Aristotle, this is a familiar theme: extended things (whether extended through time or extended in space) are not composed of non-extended things; that is, intervals or stretches of some continuous phenomenon are not composed of points.\(^{13}\) The points are there only as potential places for division. Below, I will discuss the consequences of this for time specifically.

Allow me to elaborate a little on how and why denying these composition theses and related theses deliver the desired result for Aristotle. First, according to Aristotle, time is not composed of point-instants. More specifically, every duration of time “contains” instants (in virtue of instants being potential sites of division of the duration), but instants do not compose the duration. Generally for Aristotle, points are nothing more than the edge or limit of some extended thing; in the case of time, the present instant is nothing more than “an extremity of the past . . . and again of the future.” (Physics VI.3, 233b34-234a2)\(^{14}\) Above I suggested that Aristotle interprets premise (2) as follows: if something moves steadily through a period of time, it is moving through any part of that period of time. Point-instants, including the present, are not parts of time. If this is indeed the proper interpretation, then the something other than the claim that the present is a part of time is needed to establish (2). Although this is not a direct denial of (2),\(^{15}\) Aristotle sees himself as pulling the rug out from under Zeno.\(^{16}\)

A common charge against Aristotle here is that composition is irrelevant to this

---

\(^{13}\)See §2.6 above for further discussion of Aristotle.

\(^{14}\)All Aristotle citations are from Barnes’ Complete Works of Aristotle, [5], with Stephanus pagination.

\(^{15}\)To suggest otherwise would be to commit the fallacy of denying the antecedent.

\(^{16}\)I have already suggested that Aristotle sees premise (2) unlike Zeno. This will be discussed more below. In the mean time, the reader hopefully can see the strategy that Aristotle takes.
argument. Against Aristotle, people complain that he is committing the familiar fallacy of composition when he makes the following remark:

[T]he flying arrow is at rest, which result follows from the assumption that time is composed of moments: if this assumption is not granted, the conclusion will not follow. (Physics VI.9, 239b30-33)

I do not think Aristotle is committing any fallacy. Not all examples of division or composition are fallacious. For example, every stone in the pyramid at Giza is larger than my head; therefore, the pyramid at Giza is larger than my head. Similarly, for all ordinary extended regions of time $T$, if an object is in motion for the parts (whole) of $T$, then it is in motion for the whole (parts) of $T$. If Aristotle is obviously mistaken, it is perhaps only because one has already assumed the truth of the calculus, where the idea of motion-during-the-present is considered conceptually stable. Since all other cases of motion bear witness to these division and composition principles, Aristotle hardly commits a fallacy by thinking the analogue would hold for such difficult cases.

**Instants and Mechanical Categories**

Aristotle’s overall views involve more than simply denying that instants compose extended periods of time. Lear draws his reader’s attention to remarks Aristotle makes (in Physics VI.3) that go beyond but are informed by his background views of the continuum. Aristotle suggests that the concepts at rest and in motion do not apply at all to point-instants. For the concepts of motion or rest to apply at all, there must be an extended period of time over which an object is either resting or moving. I suggested above that Aristotle uses an argument from the composition of the continuum to undermine premise (2). Now, since Aristotle assumes that the

---

17To be slightly more perspicuous, $T$ is self-connected and the motion is constant through $T$. No one would make the mistake of thinking that if something is moving during a period of time, then it would necessarily be moving through all of it. Generally, when I drive from Columbus to Cleveland, I could say of me that I moved from when I left until I arrive at my destination, though I definitely stop at red lights and the like.
present is an instant and that neither motion nor rest can occur during an instant, he outright denies premise (2): *A moving arrow, while it is moving, is moving in the present.*

First, against motion during an instant, Aristotle begins with the thesis that individual instants are points; i.e., unextended and indivisible. Suppose for contradiction that motion is coherent during an instant. Let objects $A$ and $B$ be moving at speeds $a < b$ during the instant and traversing distances $\alpha < \beta$ respectively during the instant. Since $\alpha < \beta$, $B$ will have travelled $\alpha$ distance before having gone distance $\beta$.\(^{18}\) Necessarily, since $B$ travels $\beta$ in an instant, it will have travelled $\alpha$ in a time strictly shorter than the instant, thus dividing the instant—contradiction.

Second, against rest in an instant, Aristotle provides a pair of arguments one of which is more difficult, if only from being extremely compact in presentation. The more straightforward argument begins from the point of view that since resting is the conceptual complement of motion, all and only candidates of motion are candidates of rest. Another way of putting this is that if some thing’s nature prevents one from ascribing to it various dynamic properties, then the same aspects of its nature would prevent him from ascribing the parallel static properties. Since it is a category mistake to say that something is moving in an instant, it is just as much a category mistake to say something is resting in an instant. Just because the number two is not the color blue that does not entitle one to think it is some other color. So it is with motion and rest.

The more difficult argument begins with the thesis mentioned above that an instant is the extremum of an extended period of time; indeed, it is the extremum between two extended periods of time. This argument is meant to show more that neither motion nor rest can occur in an instant. Let some object $A$ be moving for $\alpha < \beta$ follows from the fact that $A$ is moving faster than $B$ for a fixed time. In other words, multiply speeds $a < b$ by a fixed positive time to yield lengths $\alpha < \beta$.\(^{18}\)
some time, slow down to a stop and rest for a period. Between A’s moving and its resting, there is an instant which is the extremum of both periods of time. To exhaust all possibilities, consider all permutations: in the instant between these two periods, A be at rest, in motion, both or neither. That A is both in motion and at rest is an option I will not consider here.\(^{19}\) Now, the only reason (according to Aristotle) that one might give to suggest that A is at rest during the instant—\textit{viz.} that it is the extremum of the period of rest—equally applies to hypothesis that A is in motion during the instant. Hence, one must conclude that in the instant, A is neither in motion nor at rest.

**Looking Elsewhere**

Given Aristotle’s overall worldview, he has been effective in disarming Zeno. However, there are reasons for seeking other responses. The first reason is comparable to the last one: Zeno does not interpret premise (2) the way Aristotle does.\(^{20}\) In that way, a Zeno sympathizer would not be satisfied. He would not be drawn in by Aristotle’s arguments from composition. Second, and relatedly, Zeno need not believe that the present is a point-instant. For Aristotle’s arguments against motion and rest during a point-instant to link up with the paradox, the present must be a point-instant. In fact, one can grant Aristotle the presupposition that motion during an instant is incoherent and yet still provide a stable notion of motion-during-the-present. This requires a picture of time where the present is extended. It is to this exact idea that I now turn.

\(^{19}\)To explore this possibility, see Graham Priest’s \textit{In Contradiction}, [156].
\(^{20}\)This was discussed in §§3.1 and 3.2.1.
Chapter 3.3

White on Zeno

To answer Zeno, White uses nonstandard analysis and its accompanying number system, the hyperreals, rather than the limit-based real analysis. Instead of trying to make sense out of motion during an unextended point-instant, White attempts to understand the present moment as an extended interval of time. This approach enables him to sidestep the problems described above: motion during intervals of time is all he needs if the present is not a point-instant. Let the reader note that on this view, points are still around. However, points in time play a purely formal role and are not meant to represent anything about time. Points are like markers and no more. One might be concerned that there is no gain over the limit-based perspective, since points are still a part of the theoretical framework. To view points this way is not theoretically stable in combination with the limit-based answer. On the limit-based picture, points are a natural choice for representing the present. Instead, White introduces infinitesimal intervals to play the part of the present. I will begin by providing a brief description of the hyperreals and how White puts them to use. I will then proceed to discuss some difficulties that emerge from using Robinson’s hyperreals.

3.3.1 Robinson’s Hyperreals

The hyperreals extend the real numbers (\( \mathbb{R} \)); i.e. there is a copy of the real numbers inside the set of hyperreals. However, there are infinitely large numbers and infinitely small numbers (infinitesimals) among the hyperreals as well. The original everyday real numbers are referred to as the standard reals. An infinite number \( \Omega \) is one where for every (standard) real number \( x \), \( \Omega > x \). Such infinite numbers exist in an “outer
orbit” beyond all the standard real numbers. On the other hand, a nonzero infinitesimal \( \kappa \) is one where for every nonzero (standard) real number \( x \), \( |\kappa| < |x| \). In fact, just as for 2, 3, and 4, there are reciprocal values \( \frac{1}{2}, \frac{1}{3}, \frac{1}{4} \), each infinitesimal is identical to \( \frac{1}{\Omega} \), for some infinite number \( \Omega \). In the same way that the infinite numbers are out beyond “the stratosphere” of the ordinary reals, the infinitesimals are immeasurably small but still nonzero. To use an illustration, consider a microscope that can magnify by an infinite factor. Only with a lens of such strength can one view the infinitesimals. The infinitesimals would be invisible to a microscope that magnified by only some finite factor, no matter how large the factor! See Figure 3.1.

![Figure 3.1. Hyperreals magnified by an infinite factor](image)

Finally, as a matter of bookkeeping, I will need to make use of a crucial relation among hyperreals \( x \) and \( y \): being *infinitely close*, written \( x \simeq y \). Values \( x \) and \( y \) are *infinitely close* just in case there is some infinitesimal \( \varepsilon \) such that \( x - y = \varepsilon \). The concept of the *monad* of a point \( x \) is another property needed below. It is
defined as follows: for any point $x$, the monad of $x$ is defined to be the class of points infinitesimally close to $x$.

### 3.3.2 White against Zeno

How does White put the hyperreals to use? Recall the definition of average velocity from §2.1 (here I suppress all subscripts):

$$\frac{f(t + \Delta t) - f(t)}{\Delta t}$$

Previously, I presented motion during the present as the derivative of some function (here $f$) which determines the arrow’s location as a function of time $t$. There, I specifically used the limit concept of a derivative, a concept at odds with Zeno’s commitment to presentism. White dodges this problem by selecting an infinitesimally extended interval for the present. Since infinitesimals are nonzero, White can provide a simple formulation of velocity during the present that does not already depend on the velocity of some other size as in the case of the limit-based presentation.

In order to make this work, White’s solution presupposes that space and time have the structure of the hyperreals.\textsuperscript{21} There is a deeper question about whether space or time actually has this structure. I cannot address such a large question here. I take Zeno to be trying to show that motion is impossible given some modest assumptions about time and motion. If so, it suffices to show that there is a coherent story to be told about motion when accepting those assumptions. White does not explicitly address this question himself. Nevertheless, given that his stated concern is in investigating the conceptual viability of using nonstandard analysis to solve the paradox, I suspect that White adopts a similar posture towards solving Zeno’s

\textsuperscript{21}Aristotle suggests in the Physics (232b25) that space and time must have the same structure. The essence of his claim is that variation in the measure of motion can only be captured properly if time and space match in structure. I am simply taking this on as an assumption.
As I mentioned above, with extended values available for the length of the present, one can directly define velocity during the present. This technical concept of velocity-during-the-present plays the part of ‘motion in the present’ just as instantaneous velocity does on the limit-based picture. Velocity-during-the-present is simply the distance covered in space during the present divided by the length of the present. Unlike instantaneous velocity discussed above, the structure of average velocity and velocity-during-the-present are identical; there is no need to resort to viewing things in the limit. As it turns out, White’s velocity-during-the-present is a special case of how Robinson defines the derivative. The derivative definition in nonstandard analysis requires not just one infinitesimal length (as in the case of the present) but all infinitesimals. The issue of how Robinson’s derivative compares with White’s velocity-during-the-present will be revisited below.

I will now show how White applies these concepts directly to the paradox. Here is the argument again:

1. Anything that is occupying a space just its own size is at rest.
2. A moving arrow, while it is moving, is moving in the present.
3. But in the present the arrow is occupying a space just its own size.
4. Therefore, in the present the arrow is at rest.
5. Therefore a moving arrow, while it is moving, is at rest. (Lear, [114], 91)

---

22One might worry that a proper answer to Zeno’s Arrow should only use contemporary scientific theories. A scientific theory’s success depends primarily on its internal consistency and how well it models some empirical phenomenon. The phenomenon here is, of course, motion. My Zeno is placing a priori constraints on the kinds of answers possible that would be ignored by a practicing scientist. White is (and I am) trying to supply an answer sensitive to Zeno’s a priori constraints.

23More specifically, one says that $f$ has a derivative at $t$ just in case, there is a standard $m \in \mathbb{R}$,

$$m \simeq f(t + \varepsilon) - f(t)$$

for all nonzero $\varepsilon \simeq 0$.

Even in the nonstandard setting, the value $m$ is often written as $f'(t)$. This is discussed more below in §3.4.2.
To simplify things throughout, consider the material unextended point at the very tip of the arrow and assume that it is located at some particular spatial point $z_0$ at some time $t_0$. White begins by granting Zeno his first premise, interpreting it as the aforementioned,

\[(1')\] Anything that is occupying some same space just its own size for an extended period of time is at rest during that time.

Hence, in the picture White develops, whenever the arrow is occupying a space exactly its own size for any extended period of time—even an infinitesimal one—it follows quite plausibly that the arrow is at rest.\(^{24}\)

What about premise (3)? Following (2), if one maintains that a moving arrow moves during the present, then there is some nonzero spatial segment that the arrow passes through during the present. Allow me to spell this out explicitly. Let $\varepsilon$ be a nonzero infinitesimal so that the interval $[t_0, t_0 + \varepsilon]$ counts as the present. Also, let the function $z(t)$ model the motion of the point vertically relative to the earth. Assuming that $t_0$ is some time after the arrow has been shot but before it has landed, let $z(t_0) = z_0$. Define $T_0$ to be the set of spatial points the arrow moves through during the present. It follows that $T_0$ is the interval $[z_0, z(t_0 + \varepsilon)]$. Note that $T_0$ has nonzero length $z(t_0 + \varepsilon) - z_0$, itself infinitesimal.\(^{25}\) It follows that (3) must be false, since the very tip of arrow is not occupying a point just its size but rather an extended interval in the present. If premise (3) were correct then $T_0$ would contain only the single point $z_0$. As it stands, $T_0$ is infinitesimally extended.

\(^{24}\)By interpreting (1) this way, White can avoid any difficulties associated with motion during a single point of time. This interpretation is a nod to Aristotle’s view that motion or rest during an instant is not coherent. On Aristotle’s view, motion (rest) requires a span of time over which to move (rest).

\(^{25}\)A natural choice for this function would be, $z(t) = -9.8t^2 + 73t + 1.8$ (where time is in seconds and length is meters). The length of the interval is therefore $-19.6t_0\varepsilon + 73\varepsilon + \varepsilon^2$, an infinitesimal value.
3.3.3 Looking Elsewhere

There are some concerns about viewing the present as an extended period of time. Once one leaves the comfort of representing the present by the uniqueness of a single point, there is a great difficulty in narrowing down how far the present is to extend. Almost any choice will be arbitrary on some level. White’s own solution is a step away from total arbitrariness. Nevertheless, White has not escaped the arbitrariness completely.

If one argues that the present is extended, a natural start is to pick some finite length. The difficulty then becomes, what is the duration of the present? Suppose for starters, I arbitrarily select a second in time for the length of the present. But, when I am at the gym, my heart beats in intervals smaller than half-seconds. The present should be no larger than one of these heart beats. Maybe a quarter-second would do the trick. However, any twenty-first century computer shows that multiple billions of events occur within a quarter-second. From this brief examination, it appears that the present is being squeezed from both sides by the past and future. For any arbitrarily small selected length for the present, an even smaller one suggests itself.

White’s solution has an advantage over the above approach: infinitesimals are smaller than any practically measurable length of time. In that way, infinitesimal lengths are not subject to the kinds of obvious counterexamples that ordinary finite values would be. Furthermore, his solution provides a mathematically definable class of lengths from which to select the length of the present. With that in mind, I will

---

26 Along these lines, some ancient atomists developed the view according to which matter, space and time are all composed of extended “atoms” or simples—things with no proper parts—of matter, space and time respectively. The present would count as one of these finitely extended time simples. On certain interpretations, these atomists developed these views in an effort to answer various of Zeno’s paradoxes. For rich discussion, see Richard Sorabji’s Time, Creation and the Continuum: Theories in Antiquity and the Early Middle Ages, [196], especially Part V.
now critically examine White’s solution in more detail.

White’s own remarks are very brief. He writes, “let us identify the present with an (any) infinitesimal interval of time.” (White, [213], 242) In spite of the brevity, I hope to reconstruct his reasoning by filling out this idea further. White begins by viewing the hyperreals from the standpoint of full expressive power of mathematical English with a suitably powerful background theory like Zermelo-Fraenkel set theory with Choice (ZFC). From this metalinguistic vantage point, White selects one of the hyperreal infinitesimals at random and then stipulates that the length of the present is always that value in any discussion where it is needed. The fact that White has allowed for any infinitesimal interval to count as the present suggests that even among the infinitesimals, there is significant slack in choice. As a result, just as in the finite case, for any selected interval for the length of the present, there is a smaller one that would do just as well. Such considerations suggest that although the infinitesimal strategy is a good one, something better might be available.

Chapter 3.4

Outline for a Solution

White’s solution to Zeno’s arrow is in rough outline very similar to my own. Truthfully, the differences in my approach from White’s are few in comparison to the wealth of similarities. Allow me to begin with a basic discussion of what the present is, especially in regard to its place in the discussion of the Arrow. ‘Present’ often comes with a qualification ‘the present century,’ ‘the present year’ or ‘the present minute.’

---

27I must be especially careful in this discussion because the hyperreals were traditionally constructed from a metalinguistic point of view that contrasts with a purely first-order version of analysis. On this original conception, the first-order sentences of real analysis are considered “internal” and the metalanguage is something much more powerful like ZFC. In this setting, when speaking of the object language, I mean the whole theory of hyperreals; when speaking of the metalanguage, I mean all of mathematical English used by a practicing mathematician-cum-philosopher which would include ZFC or more.
The problem is that when ‘the present’ is neither qualified directly as above or indirectly by context, one might want to know what is being indicated. Some settings seem to indicate that the present should be as small as possible, perhaps even punctiform (point-like). As an everyday example, I might naturally think while driving from Columbus to Cleveland, “My present speed is 65 mph,” where ‘present’ is used adjectivally. If I were in an especially philosophical mood (and sensitive to using ‘the present’ as a noun), I might think, “My speed during the present is 65 mph.” Whether expressed adjectivally or using a singular term, call this unqualified usage the *absolute* present.28 This final sense of ‘the present’ is what is at issue in Zeno’s arrow.

Below, I attempt to provide a solution to Zeno’s Arrow that treats the the present in some way other than a point-instant. I propose the following three requirements as guides:

(a) Arbitrariness is generally to be avoided.

(b) The absolute present must be given a fitting mathematical characterization.

(c) The absolute present is to be as small as possible.29

28For the present discussion I am ignoring the thorny issue of Einstein’s relativity. I am here considering the basic simple motion to which ordinary Newtonian mechanics apply. These are the kinds of mechanics at which I think Zeno aims his Arrow. To engage the harder questions governing motion at the microscopic or astronomical levels of observation is to jump a ways downstream dialectically.

29The reader might be interested to discover these guideposts eschew a pair of responses that have been given in the past in answer to the question, “What is the extent of the present?” To use an extended chunk of time to count as the present is not without historical precedent. The first answer is found in the ancient atomists. They envisioned the whole of time as being composed of finite, extended, indivisible time blocks. For the Atomists, the present would count as one of these. For a more detailed discussion, see Richard Sorabji’s *Time, Creation and the Continuum: Theories in Antiquity and the Early Middle Ages*, [196], or “Atoms and Time Atoms,” [?]. In the case of the Atomists, no natural algebraic characterization is available that does not somehow already conflict with either (a) or (c). If an Atomist satisfies (c) and gives a nice small margin into which the present fits, then she will run afoul of guidepost (a): one wonders why a smaller finite unit would not be better suited.

This final remark applies also to a more contemporary attempt to understand the present, viz. the specious present—the basic unit of temporal consciousness. Here are some charming remarks from William James:

In short, the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own on which we sit perched, and from which we look in two
I will use these guides below both to assess White’s solution in greater detail and also to judge the quality of solutions I offer that use nilpotents.

### 3.4.1 Avoiding Arbitrariness

Starting with (a), a first thought is, “What’s so bad about arbitrariness?” To desire to avoid arbitrariness is to tacitly endorse a form of the principle of sufficient reason: all facts admit of explanations. I do not aim to defend the principle of sufficient reason as an unimpeachable requirement on all philosophical work. I take it to be a maxim that one should hold alongside other methodological principles. Violations of this maxim can be accepted only if there is reason to believe that arbitrariness is constitutive of the topic under discussion. Constitutive arbitrariness could be established by either (i) simply making a positive case that no explanation could be had or (ii) a noticeably long history of failures to provide an explanation. Unfortunately for White, neither of these roadblocks to philosophical explanation is found for the length of the present.

One way out for White is to represent the present within the hyperreals as the monad of a given point in time, recalling from §3.3.1 that the monad of a point \( x \) is the set of all points infinitesimally close to \( x \). This has the advantage of providing a unique representation of the present. Unlike the sundry intervals, no other comparably larger or smaller set suggests itself. Selecting the monad also will help in satisfying (b), to which I now turn.

---

Besides failing along a similar vein as the atomists, the specious present is not well-suited to be a first response to Zeno. The reason is that it takes the phenomenology of time more seriously than Zeno would ever allow. Again, it is not as if Zeno’s psychology was different from anyone else’s: the ability to discriminate change is already inherent in the doctrine of the the specious present. If simply using one’s own psychological capacities is an acceptable move, then Zeno is done before he starts.
3.4.2 The Present in Mathematics

Consider now (b): The absolute present must be given a fitting mathematical characterization. This guide is more directly relevant to the Eleatic discussion, in particular because significant weight has been placed upon the extent of the present—is it an interval? Is it a point? Such questions are most naturally settled using tools from mathematics. Aristotle leads the way here in his response to Zeno and related analyses of mechanical concepts. A major part of his solution requires discussion of the geometric concept point and how points figure in larger intervals. Using mathematics to provide a precise characterization of the present is obviously valuable. However, this requirement goes deeper. The present cannot count as simply any mathematical object; a dodecahedron will not do. In other words, the selection of theory is important here as well as the selection of some constituent within the theory to play the role of the present. Here, the concept of motion is playing a central part alongside the concept of the present. It follows that I must select some theory with a concept of a derivative and then select the appropriate constituent of that theory for the present moment.

As far as White is concerned, he uses hyperreal infinitesimal intervals to play the role of the present. Nonstandard analysis is a fine selection as a theory. However, I already indicated above that there is a mismatch between White’s velocity-during-the-present and the derivative within nonstandard analysis. Within nonstandard analysis, $f$ has a derivative whenever there is some standard $m \in \mathbb{R}$, and

$$m \simeq \frac{f(t + \varepsilon) - f(t)}{\varepsilon}, \text{ for all nonzero } \varepsilon \simeq 0.$$ 

A natural thought is that whatever the present is, motion-during-the-present should match the derivative concept within the relevant mathematical theory. In the limit-
based presentation of the calculus, the present is best understood as a single point.\footnote{\textit{Best} here is meant to be relative to the choice of theory and presupposes a laying aside of Eleatic worries about the limit-based presentation.} In other words, in ordinary Cantor-Dedekind analysis, motion-during-the-present has its formal image in the limit-based concept of the derivative. One would expect an analogous move when transitioning to nonstandard analysis. The analogue for the present within nonstandard analysis would be the class of all points infinitesimally close to the present, the monad. Within nonstandard analysis, each point has a monad, and it is the monad that is at the core of the nonstandard derivative concept.

Oddly, White has chosen something different: his motion-during-the-present is related to but not identical to the derivative concept. Although White could avail himself of the derivative concept, it would remain opaque what distinct work his motion-during-the-present is supposed to do and how it relates to the nonstandard derivative. As White has defined it, motion-during-the-present is theoretically inert. Above, I indicated that White’s solution would be much improved on the arbitrariness scale, if he would select the monad for the present. The monad streamlines naturally with the formal theory already in place, since the nonstandard definition of the derivative uses \textit{all} infinitesimals, not just \textit{any} infinitesimal. This approach would allow White to dodge the arbitrariness challenge as well as develop a notion of motion-during-the-present that fits into the full theory of nonstandard analysis.

### 3.4.3 The Miniature Present

Consider now (c): the absolute present should be as small as possible. Here one must take a rudimentary glance at the concepts within the A-series: past, present and future. The central conviction behind (c) is that the present moment is being squeezed from both sides by the past and future. As such, the absolute present exhibits what G.E.L. Owen calls \textit{retrrenchability}: when called to a deliver a value for
the extent of present’s boundaries, one is wont to squeeze it into smaller and smaller regions of time. Why is this exactly? Recall my imaginary journey from Columbus to Cleveland. Suppose that near Mansfield—a city north of Columbus but south of Cleveland—I am traveling only 50 mph, perhaps because I am thinking too much about Zeno. Once I realize this, I accelerate up to 65 mph and reach that speed by the time I am on the north end of Mansfield. At some moment between those two speeds, I would have been at 60 mph. The present needs to be short enough that one can coherently talk about the moment when the car is going 60 mph, punctiform or not. Infinitesimals are effective to this end, provided that one tolerates speeds within an infinitesimal of 60mph. Within nonstandard analysis, the derivative is defined with exactly this level of tolerance in mind. Recall that a function \( f \) has a derivative in nonstandard analysis whenever there is some standard \( m \in \mathbb{R} \), and

\[
m \simeq \frac{f(t + \varepsilon) - f(t)}{\varepsilon}, \text{ for all nonzero } \varepsilon \simeq 0.
\]

So far, so good. Below, I will actually suggest that Giordano’s Fermat reals have the edge, because they are in some sense smaller than the infinitesimals of the hyperreals.

* * * * * * * *

Allow me to review the ground covered so far. White denies Zeno’s premise (3): that in the present the arrow is occupying a space just its own size. He argues that the present is extended. With the mechanical picture accompanying Robinson infinitesimals, it follows that the arrow is not occupying a space just its size, but rather a size infinitesimally larger than its own. I will deny the same premise in what follows. I also have provided the following guideposts for shaping a denial of premise (3): (a) arbitrariness is generally to be avoided; (b) the absolute present must be

\[31\text{This discussion is found in “Aristotle on Time,” [146].}\]
given a fitting mathematical characterization; (c) the absolute present is to be as small as possible. White's hyperreal infinitesimal intervals do not fare well on any of (a), (b) or (c)—the last of which is to be explored more below. If he were to model the present using the monad, then he would be better off with respect to (a) and (b). In the following, I will examine two other theories of infinitesimals along the lines of these three guideposts.

Chapter 3.5

SIA and Kock-Lawvere Nilpotents

In this section, I sketch an alternative to the point-instant picture of the present along White’s lines. In rough outline, White’s solution to Zeno’s arrow is very similar to my own. Unlike White, I invoke the theory of Smooth Infinitesimal Analysis (SIA). This system contains a special kind of infinitesimal, a nilsquare. Recall that a nilpotent $\varepsilon$ is any number where $\varepsilon^k = 0$, for positive integer $k$; nilsquares are nilpotents with $k = 2$. Now, imagine a segment of nilsquare length $\varepsilon$. Oddly enough, it follows that a geometric square with sides of length $\varepsilon$ has zero area, since $\varepsilon^2 = 0$. One can see now why they are called nilsquares: the square is nil. This is certainly not possible in ordinary (Euclidean) geometry. What makes this particular theory of nilpotents especially interesting is that it requires an intuitionistic logical consequence relation. Unlike other intuitionistic developments, this one is not connected with more elaborate philosophical considerations.\[^{32}\]

\[^{32}\]There are two philosophical schools that take intuitionistic logic very seriously. The first one is from the mathematician L.E.J. Brouwer. On this view, mathematical concepts and statements are a matter of mental construction. Generally, in order to assert any mathematical disjunction ($\varphi \lor \psi$), one must have produced either a proof of $\varphi$ or a proof of $\psi$. One may not simply assert for any statement $\varphi$, ($\varphi \lor \neg \varphi$) since at any time there are many yet unsettled mathematical statements. The other school is often referred to as semantic anti-realism, an updated form of verificationism, developed along similar lines by each of Michael Dummett, Dag Prawitz and Neil Tennant. On this view, all truth—not just mathematical truth—is bound up with verification. As a result, these thinkers’ eschewal of excluded middle involves understanding logical constants such as disjunction...
middle is not merely withheld; to assume it leads to contradiction. In the immediate sequel, I will first explain this requirement along with other essentials of this theory. Following that, I will apply SIA directly to the Arrow.

### 3.5.1 Basic Presentation

The theory of SIA begins with the intuition that any curve is the fusion of incredibly small straight lines. Bell writes:

> An infinitesimal may be taken to be an ‘ultimate part’ of a continuum: in this same sense, mathematicians have on occasion taken the ‘ultimate parts’ of curves to be infinitesimal straight lines. (Bell, [9], 3)

Curves are represented by smooth functions, those that can be differentiated as many times as one likes. There are no angles or abrupt shifts. For the sake of illustration, consider a jump rope with cylindrical plastic beads of alternating colors. The jump rope is curved but each bead is straight along this curve. So it is with smooth curves in SIA except that instead of having one hundred and fifty beads of an inch length, it is more like having infinitely many beads of infinitesimal length. Of course, given some of the more technically delicate aspects of this theory, do not extend this analogy too far. The beads are only meant to provide some helpful imagery.

The basic object of SIA is the smooth-real line, \( \mathbb{R} \). Unlike the hyperreals, \( \mathbb{R} \) is not a model-theoretic extension of \( \mathbb{R} \). One of the most central and significant differences is that all nilsquares are zero for the ordinary reals (as well as for the hyperreals). Not in \( \mathbb{R} \)! This is because in order to capture the idea that curves are straight-in-the-small, the nilsquares in \( \mathbb{R} \) must be nondegenerate; i.e. not all nilsquares are zero.

In keeping with Bell’s exposition on this topic, I will denote the set of nilsquares as in a manner similar to the one described above. One is entitled to assert some disjunction only if she is already entitled to assert at least one of the disjuncts. A nice discussion of these views can be found in Neil Tennant’s *The Taming of the True*, [200]. Neither of these historical motivations for intuitionistic logic is behind SIA, but rather a completely different strand of thought that emerged by developing mathematical logic from within Category Theory.
\[ \Delta.a \] Also, for any point \( a \) in \( \mathcal{R} \), let \( \Delta_a \) denote the micro-neighborhood around \( a \); that is, the nilsquare “haze” around a point \( a \). Formally,

\[ \Delta_a = \{ x \in \mathcal{R} : (\exists \varepsilon \in \Delta) \ x = a \pm \varepsilon \}. \]

Before showing the nondegeneracy, I must introduce the Kock-Lawvere Property (KLP), which very loosely says that every function is straight on the nilsquares (or straight-in-the-small). Formally,

\[ (\forall f : \Delta_a \to \mathcal{R})(\exists! m \in \mathcal{R}) \ (\forall \varepsilon \in \Delta) \ f(a + \varepsilon) = f(a) + \varepsilon \cdot m. \]

The idea behind KLP is that at any point \( a \), and function \( f \), there is a tangent line for \( f \) with slope \( m \) at point \( a \) spanning the nilsquares. Notice that KLP is the natural representation for differentiation and therefore velocity within SIA where one takes \( m = f'(a) \).

I will now begin to unravel some of the interesting results. It follows very quickly from KLP that the nilsquares are nondegenerate, in particular zero is not the only nilsquare. For contradiction, suppose the nilsquares are degenerate, i.e. \( \Delta = \{0\} \).

According to KLP, there is some unique \( m \) such that

\[ f(a + \varepsilon) = f(a) + \varepsilon \cdot m, \text{ for all } \varepsilon \in \Delta. \] (3.5.1)

However, since \( \Delta = \{0\} \), statement (3.5.1) above is equivalent to \( 0 = 0 \cdot m \). The problem is that \( 0 = 0 \cdot x \) for all \( x \), not just this \( m \). So, \( m \) is not unique—contradiction.

\[ ^{33}\text{I draw extensively from Bell’s lucid monograph, [9], A Primer of Infinitesimal Analysis, 2nd Edition.} \]

\[ ^{34}\text{The expression ‘}\exists!\text{’ expresses unique existence and can be paraphrased in the familiar way using first-order logic with identity.} \]

\[ ^{35}\text{The reason it is written this way is that } 0 \in \Delta \text{ and so one cannot simply divide by all of the } \varepsilon \in \Delta \text{ as in the usual expression of the derivative as a quotient.} \]
Therefore, the nilsquares are nondegenerate. Regarding this proof sketch, note that I have only used intuitionistic reasoning. Furthermore, I have only assumed that $\mathcal{R}$ satisfies the field axioms.\[36\]

The fact that $\mathcal{R}$ is a field is at the center of the most surprising result in SIA: there are no nonzero nilsquares. For contradiction, suppose there is some nonzero nilsquare $\varepsilon$. Then $\varepsilon$ has a multiplicative inverse, $\frac{1}{\varepsilon}$. Now, since $\varepsilon^2 = 0$, it follows that

$$0 = \varepsilon^2 \cdot \frac{1}{\varepsilon} = \varepsilon \cdot \left( \varepsilon \cdot \frac{1}{\varepsilon} \right) = \varepsilon \cdot 1 = \varepsilon.$$ (3.5.2)

Tracing the equality through (3.5.2), it follows that $\varepsilon = 0$—contradiction. Therefore, there are no nonzero nilsquares. This is a very peculiar result since I have that not all nilsquares are zero but it is false that there is some nonzero nilsquare. Such an extraordinary tension is maintained only by the necessary use of a logic no stronger than intuitionistic logic. In classical logic, the claim that not all nilsquares are zero implies that there is some nonzero nilsquare.\[37\] This tension is exactly provides the special resources to answer Zeno’s Arrow.

### 3.5.2 SIA and the Arrow

In this section, I will use the smooth-reals instead of the hyperreals to model space and time.\[38\] Unlike White, I do not exactly argue that the present is extended; in fact, it is not extended. However, it is not a point either. This is enough to resolve the paradox. I claim that the present should be represented within SIA as the nilsquares, $\Delta$. In effect, the present is no shorter than any $\varepsilon \in \Delta$. Like White, I can avoid treating the present as a point, and thereby dodge Zeno’s Arrow. The fact that the

---

\[36\]See §2.3.3 for the field axioms. Detailed discussion of SIA is also found in §2.4.3.

\[37\]This peculiarity can be stated so as to directly undermine the law of excluded middle. In particular, in SIA, it is provably false that for any nilsquare $\varepsilon$, either $\varepsilon = 0$ or $\varepsilon \neq 0$.

\[38\]Recall from §3.3.2 that I am not concerned to settle the structure of space but rather deny Zeno’s contention that motion is impossible.
nilsquares are nondegenerate allows me to deny Premise (3): in the present the arrow
is occupying a space just its own size.

How does that work? This will take the same shape as White’s solution. Let
the present be the micro-neighborhood around some point $t_0$ in time, written $\Delta_{t_0}$.
Let the arrow’s flight remain the same as above (see §3.3.2); that is, $z(t)$ charts the
arrow’s flight and $z(t_0) = z_0$. Now, define $T_0$, as above, to be the set of points where
the arrow is located during the present. In this case, $T_0$ is the set of all the places
where the arrow is located at time $t$ where $t$ can vary within the micro-neighborhood
of $t_0$. Formally,

$$T_0 = \left\{ z \in \mathcal{R} : (\exists t \in \Delta_{t_0}) \ z = z(t) \right\}.$$ 

In order to deny that in the present the arrow is occupying a space just its own size,
suppose for contradiction it is true; i.e. assume for contradiction that $T_0 = \{z_0\}$. It
follows from this that the micro-neighborhood around around $t_0$ is degenerate, $\Delta_{t_0} =
\{t_0\}$. But it was demonstrated above that no micro-neighborhood is degenerate—
contradiction. Because SIA uses intuitionistic logic, this conclusion does not grant that
$T_0$ is extended; indeed, $T_0$ is not extended. The fact that $T_0 \neq \{z_0\}$ is just enough
to deny Zeno his premise (3). Although the set of infinitesimals is nondegenerate,
zero is the only infinitesimal that provably exists, unlike with the hyperreals. This
peculiar aspect of SIA grants victory with the same costs as before.

3.5.3 Evaluating SIA

I will now evaluate SIA according to the guideposts given above:

(a) Arbitrariness is generally to be avoided.

(b) The absolute present must be given a fitting mathematical characterization.

\footnote{Strictly speaking, to make this move, I am assuming that micro-neighborhoods are topologically
connected. Therefore, although possibly $z(t) = z_0$ for some $t \neq t_0$, if $t \in \Delta_{t_0}$ and $z(t) = z_0$, then
t = t_0.}
(c) The absolute present is to be as small as possible.

Starting with guidepost (a), SIA stands proud. White’s own suggestion of the infinitesimal intervals is arbitrary: there is no obvious reason for selecting one rather than another. Of course, I suggested a way for White that takes the present to be a monad. Compared to this suggestion, SIA and nonstandard analysis are fairly comparable.

I will return to (b) and now consider (c), since comparatively little will be said. SIA distinguishes between infinitesimals (members of $\Delta_a$) and points (member(s) of $\{a\}$) in the weakest manner possible. No exemplar from $\Delta$ is available to demonstrate their distinctness in the ordinary way. This fact makes the present as small as possible without actually being zero, even to the extent of bending logical consequence into conformity. This is an improvement even over the use the hyperreal monad, since (ignoring the arbitrariness for the moment) there are numerous classes of infinitesimals included in each monad; indeed, every infinitesimal interval around any point $a$ is included in $a$’s monad. Another way of putting this is that in SIA, there is nothing smaller than $\Delta_a$ included in $\Delta_a$ than $\{a\}$—the latter being ruled out already. By comparison, in nonstandard analysis, there are uncountably many classes strictly between (in terms of set-theoretic inclusion) $\{a\}$ and the monad of $a$. In fact, for any such class around $a$, there is one that is yet smaller but not identical to $\{a\}$. This suggests that SIA has the edge in terms of size.

Finally, I consider guidepost (b). I indicated in §3.4.2 that (b) can be violated by a poor choice of theory or by a poor choice of representational constituent from within a theory. I will call the former into question even though $\Delta$ is a fitting choice for the present once SIA is selected. Just like the monad for the hyperreals, $\Delta$ streamlines naturally with the presentation of motion-during-the-present for SIA, since the derivative concept is always defined using all of $\Delta$ as expressed by KLP.
Unfortunately, sIA is not an ideal choice of theory more generally, independent of how well it might model the present. Limiting oneself to intuitionistic logical consequence is not theoretically cheap. Intuitionistic logic prevents one from making many natural, rational maneuvers found in classical logic, not the least of which is the Law of Excluded Middle. Furthermore, the principles used to argue in favor of the use of exclusively intuitionistic reasoning are often less obvious and natural than the logical laws one is being asked to abandon.40 In the case of Zeno’s Arrow, more is needed than a paradox of motion to force one to adopt such logical strictures.41

There is another problem worth mentioning for sIA. The strange nature of the set of nilsquares forces one to surrender a rather closely held property from Classical Extensional Mereology, **Weak Supplementation**, (ws). (Ws) states that for any x and y, if x is a proper part of y, then there is some z that is proper part of y and shares no parts with x.42 When considering the relation of a part to a whole within a universe of sets, parthood is interpreted to be the relation of nontrivial inclusion or nonempty subsethood. Translated into set-theoretic terminology, (ws) states: for nonempty X and Y,43 if X ⊊ Y, then there is some nonempty Z ⊊ Y and Z ∩ X = ∅.

I claim that in sIA, (ws) is false. For contradiction, suppose (ws) is true. Note that

---

40 Thomas Nagel makes a similar point in *The Last Word*, [142], though not directly about intuitionistic logic. The basic idea is that logical principles can only be doubted by more fundamental kinds of facts. Attempting to discover facts more fundamental than logical facts strikes me as futile.

I will admit that I find arguments from harmony regarding the introduction and elimination rules of logical particles compelling. Yet, I must confess that the implausibility of the refusal to affirm the Law of Excluded middle is greater than the persuasiveness of the argument from harmony.

41 In the spirit of tolerance, I am inclined to reserve a place for sIA, even if I do not favor it in the end. It’s noteworthy that sIA has models in category theory: a fairly recent, classical abstract mathematical system that, among other things, provides an alternative to set-theoretic foundations. Furthermore, category theory has gained serious respect among working mathematicians, specifically those working in the area of algebraic geometry. The point of all this is that sIA satisfies the best standards of rigor that the typical working mathematician would expect. For more on these issues, Geoffrey Hellman provides an excellent analysis in his “Mathematical pluralism: The case of smooth infinitesimal analysis,” [88]. John Bell devotes some discussion to this issue in the Introduction to *A Primer of Infinitesimal Analysis*, [9]. Finally, see Awodey’s “An Answer to Hellman’s Question: ‘Does Category Theory Provide a Framework for Mathematical Structuralism,’ ” [3].

42 See Peter Simons’ *Parts*, [190].

43 If Y is nonempty, then this principle is false; if X is nonempty, then it is vacuous.
\{0\} \subset \Delta. By (ws), there is some \textit{nonempty} \( Z \subset \Delta \) but \( Z \cap \{0\} = \emptyset \). From the fact there are no nonzero nilsquares, if \( Z \) shares no members with \( \{0\} \), then \( Z = \emptyset \) — contradiction. Therefore, (ws) is false. This is very surprising given that (ws) is a rather weak requirement to place on parthood, let alone the fact that ordinary set theory satisfies incredibly strong mereological principles.\footnote{Using the limitation of size presentation of set theory, one can see that even General Extensional Mereology is satisfied provided that the number of things forming a fusion is small (in the technical sense, where any collection is small just in case the universe cannot be injected into it.) For lengthy discussion of the interrelationship between set theory and mereology, see Lewis’ \textit{Parts of Classes}, [124], and Caplan, Tillman and Reeder’s, “Parts of Singletons,” [33]. The latter also contains numerous references to other relevant literature.}

\textbf{Chapter 3.6}

\textbf{Giordano’s Fermat Reals}

In this section, I will use a number system recently developed by Paolo Giordano, the Fermat reals, to model space and time. The system is similar to \textit{sia} both because it contains nondegenerate nilpotents (nonzero values \( \varepsilon \) where \( \varepsilon^k = 0 \) for some positive integer \( k \)) but also because it satisfies a version of KLP, \textit{viz.} for open \( A \subseteq R \),\footnote{Here I use a generic letter \( R \) since the property could be stated either for either \( \mathcal{R} \), \( \mathbb{R} \) or in this case, the Fermat reals defined below. Of course, it is false for \( \mathbb{R} \) on account of the non-uniqueness of \( m \) in \( \mathbb{R} \).} if \( f: A \to R \) is smooth, then

\[
(\exists! m \in R) (\forall \delta \in D) \quad f(a + \delta) = f(a) + \delta \cdot m,
\]

where \( D \) is the set of nilsquares (the set of values \( \varepsilon \) with \( \varepsilon^2 = 0 \)) within the Fermat reals. One crucial difference is that the consequence relation is classical. The reader will see later reintroducing classical logic comes with a few costs, both algebraically and also from the point of view of answering Zeno.
3.6.1 Basic Presentation

Allow me to begin with a basic characterization of the Fermat reals. The Fermat Reals, just like Robinson’s hyperreals, extend the ordinary Real numbers, i.e. for each real number, there is a corresponding Fermat real that plays the exact same part. In what follows, I will review some basic properites of the Fermat reals from §2.5.4. Following that, I will consider some algebraic properties the system has, viz. those that are important for geometric and physical investigation. The reader may skip this section without significant loss.

Constructing the Fermat Reals

The basic idea is that one treats any constant function \( f(t) = c \), for some \( c \in \mathbb{R} \) as playing the part of \( c \) in the new larger set of Fermat Reals, \( \bullet \mathbb{R} \). One obtains greater diversity by allowing for polynomials over the reals to play the role of infinitesimals. Officially, the Fermat reals are equivalence classes of members of what I will call Fermat polynomials, \( \mathbb{R}_o[t] \).\(^{46}\) Recall from §2.5.4 that each \( x_t \in \mathbb{R}_o[t] \) may be written as,

\[
x_t = \sum_{k=0}^{n} \alpha_k t^{b_k} + o(t) \quad \text{as} \quad t \to 0^+ \quad \text{where} \quad n \in \mathbb{N}, \alpha_k \in \mathbb{R}, b_k \in \mathbb{R}^+.\(^{47}\)
\]

\(^{46}\)Giordano calls these ‘little-oh polynomials.’ I find this slightly confusing given some other terms he uses, hence the new title.

\(^{47}\)Before using little-oh notation exclusively from now on, I remind the reader that the \( o(t) \) in (3.6.1) ensures that

\[
\lim_{t \to 0^+} \frac{x_t - y_t}{t} = 0
\]

where

\[
y_t = \sum_{k=0}^{n} \alpha_k t^{b_k}
\]

In other words, if \( y_t \) is the \( n \)-ary summand only, then one may add a little-oh function to it and still be within \( \mathbb{R}_o[t] \). These so-called little-ohs enable both notational simplicity and technical maneuverability in the construction before washing them out with the equivalence relation.
These have the shape of ordinary polynomials but instead of positive integral powers, any positive power is eligible (hence, the \( b_k \)'s). Also, I need the equivalence relation responsible for the aforementioned equivalence classes of Fermat polynomials. Let \( x_t, y_t \in \mathbb{R}_o[t] \) and define \( x_t \sim y_t \) to mean that

\[
\lim_{t \to 0^+} \frac{x_t - y_t}{t} = 0. \tag{3.6.2}
\]

Another way of writing this (which will avoid gratuitous use of fractions) is

\[
x_t = y_t + o(t) \text{ as } t \to 0^+. \tag{3.6.3}
\]

Finally, remove the little-oh\(^{48}\) to reveal the Fermat Reals, \( \bullet \mathbb{R} \). Explicitly, the Fermat reals are defined to be the set of equivalence classes from the Fermat polynomials, \( \mathbb{R}_o[t] \), under the relation \( \sim \) from (3.6.2). Recall that the infinitesimals \( d \) within the Fermat reals are such that \( d^k = 0 \) but \( d \neq 0 \); this is the extraordinary and distinctive feature of nilpotents.

Now for some bookkeeping. In the future, I will need to refer to what’s called the “standard part.” This is discussed at length in §2.5.4, so I will simply define it outright here. If \( y(t) \in \bullet \mathbb{R} \), then the standard part, written as \( \circ y \), is \( y(0) \). The beauty of this definition is that the infinitesimal fringe is determined by the polynomial part with nonzero degree; most specifically, the part which vanishes when \( t = 0 \). Naturally, one may define the infinitesimals of the Fermat reals to be those values \( D_\infty = \{ x \in \bullet \mathbb{R} : x(0) = 0 \} \).

\(^{48}\) I say ‘remove’ only to be colorful. The “removal” is actually by defining the quotient \( \mathbb{R}_o[t]/\sim \).
Algebraic Properties of Fermat Reals

Here I will discuss the algebraic properties of the Fermat Reals and contrast them with important properties for Robinson’s hyperreals. The first and most important difference is that the hyperreals form a linearly ordered field, whereas the Fermat reals form a linearly ordered, commutative ring. What this amounts to is that the Fermat Reals do not guarantee that all nonzero values have multiplicative inverses; i.e. for all \( x \neq 0 \), there is some \( y \) such that \( xy = yx = 1 \). Fields are very strong because they provide one not only with algebraic command but also are the most natural representatives for representing space.

As it turns out, the loss is minor for two reasons. The first reason is that the Fermat reals have a cancellation law that mimics invertibility: For nonzero \( x \in \mathbb{R} \) and \( a, b \in \mathbb{R} \)

\[
xa = xb \implies a = b
\]

(3.6.4)

In other words, provided that \( a, b \) are ordinary real numbers without any infinitesimal “fringe,” the \( x \)'s can be cancelled. The second reason is that as long as \( x \in \mathbb{R} \) is not infinitesimal, then \( x \) has a multiplicative inverse. Another natural cancellation law follows immediately from this. For any \( x, a, b \in \mathbb{R} \)

\[
^o x \neq 0 \& xa = xb \implies a = b. \quad 49
\]

(3.6.5)

The reason why these cancellations laws are slightly weaker than the typical ones is that the nilsquares provide direct refutation of the ordinary unqualified principle. For any two distinct nilsquares, \( d_1, d_2 \), their product is zero. Hence, there are cases where given some further nilsquare \( d_3 \), \( d_3d_1 = d_3d_2 \) but \( d_1 \neq d_2 \) by stipulation. All of this

\[49\]Let the reader note that because the Fermat reals are multiplicatively commutative, these cancellation laws could have been presented with \( ax = bx \) as the first clause.
is important in trying to understand the trade-offs between Robinson hyperreals and the Fermat reals. I will discuss this more below in §3.6.3.

3.6.2 The Fermat Reals and the Arrow

When it comes to determining the effect of the Fermat reals on the Arrow, one must understand how derivatives or velocity are handled. As with SIA, derivatives are negotiated using the Kock-Lawvere Property:

\[(\exists! m \in \mathbb{R}) (\forall \delta \in D) f(x + \delta) = f(x) + \delta \cdot m.\]

For explicitness, recall that $D$ is the set of nilsquares. With the right kind of algebraic rearranging, this is quite similar to the property given above for the hyperreals. The reason it is written this way is that $0 \in D$ and so one cannot simply divide by all of the $\delta \in D$.

So much for KLP—how does this help with the Arrow? The strategy is exactly the same as before. The present will be modeled in the Fermat reals as $D$, given $D$’s role in the derivative concept. By now, the reader has already seen in very similar systems how one might deny premise (3): in the present the arrow is occupying a space just its own size. For that reason, I am not going to rehearse things in the same detail as above. Especially given that the theory of Fermat reals uses a classical consequence relation, demonstrating that (3) is false here will look very similar to the work done for nonstandard analysis. Since the present is extended along $D$ in time, the arrow occupies a space just slightly larger than its own size during the present.

3.6.3 Evaluating the Fermat Reals

Here are the three guideposts again for the reader:

(a) Arbitrariness is generally to be avoided.
(b) The absolute present must be given a fitting mathematical characterization.

(c) The absolute present is to be as small as possible.

While assessing the Fermat reals, I will also remind the reader of how the others fared and if there are any strengths or weaknesses of $\mathbb{R}$ in comparison.

Starting with (a), how do the Fermat reals fare on issues of arbitrariness? Again, in comparison to White’s published solution, $D \subseteq \mathbb{R}$ is an improvement. When compared to the monad suggestion I gave, all three selections for the present (the monad, $\Delta$ and $D$) are equally non-arbitrary.

In this case, satisfying (b) is much more straightforward than for sia. As above, selecting the nilsquares $D$ as playing the part of the present is appropriate given the role they play in $\text{KLP}$ and therefore the derivative concept. The question remains whether the Fermat reals are a fitting choice of number system more generally.

For the Fermat reals, the classical consequence relation is reintroduced. This puts the Fermat reals above sia in my mind, but relatively equal with the hyperreals.

How do the Fermat reals compare to the hyperreals in other ways? Admittedly, the same level of algebraic richness enjoyed by the field of hyperreals is unavailable. Fortunately, multiplicative cancellation is available in the Fermat reals for a vast array of values as indicated by (3.6.4) and (3.6.5) above in §3.6.1. This delivers a decent proxy of the algebraic command provided by having multiplicative inverses in fields, like the hyperreals. Also, unlike the hyperreals, infinitesimals are not invertible; that is, there are no infinite Fermat reals. I take this to be an advantage actually. The lack of infinite values in the Fermat reals provides for a more natural representation of space and time.\footnote{Projective geometry employs points at infinity. For the more pedestrian purposes here of an arrow’s flight, if one can eschew points at infinity without significant loss, all the better.}

This way, one can have small values available for one’s needs without being committed to infinitely distant points in space.
Now, consider (c). Since I am using the nilsquares, these values are already very small. The fact that one may restrict an especially small class of values within the infinitesimals indicates improvement over White. The structure of the hyperreals does not admit of the same kind of fine-grained discrimination among its infinitesimals. There is an order defined on them, but its infinitesimals cannot be cordoned off as especially small or large infinitesimals as is the case with the nilpotent theories. John Bell writes,

It is to be noted that the property of being a nilsquare infinitesimal is an intrinsic property, that is, in no way dependent on comparisons with other magnitudes or numbers. (Bell, [9], 2, emphasis original)

Bell says this in regards to sia but it holds for the Fermat reals too. This is not so for classes of hyperreal infinitesimals. Any proper subset $I$ of hyperreal infinitesimals, with $I \neq \{0\}$, cannot be defined without explicit reference to a previously fixed value.\[^{51}\] When considering guidepost (c), *viz.* the smallness of the present, this intrinsicality feature of the Fermat infinitesimals gives it an edge over the hyperreals in using it against Zeno.

**Chapter 3.7**

**Conclusion**

I examined above what an answer to Zeno would look like that simultaneously respects contemporary mathematical developments and the supposition that the present is extended. To that end, White’s solution is satisfactory: infinitesimals are of the same form as other examples of motion across an interval. However, I have attempted to

\[^{51}\]It might be tempting to argue that this is also true of the nilsquares since they are defined using 2 and 0 in the $x^2 = 0$. Let the property $\Phi(x) := x^2 = 0\downarrow$. So, one may define the nilsquares to be \(\{x \in \mathbb{R}^*: \Phi(x)\}\). Note, however, that $\Phi(x)$ can be rewritten as $(\exists y)(\forall z)(y + z = z + y = z) \& x \cdot x = y)$. Such algebraic control is unavailable in the hyperreals, specifically because of the fact that the real numbers and hyperreals satisfy all the same first-order statements. This latter property is often referred to as the Transfer Principle.
persuade the reader that alternative infinitesimal systems might do better work in answering Zeno than the one White uses, the system of hyperreals.
Chapter 4

Contact

Touch me not so near:
I had rather have this tongue cut from my mouth
Than it should do offense to Michael Cassio
—Iago, Shakespeare’s Othello

[E]very body is, on every side of it, either in contact with some other body, or in the absence of any other body, in contact with pure ether.
—Bolzano, Paradoxes of the Infinite

The primary goal of this chapter is to examine the extent to which an alternative mathematical theory of the continuum could shed light on the concept of contact, that is, when two bodies touch. Contact is fraught with difficulty, specifically because of the problem of contact: pairs of bodies with certain topological profiles—regardless of their mechanical or electro-chemical profiles—are unable to touch without counterintuitive consequences. In what follows, I offer a novel solution that involves the introduction of infinitesimals—non-zero lengths as small as the infinite is large—into the ontology of space. Roughly, I will argue that two bodies should count as being in contact provided there is no more than an infinitesimal distance between them.

In what follows, I start by presenting some basic notions of part and place that will be necessary for framing or answering the contact problem. In §4.2, I will expand upon the contact problem, provisionally illuminating along the way some of the major

\footnote{See Bolzano, [17], §67.}

\footnote{I will use ‘contact’ and ‘touch’ (and their cognates) interchangeably.}

\footnote{Extended discussion of infinitesimals is found in Ch. 2 above.}
hurdles to be cleared by any particular theory of contact. In §4.3, I will offer my own solution followed by a cluster of theoretical desiderata in §4.4. In this latter section, I will also try to judge how well my solution fits the desiderata. The solution I offer, like any theory, has its costs. For those desiderata that I prioritize, it is the only one of its kind; however, I leave it to the reader to decide between it and other serious contenders depending on his sense of the value of the various desiderata. Following this, in §4.5, I will present some of the alternative solutions. Lastly, in §4.6, I will offer some refinements of the desiderata and see if this allows one to maneuver around some of the tensions created by the original, more demanding desiderata.

Chapter 4.1

Parts and Places

Allow me to develop a repertoire of concepts for framing the problem and its solutions. The two primary concepts that I will need are those of parthood (the relationship of a part to a whole) and location (the relationship of a material body to the space it occupies).\footnote{Authors vary when presenting theories of location whether the locations over which the quantifiers range actually exist. In other words, do theories of location force one to be committed to the thesis of absolutism about space? I do not have any strong opinions on this question and hope to remain this way.} In what follows I will present both core concepts used by participants in the debate over contact. I do not provide any novel definitions here. If the reader is familiar with these concepts, he may skip directly to the puzzle as presented §4.2.

4.1.1 Parts

In this section, I will present the simplest mereological system sufficient to describe the contact problem and its solutions. Before so doing, I must clarify that $x$ is a proper part of $y$ just in case $x$ is a part of $y$ and $x \neq y$. The system comprising
the following four axioms is called *Minimal Mereology* by Roberto Casati and Achille Varzi in their *Parts and Places*:\(^5\)

**Reflexivity:** For any \(x\), \(x\) is a part of \(x\).

**Anti-symmetry:** For any \(x\) and \(y\), if \(x\) is a part of \(y\) and \(y\) is a part of \(x\), then \(x = y\).

**Transitivity:** For any \(x, y\) and \(z\), if \(x\) is a part of \(y\) and \(y\) is a part of \(z\), then \(x\) is a part of \(z\).

**Weak Supplementation:** For any \(x\) and \(y\), if \(x\) is a proper part of \(y\), then some proper part \(z\) of \(y\) shares no parts with \(x\).\(^6\)

Before adding a final principle on top of *Minimal Mereology*, there are a few definitions to be considered. First, \(x\) and \(y\) *overlap* just in case they have at least one part in common. Second, \(x\) is the *fusion* of some things just in case each of them is a part of \(x\) and every part of \(x\) overlaps one of them. Finally, we give the following fusion principle:

**Finite Fusion:** For any \(x\) and \(y\), there is a fusion of \(x\) and \(y\).

Allow me to use the name *Minimal Mereology*\(^+\) (\(\text{MM}^+\), henceforth) to mean *Minimal Mereology* along with *Finite Fusion*. Logically speaking, \(\text{MM}^+\) is strictly between *Minimal Mereology* and *General Extensional Mereology* (\(\text{GEM}\)). The latter allows any things to have a fusion. **Finite Fusion** guarantees only arbitrarily large finite fusions. I introduce here a slightly weaker version than \(\text{GEM}\) because it is the simplest system I need in order to describe the problem and some of its solutions.

Now that a base-level theory of parthood is in place, I need to introduce some principles that are more contentious and are actually at the heart of Dean Zimmerman’s “Could Extended Objects Be Made Out of Simple Parts?” [225]. These principles are the two sides of the simples debate: *Is everything composed of ultimate parts*

---

\(^5\)See Casati and Varzi, [36], p.39. This system is also discussed among others at length in Peter Simons’ *Parts*, [190], Part I.

\(^6\)This statement of Weak Supplementation is slightly different than usual. In the consequent, the ‘proper part’ is often stated as just ‘part.’ As it turns out, these are provably equivalent.
(simples), i.e. objects with no proper parts? Or, alternatively, is everything “gunky,” i.e. does everything have a proper part? There are mixed views as well: some things have simples as parts, some things are gunky (each of their parts has a proper part). To make future discussion easier, I will present each of these views with a name:

**Simples:** Everything has a simple as a part.

**Gunk:** There are no simples; i.e. everything has a proper part.

**Mixed** There is a simple and there is something that has no simple as a part.

These principles do not follow from or contradict $MM^\dagger$. In fact, they are even logically independent of GEM, which is why I introduced them separately.

4.1.2 Places

The most basic notion of place needed to frame the problem of contact is that of being a *receptacle* of some material body. Explicitly, $M$ has receptacle $R$ just in case $M$ exactly occupies $R$. Intuitively, $M$ exactly occupies region $R$ when $R$ hugs the very contours of $M$: (i) every subregion is filled and (ii) no part of $M$ is outside of $R$.\(^7\)

To understand *receptacle* is only the beginning. Now allow me to develop the topological notions required for understanding the problem, *viz.* that pairs of bodies with certain *topological profiles* are unable to touch without counterintuitive consequences. Here I will put the framework into place necessary for capturing the relevant topological profiles.

First, allow me to introduce the Euclidean metric—a distance function between

\(^7\)More detailed analyses of receptacle and occupation are available, but for understanding the contact problem the above gloss will do. The interested reader will find helpful either Parsons' “Theories of Location,” [147], or Casati and Varzi’s *Parts and Places*, [36].
points—on any n-dimensional space:\(^8\)

\[ d(\vec{x}, \vec{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \]  

(4.1.1)

Here I take points in space to be basic;\(^9\) regions are (nonempty) sets of points. Below I use variables \(X,Y,Z\) for sets of points, \(x,y,z,p,p',p''\) ... for points and \(r\) for scalar values\(^10\) of the space under discussion. Below are common topological properties and concepts that are undergirded by the metric defined by (4.1.1).

**Neighborhoods:** For any \(r > 0\) and point \(p\) the \(r\)-neighborhood of \(p\) is defined to be the set \(N_r(p) = \{x : d(x,p) < r\}\).

**Limit Points:** \(p\) is a limit point of a set \(X\) iff for all \(r\), there is a \(q \in X\) where \(q \neq p\) and \(q \in N_r(p)\).

**Interior Point:** For any set \(X\), a point \(p\) is an interior point of \(X\) iff there exists an \(r\) such that \(N_r(p) \subseteq X\).

**Boundary Points:** For any set \(X\), a point \(p\) is a boundary point of \(X\) iff \(p\) is a limit point of \(X\) and not an interior point of \(X\).

**Open:** \(X\) is open iff every \(p \in X\) is an interior point of \(X\).

**Closed:** \(X\) is closed iff \(X\) contains all of its limit points.

**Mixed:** \(X\) is mixed iff there are distinct boundary points of \(X\), \(p\) and \(p'\), such that \(p \in X\) and \(p' \notin X\).

**Connected:** \(X\) and \(Y\) are connected iff there is a boundary point \(p\) of both \(X\) and \(Y\) and \(p \in X \cup Y\).

**Self-Connected:** \(X\) is self-connected iff any two sets \(Y,Z\) partitioning\(^11\) \(X\) are connected.

---

\(^8\)Here I am using the convention of \(\vec{x} = \langle x_1, \ldots, x_n \rangle\). I use the arrow here to disambiguate points (here, \(\vec{x}\)) and the location of a point along a dimension (here, \(x_i\) for \(1 \leq i \leq n\)). When ambiguity is absent, the arrow will be dropped.

\(^9\)For the present discussion, I do not aim to weigh in on what is basic or fundamental. I only mean that points function as *definiens*. A very lengthy discussion of fundamentality can be found in Schaffer’s “On What Grounds What,” [177].

\(^10\)A scalar value is one that has no dimension and does not exist in the space, strictly speaking but may be employed in describing or performing transformations on the space.

\(^11\)\(Y\) and \(Z\) partition \(X\) iff \(Y \cup Z = X\) and \(Y \cap Z = \emptyset\).
Distance between Regions: For \( X \) and \( Y \) \( d(X, Y) = \) the greatest lower bound of

\[
\left\{ d(x, y) : x \in X, y \in Y \right\}.
\]

Sphere: \( X \) is a sphere iff for some scalar \( r \) and some distinguished (center) point \( p \), \( \{ x : d(x, p) < r \} \subseteq X \subseteq \{ x : d(x, p) \leq r \} \).

* * * * *

Now with right tools for the job, on to the contact problem!

Chapter 4.2

The Problem

Above I stated that the contact problem is that there are pairs of bodies with certain topological profiles that are unable to touch without counterintuitive consequences. In this section, I explicitly articulate which pairs are trouble and how they are so. At this stage, I will only point casually towards preliminary metaphysical discomforts one might sense. It is not until §4.4 that I will spell out more seriously what one should avoid in a theory of contact and why.

The different kinds of contact can be reduced to three cases: (i) mixed pairs, (ii) open pairs and (iii) closed pairs.\(^\text{12}\) What exactly are these? A mixed pair of bodies includes one body with an open receptacle and one body with a closed receptacle; that is, the former’s receptacle contains none of its boundary points and the latter contains all of them. Scenario (ii) is when both bodies have open receptacles and (iii) covers those cases where both bodies have closed receptacles. There are problems for contact primarily for cases (ii) and (iii). For simplicity, let the exemplary material

\(^{12}\)In any case where the pair of bodies is more complex, any point of contact between the bodies will reduce to (i), (ii) or (iii) and will suffer from whatever difficulties are present for them.
body be a ball—a ball is any solid material body whose receptacle is a sphere. Since I will need to regularly refer to some pair, I will call them ‘Redball’ and ‘Greenball’ (or sometimes just ‘Red’ and ‘Green’).

4.2.1 Mixed Pairs

Figure 4.1. A Mixed Pair: the closed ball (left) includes the shared boundary point.

Take a mixed pair. Let Redball be open, so its receptacle does not contain its boundary points, but let Greenball be closed, so its receptacle does contain its boundary points. In this scenario, the two receptacles can be connected—loosely, one could draw a line from Red to Green without lifting the pencil. Red and Green come together perfectly and their fusion has as its receptacle a self-connected “peanut”: two spheres connected at exactly one point with half of the peanut open and half of it closed. This scenario is presented mainly for exhaustiveness, since mixed pairs do not impugn the possibility of contact. As far as I know, provided that receptacles of both types are permitted, there is no extant theory of contact where mixed pairs cannot touch.

\[\text{I use ‘solid’ somewhat casually here. If someone were to push me harder to define the notion of a solid, then I would offer the following. A region } X \text{ is } \text{solid} \iff \]
\[\begin{align*}
1. & \text{ } X \text{ is self-connected;} \\
2. & \text{ There is some sphere } S \text{ such that } S \subseteq X; \text{ and,} \\
3. & \text{ For any boundary point } p \text{ of } X, \text{ there is a closed sphere } S \text{ with boundary point } p \text{ such that } S \cap X = \{p\} \text{ and for any other closed sphere } T \text{ with boundary point } p, S \subseteq T \text{ only if } T \cap X = \{p\}.
\end{align*}\]
4.2.2 Open Pairs

Now, take an open pair: things start to become a little bit tricky. Redball and Greenball while occupying distinct regions will always be separated by a mutual boundary point—see Figure 4.2.

![Figure 4.2. An Open Pair: the shared boundary point is not included in either ball.](image-url)

This (tiny) distance cannot be closed without interpenetrating, i.e. parts of each of Red and Green—though composed of distinct matter—would (possibly non-exactly) occupy the same region. A natural thought is that two things are touching just in case there is nothing between them. In the present case, there is *something* between them, *viz.* an empty spatial point. Maybe a mere spatial point should not be intimidating. One could take two objects to be in contact just in case there is a zero distance between them—here satisfied by Red and Green. Call this **Measure-Zero Contact**. The problem with this definition is that if there is an object—call it Skinny—of less than three dimensions, I could slip Skinny between Red and Green. Even if Skinny and his family do not exist, there is still the space between Red and Green mentioned above. This presents a choice. Either one must accept **Measure-Zero Contact** and with it the gaps between touching objects or one must reject **Measure-Zero Contact** and thereby deny that Red and Green can touch. Here is a first impasse:

---

14 This is presented as (CONTACT-1) in Hudson’s [92] and [94].

15 Skinny is still a problem for me, but it is also a problem for the Gunky solution described below. I should also add here that in the next section, the reader will notice that **Measure-Zero** runs afoul of the closed pairs.
it seems strange to deny that open receptacles and contact are compossible.

4.2.3 Closed Pairs

Figure 4.3. A Closed Pair: Red and Green are always separated.

Call the closed versions of Redball and Greenball, ‘C-Red’ and ‘C-Green’; call the open versions ‘O-Red’ and ‘O-Green’. Because C-Red and C-Green’s receptacles both contain their boundary points, they cannot even get as close as O-Red and O-Green can without interpenetrating. In this case, C-Red and C-Green can get within any non-zero distance of one another, but never closer than that. Even if one adopts the liberal Measure-Zero definition, C-Red and C-Green still fail to be in contact; indeed, now an extended object can be placed between Redball and Greenball. As in the case of O-Red and O-Green, either one must deny the existence of C-Red and C-Green (and all their closed friends) or one must refuse contact for them.

Perhaps a further liberalization of the definition of contact will help. It is hard to imagine how this would work without interpenetration—an issue I will discuss at length in §4.4. One option is to see contact as a matter of degree. Call this Measure-Epsilon Contact: two objects are in contact to degree $\varepsilon > 0$ just in case they are $\varepsilon$ distance from one another. In this case C-Red and C-Green would be in some form of contact no matter the distance between them. A large distance would simply suggest that they would be in a very low degree of contact. This is weird. Take note: the weirdness is not as much with introducing degrees. In ordinary usage, one speaks
of degrees of contact, but such degrees are a matter how much of the total surface area of the touching bodies are in contact in some more basic and absolute sense. Furthermore, such degrees have both an upper and a lower bound on them. On the other hand, Measure-Epsilon’s degrees are not based on some basic or absolute notion of contact from which one can define the degrees. Measure-Epsilon purports to be the basic notion, in spite of its flimsiness. Worse still, these degrees have neither an upper nor a lower bound: for any degree there is both a greater and a smaller degree of contact. Measure-Epsilon Contact pushes the concept of contact to the limit of comprehension—and beyond.

Chapter 4.3

An Infinitesimal Solution

In this section, I want to offer a picture according to which space contains infinitesimal lengths or magnitudes. The primary idea behind this solution is to make the relationship between C-Red and C-Green look more like the one between O-Red and O-Green by introducing a new category of distances. In other words, I allow for a non-zero infinitesimal distance to exist between C-Red and C-Green. The idea here is that the phenomenology—if I may—of infinitesimal magnitudes is punctiform: infinitesimals behave in many ways just like points, only they enjoy the mathematical properties of extended regions. Unlike the original case of C-Red and C-Green, there is a principled bound on how far apart two closed bodies can be and still count as touching. This removes a measure of the arbitrariness in choosing how close is close enough for C-Red and C-Green to count as being in contact. Before I get ahead

---

16Intuitively, an example of a high degree of contact in the more familiar sense would be a pillow and the inner surface of the pillow case enclosing the pillow.

17Strictly speaking, zero is a lower bound, but zero is not among the degrees. Measure-Epsilon degrees approach but never reach zero.
of myself, I want to echo Kilborn that “there are important costs that come with endorsing any response to the puzzles of contact.” (Kilborn, [106], 267) I address the costs of my view directly in §4.4.3.

4.3.1 Mathematical (Re-)Orientation

In section §2.5 above, I discussed the nature of infinitesimals at length. To some extent, this section will serve to jog the reader’s memory, while offering suggestive commentary relevant to the problem of contact. Here I will offer a fairly cursory presentation of what space would be like if there were infinitesimal magnitudes in it. Call this view of space the Leibnizian perspective (not necessarily the historical Leibniz’s perspective). Let the reader recall from §2.5 that Leibniz treated circles as infinilateral polygons. In Figure 4.4, I have reproduced one of the diagrams from Chapter 2.

![Figure 4.4. Sequence of Regular Polygons Converging to a Circle](image)

Commonly, circles are defined to be the set of points within a given radius around a point. I indicated above that infinitesimal magnitudes are phenomenologically punctiform. This phenomenology is consonant with the Leibnizian perspective: each infinitesimal magnitude around the infinilateral polygon is just like a point, as in the familiar definition. The Leibnizian perspective enabled the historical Leibniz and his contemporaries to solve numerous geometric problems. Likewise, by adopting the Leibnizian perspective, I hope to enable a solution to the contact problem.

The official theory will require a sharper characterization than the one just given. I will now introduce a number system that will enable me to present a brief but precise
theory of contact in line with the framework developed in §4.1. Generically, the number system that I will use is an uncountable non-Archimedean ordered field. To do this take an uncountable ordered field like the real numbers (\(\mathbb{R}\)) and then adjoin an infinitely large number. After adding the infinite value, the ordered field axioms remain true. Recall that every member of a field has a multiplicative inverse: for any non-zero \(x\), there is some unique \(y\) such that \(x \cdot y = 1\). It follows that every infinite \(\kappa\) must have an infinitesimal value, \(\lambda\), such that \(\kappa \cdot \lambda = 1\). Within a field of this sort that contains infinite (infinitesimal) values, there are uncountably many infinite values (infinitesimals); indeed, there are infinitely many orders of infinity (infinitesimals). This latter fact requires us to carefully relativize the order of infinitesimal. With this in mind, I now offer some official definitions:

**Infinitesimal:** \(\lambda\) is infinitesimal iff \(\text{df}\) for any \(r \in \mathbb{R}\), \(|\lambda| < r\).

**Infinitesimal Relativity:** \(\lambda_1\) is infinitesimal relative to \(\lambda_2\) iff \(\text{df}\) for any \(n \in \mathbb{N}\), \(n \cdot |\lambda_1| < |\lambda_2|\).

### 4.3.2 The Theory

Above I gave a fairly general characterization of the non-Archimedean ordered fields. Here I will explicitly invoke the hyperreals (\(^*\mathbb{R}\)). Recall from §4.1 a few crucial definitions necessary to define the contact relation:

**Distance:** For \(\vec{x}, \vec{y} \in ^*\mathbb{R}\), \(d(\vec{x}, \vec{y}) = \text{df} \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}\);

**Distance between Regions:** For \(X, Y \subseteq ^*\mathbb{R}\), \(d(X, Y) = \text{the greatest lower bound of} \bigg\{d(\vec{x}, \vec{y}) : \vec{x} \in X, \vec{y} \in Y\bigg\}\).
I will also use the following concept to define contact:

**Diameter:** The diameter of a set \( X, \overline{\text{diam}}_X = \) the diameter of \( S \) for the smallest closed sphere \( S \) such that \( X \subseteq S \).\(^{22}\)

Now, with the above resources, I define contact.

**Infinitesimal Contact:** Necessarily, for any material objects \( x \) and \( y \), \( x \) is in contact with \( y \) \( \iff \) (i) there is some part \( z \) of \( x \) and some part \( w \) of \( y \); (ii) there are regions \( R_z \) and \( R_w \) such that \( z \) and \( w \) have them as receptacles respectively; (iii) \( d(R_z, R_w) \) is infinitesimal relative to the minimum of \( \{\overline{\text{diam}}_{R_z}, \overline{\text{diam}}_{R_w}\} \); and (iv) \( z \neq w \).

Let me unpack the various aspects of the definition and what role they play—(i) and (ii) are simply set-up. With clause (iv), I follow Hud Hudson in [94] specifically in avoiding the degeneracy of having all objects be in contact with themselves. I do not mind cases where an object is in contact with itself, as in the case of an open U-shaped body with a single point separating the two sides. The idea is that if a single object touches itself, there should be two distinct parts that come into contact. The cost of this is that no simple—an object with no proper parts—can be in contact with itself.\(^{23}\)

Clause (iii) is where a lot of the unique action is found.\(^{24}\) To begin, note that a zero distance counts as infinitesimal. This is achieved if at least one of the objects is missing its boundary point at the point of contact. If both objects have their boundary points in the relevant place, then one must use non-zero infinitesimals.

---

\(^{22}\)One familiar with the hyperreals might be anxious at this point: are the above notions well-defined? For example, the set of infinitesimals does not have a least upper bound. Similarly, the notion of a diameter depends upon there being a least closed sphere; there is no such thing for the set of infinitesimals. The simplest way to avoid this problem is to develop the hyperreals from within Edward Nelson’s *Internal Set Theory* (IST), a conservative extension of Zermelo-Fraenkel Set Theory with Choice (ZFC). Within (IST), the problematic sets do not even exist. If one wants to stay within the framework developed in Chapter 2, then one should restrict the sets of the theory to first-order definable sets. All first-order definable sets satisfy the above definitions.

\(^{23}\)If one finds the idea of a simple in contact with itself more plausible than all objects being in contact by necessity, then he can simply remove (iv) without doing serious damage to the theory.

\(^{24}\)I must acknowledge that a lot hinges on the metrical aspects of space rather than exclusively topological aspects. Topological concepts can be defined from metrical ones, but this theory is distinguished in part by providing an answer that requires explicit reference to a metric.
How does this work? Look at the smallest of the two bodies coming into contact and ensure that the distance between them is infinitesimal relative to it. As I mentioned above, I am trying to make closed pairs look like open pairs. To do this, I exploit the smallness of the infinitesimally extended distance to fashion something that one might call *phenomenologically punctiform*—as far as anyone can tell, it is a point. These infinitesimals are *per impossibile* extended points. As long as some bodies get infinitesimally close, then they are in contact.

**Chapter 4.4**

*Desiderata*

Before examining any of the solutions provided in the literature, I will explicitly say what I want. In each case, I will offer some arguments for what I want, but I admit that not everyone will be pleased and not everyone will be convinced. Below, I consider an annotated cluster of theoretical *desiderata*. I prioritize the first three *desiderata* over the others. These are my *First Priorities*. One of my aims in this chapter is to see if a theory of contact can be developed that satisfies them first. The second three I provide are also of interest. If I could have them too, I would. The reader will see that I cannot have all of them.

### 4.4.1 First Priorities

**CONTINUITY:** *Space is continuous*—Space is paradigmatically continuous.25 Continuity’s major contender for the structure of space is discreteness.26 Discrete sys-

---

25 In “Is Space-Time Discrete or Continuous? – An Empirical Question,” [69], Peter Forrest offers a nice discussion of this issue. Due to some fairly subtle issues of theoretical economy, he argues that discrete space is at least not implausible.

26 There are models of space of intermediate strength—weaker than continuous but stronger than discrete—but they are rarely if ever considered. In “Models and Reality,” [159], Hilary Putnam discusses the role of rational space-time, but his concerns there are of a very different nature than mine here. Furthermore, the main argument advanced against discrete space can be effectively
tems are those modeled by countable number systems that allow for successors and predecessors, i.e. the integers, $(\mathbb{Z})$.\(^{27}\) Although there are number systems that are equinumerous with the integers, like the rationals, they fail to be discrete in the most important way owing to their density; that is, the rational numbers are such that for any $x$ and $z$ with $x < z$, there is some $y$ such that $x < y < z$—so there is no “next” point in space (or, concerning time, no “next” moment in time.) Put another way, the density property allows for infinite divisibility,\(^{28}\) the very property that champions of discrete space reject. A discrete world would be something like a three-dimensional version of the two-dimensional television screen: as the television’s individual pixels can be lit up with different colors to create a single image, so too each node in a three-dimensional spatial array can be occupied by the various parts of material bodies.

What’s the problem? The most serious problem for a discrete picture is that one must abandon very simple and fundamental theorems of geometry. For a straightforward example, one must abandon the Pythagorean theorem. Take for example a triangle with legs of length 3 and 4 (using the basic building blocks of space). According to the Pythagorean Theorem, this triangle should have hypotenuse of length 5. Consider Figure 4.5. below.

---

\(^{27}\)In Forrest’s “Is Space-Time Discrete or Continuous?” [69], he suggests that the mark of discreteness in space is that a notion of adjacency in space gains purchase: when looking at the basic building blocks of space, there is something that comes next in line. I capture this intuition using the notion of successors (or predecessors).

\(^{28}\)Recall from §2.1: Infinite Divisibility—An object $X$ is infinitely divisible just in case it is extended and any extended part of $X$ has at least two disjoint (non-overlapping) extended parts.
Figure 4.5. Discrete 3-4-5 Triangle.

First, this figure is strictly speaking not a triangle at all: it is a Lego-like figure that resembles a cross-section of an Mesopotamian ziggurat. It is also not at all clear how one should go about measuring the length of the hypotenuse. For example, one might choose to count all those squares through which the actual hypotenuse cuts: six. If one were to only count those blocks along the right edge, then there are four—blocks \( A, C, E \) and \( F \). Perhaps instead one might exclude only Box \( E \) from the six, since too little goes through it. This would deliver the requisite five blocks. However, why Box \( E \), but not Box \( C \), since not much goes through \( C \) either? In any case, even with a hypotenuse with an integral value, there are problems—let alone those with irrational values! Upon admitting irrational hypotenuses, no clever counting strategy will deliver the appropriate length. Furthermore, the Pythagorean theorem is a fundamental result in basic planar geometry. If one cannot have this, very little else geometrically will be available to us in a discrete world.

Generally, one should avoid tying the hands of progress. If a principle has extraordinary but fruitful consequences, then that is in its favor. Suppose that we can provide an adequate science of discrete space, comparable in purpose to Euclidean geometry. I remain hesitant due to the theoretical maxim of Minimum Mutilation, also known as Familiarity of Principle. One follows Minimum Mutilation to avoid the kinds of catastrophes resulting from a conceptual overhaul.\(^\text{29}\) In any case, a shift

\(^{29}\text{For more, see Quine’s Word and Object, [162], §5.}\)
to discreteness from continuity would require a profound shift in the mathematics employed for the last two and a half millenia to describe space. One would have to re-engineer from the ground up the mathematical tools employed in contemporary physics and more if one chose such a path. As far as I can tell from my limited vantage point, this weakness is too great to accept for the sake of simplifying our account of contact (and similar puzzles).\textsuperscript{30}

TOTAL CONTACT: \textit{No matter their topological properties, it is possible that any two bodies can touch}—Bernard Bolzano developed a view of touching that required the two bodies’ receptacles to be connected; that is, at least one shared boundary point is contained in at least one of the receptacles. Famously, Franz Brentano called this view monstrous.\textsuperscript{31} Although I would not call the Bolzanian view ‘monstrous’, it strikes me as a mark of theoretical deficiency to refuse contact to certain classes of bodies when the reason is strictly topological and not motivated by further physical reasons (mechanical, electrochemical or otherwise). More explicitly, when two things cannot touch, one should not blame the shape of the objects for this problem. These facts play out in “Could Extended Objects Be Made Out of Simple Parts?” where

\textsuperscript{30}A worry might have emerged: discrete systems use the familiar, everyday integers to understand space. The present theory is wildly extravagant: uncountably-many infinitesimals, infinite-levels of infinity. \textit{How is this view of space better than discrete space? At least, isn’t this system of infinitesimals also guilty of violating Minimum Mutilation?} As fair as this charge is, I aim to defuse it with a pair of remarks. First, although the integers are familiar \textit{simpliciter}, they are anything but familiar in spatial contexts (as argued above). If one adopts a discrete space, many theoretical sacrifices must be made in the most basic mathematics and, consequently, mechanics. The feeling of familiarity one gets from the integers is chimerical. Second, the present view is not as much of a mutilation as it seems. The most familiar example of a non-Archimedean field is provided by the hyperreals \((^\ast\mathbb{R})\)—discussed in detail in \S 2.5.3. The accompanying first-order statements about the hyperreals are actually \textit{conservative} over first-order statements of the reals (in more precise terms, the \textit{theory} of hyperreals is conservative over the \textit{theory} of the reals). In other words, for any statement \(\varphi\) regarding the reals, if the theory of hyperreals proves \(\varphi\), then the theory of the reals does too. Conversely, anything that one might have wanted to say before is still available here. The way in which the new theory goes beyond the original only increases our conceptual and computational command. Many proofs involving limit concepts can be shortened using the resources made available by the theory of the hyperreals. In effect, the theory of infinitesimals does not overturn or replace one’s original understanding; rather, it enhances and improves it.

\textsuperscript{31}See \textit{Philosophical Investigations on Space, Time and the Continuum}, [24], p. 146.
Zimmerman suggests that, in the absence of an explanation of these gaps between the open pair or between the closed pair, one must posit the existence of forces that repel the two objects from one another.\textsuperscript{32} The nature of these forces is obscure; nevertheless, one cannot help but notice that certain bodies refuse to get past a certain distance and that shape alone is not well-suited to explain this phenomenon. There is something peculiar about being required by reason alone to introduce otherwise occult \textit{forces} because of an essentially geometric mystery. This \textit{desideratum} is simply the requirement that, \textit{in the absence of physical interference}, two bodies ought to be able to touch.

\textbf{NO INTERPENETRATION:} \textit{Interpenetration between material objects is impossible—} First, I begin with the fact that many of our pre-philosophical selves balk at the possibility of two distinct material bodies occupying the exact same space at the same time: a truism as some put it.\textsuperscript{33} Intuitions can be surrendered—no doubt—but if they can be maintained, all the better. At root, ordinary everyday bodies seem to resist occupying the same space at the same time. Any attempt to get the 8-ball and the cue ball to fill the same space would result in destroying at least one of them. A more theoretical way of capturing this intuition is by insisting on the following principle:

\textbf{One-Place-One-Object:} For any point $p$, if material objects $M$ and $N$ occupy $p$, then $M$ overlaps $N$.\textsuperscript{34}

To hold the principle of \textbf{One-Place-One-Object} is to extend this intuition that only one material body can exactly occupy a given region at a given time.\textsuperscript{35}

\textsuperscript{32}Although Zimmerman does not say as much, I suspect that he has the Principle of Sufficient Reason in mind when he says that there must be these repulsive forces. In “Simply Possible,” Theodore Sider challenges Zimmerman on the existence of these forces.

\textsuperscript{33}See Wiggins’ “On Being in the Same Place at the Same Time,” [219].

\textsuperscript{34}The occupation relation here is \textit{not} exact occupation. $M$ or $N$ might exactly occupy $p$, but it is not required. $M$ or $N$ might spill over $p$. Thanks to Ben Caplan for pressing the need for a stronger principle.

\textsuperscript{35}A weaker principle that can be derived from this one is where one holds that the receptacle
The reader might be tempted to dismiss no interpenetration, because it resembles more familiar cases from the literature of two distinct objects occupying the same space and thereby violating One-Place-One-Object. For example, consider a cat, Tibbles. Like most cats, Tibbles has a tail at time \(t_0\). Unfortunately, owing to unforeseen circumstances, Tibbles loses her tail at some time \(t > t_0\). At \(t_0\), consider the proper part of Tibbles that is Tibbles without a tail—called Tib. At time \(t_0\), Tib and Tibbles are distinct and exactly occupy distinct spaces, but at time \(t\), Tibbles and Tib exactly occupy the same space. Furthermore, given that Tib never had a tail and Tibbles lost her tail, it seems that Tibbles and Tib are distinct. Returning to concerns of interpenetration, the thought is, *If Tibbles and Tib can both exactly occupy a region of space, what’s wrong with interpenetration?*

In the above case, Tibbles and Tib are composed of exactly the same matter. If these objects are genuinely distinct, they are at least materially indistinguishable. The relation is a one-one function. Explicitly, if \(R, S\) are receptacles of \(x\) then \(R = S\) and if \(x, y\) have receptacle \(R\), then \(x = y\). Recall that the receptacle relation is the exact location relation on my reading.

In *Paradoxes of the Infinite*, Bolzano writes the following:

> I deny categorically that any such interpenetration [between substances] be possible; since, so far as I can see, it is contained in the very idea of a simple place (or point) that it be a place capable of affording ubication [position] only to a single (simple) substance. Wherever we have two atoms, there do we have two places. (Bolzano, [17], 158, emphasis original)

The final statement suggests that Bolzano held some principle comparable to One-Place-One-Object.

36See, for example, Karen Bennett’s [15], Allan Gibbard’s [75], Michael C. Rea’s [165] and [164], Theodore Sider’s [188], Ch. 5, and David Wiggins’ [219] among many others.

37The existence of Tib is guaranteed by the *Doctrine of Arbitrary Undetached Parts* (DAUP); for every subregion \(S\) of a body \(M\)’s receptacle \(R\), there is a part of \(M\) with \(S\) as a receptacle. In order to avoid DAUP, one might set the problem up by thinking of Tibbles as the fusion of her head, legs, trunk and tail; Tib is just the fusion of the head, legs and trunk (both guaranteed to exist by *Finite Fusions*). If someone is bothered by the metaphysical presuppositions behind the setup of the puzzle, there are many other comparable examples. For a conceptual framework for engaging these puzzles and numerous examples, see Rea’s, “The Problem of Material Constitution,” [164].

38For some, Tibbles is identical to Tib at time \(t\) despite their separate histories because Tibbles and Tib exactly occupy the same space and have the same intrinsic properties. I am not interested in settling problems of material constitution—difficulties that arise from the relationship between a material object and the matter of which it is constituted—but merely in distinguishing that problem from the present one.
case of interpenetration is different, though. For two distinct bodies to interpenetrate, there is some part of Red and some distinct part of Green that occupy (possibly non-exactly) the same region. Unless one of them is annihilated, there will be some point of space \( p \) such that two bodies (possibly non-exactly) occupy \( p \). Unlike Tibbles and Tib, Red and Green are *materially* distinguishable. Suppose one were to restrict the quantifiers of **One-Place-One-Object** to just the matter shared by Tibbles and Tib—call it pure matter—and allowing one to respect the difference between Tibbles and Tib:

**One-Place-One-Object***: For any point \( p \), if pure material objects \( M \) and \( N \) occupy \( p \), then \( M \) overlaps \( N \).

Even on this adjustment, Red and Green cannot interpenetrate. The point of all this is to show that interpenetration is not simply just another case like Tibbles and Tib.\(^{39}\)

There are two more arguments for **NO INTERPENETRATION** but these are specific to problems for contact. One of these arguments is found in William Kilborn’s “Contact and Continuity,” [106], described below in §4.5.3. The other I will provide now. Within a mereologically complex self-connected body, one commonly says that those parts that share boundary points are in contact. If one allows for two bodies to interpenetrate, then it is possible for them to be closer than the “innards” of a single self-connected whole. The motivating idea is that contact should be no closer than what is had by parts of a single whole.\(^{40}\) For example, two closed hemispheres that are touching via interpenetration are closer to one another than two halves of a single whole.

\(^{39}\)There are difficult questions near the surface that I hope to avoid, but one thing is clear: the *matter* composing Tibbles and Tib satisfies **One-Place-One-Object**. If one wants to prevent this principle from forcing ordinary objects like Tibbles and Tib to be identical, one might need to develop a view of ordinary objects that involves more than matter. Illuminating discussion of this is found in Kit Fine’s “Things and Their Parts,” [66], and “The Non-Identity of a Material Thing and Its Matter,” [67].

\(^{40}\)This idea is the dual to what I call **SUPREME CLOSENESS** below from Zimmerman’s [225]. The structural similarity between these two principles certainly favors taking both at once. However, I will argue later why we might want to forsake **SUPREME CLOSENESS** even while embracing the present idea. The asymmetry may appear unprincipled for now, but is motivated in §5.3.
closed sphere: they are virtually indistinguishable except that the former contains a
circular planar region of space each point of which has double matter for the price of
one. All told, it strikes me as natural that the relationship between a self-connected
body ought to be the upper-bound on the proximity between two bodies in contact.

4.4.2 Second Priorities

SUPREME CLOSENESS: ”If two objects are in contact, then it is impossible for two dis-
tinct non-overlapping objects to be closer together than the two objects in question”—
This principle is an exact quote from Zimmerman.41 According to this, if two bodies
are in contact, then nothing could be closer.42 Zimmerman lays this principle down
as a presupposition about contact: I view it as a mere desideratum—as with all the
others.43 The advantage to this requirement has come up before: if one ignores this
desideratum, then one becomes vulnerable to the possibility of ascribing contact to a
pair and yet be able to slide Skinny—a body of less than three-dimensions—between
them.44

SIMPLE INDEPENDENCE: Neither Simples nor Gunk is ruled out; nor is either
of them forced upon us—Recall that MM+ does not settle Simples or Gunk; neither
does any stronger mereological system that does not adopt either of them (at least
implicitly). If I can avoid commenting on these mereological theses while settling
issues of contact, that is theoretically desirable. The less constrained one is in formu-
lating a solution, the better.

41Zimmerman, [225], p. 9
42This is very similar to the final remarks in NO INTERPENETRATION. It suggests that the closeness
of a single mereologically complex self-connected body should be the lower-bound on the proximity
between touching bodies.
43Below I will show that this is the very desideratum that I forfeit. It is desirable, yes, but in the
end I surrender this desideratum at the cost of satisfying others.
44I forfeit this desideratum. I also argue that Zimmerman’s attempt to dodge this problem creates
problems for him elsewhere. This is discussed in §4.5.2.
RECEPTACLE INDEPENDENCE: The Liberal View of Receptacles is neither ruled out, nor forced upon us—The Liberal View of Receptacles states that the points in absolutely any nonempty class of points compose a receptacle. This view is counter-intuitive in certain ways, but provides serious ideological simplicity. It streamlines naturally with ordinary point-set topology and allows one to accept Unrestricted Composition—the view according to which for any things, there is a fusion.\footnote{To some, the latter is a deficiency. For an excellent survey of related issues, see Markosian’s “Restricted Composition,” [130]. As the title suggests, Markosian himself ends up ruling against Unrestricted Composition.} As in the immediately preceding desideratum, it is ideal to avoid over-committing oneself to various metaphysical theses in developing a theory of contact. Nothing \textit{prima facie} about contact seems to dictate that some regions could be receptacles or not.

* * * * * *

Each of the latter three principles is a second priority. I mention them as desirable but will not commit myself to their fulfillment later. Below, the reader will see that quite a bit of tension surfaces between RECEPTACLE INDEPENDENCE and TOTAL CONTACT, the latter of which I will favor.

4.4.3 Standing Up to the Desiderata

Let us now consider step-by-step how \textbf{Infinitesimal Contact} fares on each of our \textit{desiderata}:

CONTINUITY: Whether the hyperreals form a continuum is developed at length in chapter 2 above, specifically §2.5. The hyperreals have been received within the mathematical community as providing a theoretically adequate but distinct understanding of continuity. In particular, the hyperreals satisfy a long-standing intuition that a
continuum is composed of infinitesimal magnitudes.

TOTAL CONTACT: At present, the standard trio of pairs of solid regions are capable of being in contact. Returning to Redball and Greenball, note that (as before) the open-closed pair forms a self-connected “peanut.” Also, O-Red and O-Green still have only one point between them. Although the formal situation is still the same, now C-Red and C-Green are counted as in contact when they are an infinitesimal distance apart. I must admit at this point that formally speaking, for any distance between the two closed spheres, they can get closer. However, the situation is better than Measure-Epsilon Contact. First, there is a nonarbitrary criterion by which to distinguish distances, thereby eliminating the possibility of treating any distance as some degree of contact. Now there is a sharply defined concept of what is “close enough”: infinitesimally close is close enough. Furthermore, one is merely closing an infinitesimal distance by infinitesimal increments. To return to our picturesque language, the space between these two objects would be qualitatively or phenomenologically punctiform: the distance is so small that no finitely-sized body could interpose it.46

Note that provided that one accepts the existence of point-sized simples, there is still successful contact across the range of peculiar cases suggested by Hud Hudson in his “Touching,” [92]. This view is discussed in greater detail below in §4.5.4. For present purposes, Hudson’s theory allows for scattered arrays of point-simples (gritty

46 One might be thinking, What’s all this talk of phenomenological punctiformity? If infinitesimals are relative, then this phenomenological ascription is inappropriate, since it presupposes an absolute notion. There are two possible ways out of this. The simplest way is to think of all this machinery in a measurement-theoretic way. On that view, find some ordinary body of interest (a baseball, a mountain, a star cluster) in the relevant world and fix it as having finite measurements. Such a body is a reference point for an effective representation. Of course, maybe the world under consideration is too exotic to have ordinary bodies. In that world, maybe bodies are more like sets of points in mathematical space and nothing more. In this latter case, the values of the number system placing the bodies in some constellation of locations can be treated as absolute and not imposed by us from the outside as a description of the arrangement of bodies. From that perspective, whatever values are finite will be treated absolutely.
bodies) to be in contact through Zeno-like sequential convergence. **Infinitesimal Contact** actually allows for even more types of gritty objects to be in contact provided that each body is within a (relatively) infinitesimal distance of the other. In other words, none of the progress made by Hudson in his presentation of contact is lost by **Infinitesimal Contact**. This is noteworthy because Hudson’s theory of contact is tailor-made for the *Liberal View of Receptacles*—the major thesis at odds with **total contact**. Unfortunately, two points cannot come into contact because there is no sense to be made of infinitesimal relativity for them. At present, no extant view (including Hudson’s) allows points-only contact. Therefore, *Even if one accepts the* Liberal View of Receptacles, **Infinitesimal Contact** *is no worse off than the next.*

**NO INTERPENETRATION:** **Infinitesimal Contact** neither *contradicts* nor *requires* interpenetration.\(^\text{47}\) Since I take interpenetration to be impossible and **Infinitesimal Contact** does not require it, this strikes me as a supreme advantage.

**SUPREME CLOSENESS:** To begin, I admit that this *desideratum* is not satisfied. In what specific ways for this theory is there a problem? First, if infinitesimally extended bodies exist, **Infinitesimal Contact** allows one to slide an *extended* object between them.\(^\text{48}\) Secondly, for **Infinitesimal Contact**, open-closed pairs of objects can get *closer* than open pairs or closed pairs. Similarly, the closeness that obtains between adjacent parts of a single connected material object is greater than the closeness of certain things that are certified as in contact. A final problem to be described in greater detail below involves a failure of contact between the respective parts of two touching bodies.

\(^{47}\)The Bite-the-Bullet view, discussed in §4.5.1, is one such view.

\(^{48}\)This might be reason to decide the status of the *Liberal View of Receptacles* and not allow for extended objects of infinitesimal volume or dimension.
I think the first problem is indeed strange. The only conciliatory remark I can offer on this point is that although these objects are extended, they are only infinitesimally so. Although, formally speaking, these are on a par with any three-dimensional hunk of space, phenomenologically, these objects as more like planes, lines or points.\textsuperscript{49} If this is unsatisfying, maybe this is my bullet to bite.

On the second point, I want to draw attention to the fact that for some philosophers, this is an advantage. I want to suggest that the requirements for a pre-theoretic notion of contact fall short of the full self-connectedness found in solid wholes. Often contact is thought of as a relation between two recognizably distinct bodies; by contrast, self-connection is treated as the extremum when those two bodies are no longer distinct but form a unified whole. If this is so, then self-connectedness is a kind of maximal contact that is enjoyed by a unified whole and not required of other forms of contact.\textsuperscript{50}

The final worry is best explained using a concrete case. Recall C-Red and C-Green, colored balls with closed receptacles. Using the variables from \textit{Infinitesimal Contact} above, instantiate \(x\) with C-Red, and \(y\) with C-Green. Suppose that the relevant part \(z\) of \(x\) (=C-Red) has some closed, extended part \(q\) that is infinitesimal relative to \(z\). Furthermore, suppose that \(w\), the relevant part of \(y\) (=C-Green), is not infinitesimal relative to \(z\). Viewing \(q\) in isolation, it follows that \(q\) is not touching C-Green. This is peculiar, since one does not ordinarily think of objects falling in and out of contact on the basis of their parts.\textsuperscript{51}

Although this is strange, to view infinitesimally extended bodies as phenomenologically punctiform provides some help here again. Any view that allows for gaps of any kind in touching will deliver a comparably strange result where one takes \(q\) to

\textsuperscript{49}Whether we count an extended object as a plane, line or point will depend upon the number of dimensions in which the body is infinitesimally extended.

\textsuperscript{50}For related discussions see van Inwagen, [204], Chapter 3 and Casati and Varzi, [36], Ch. 4.

\textsuperscript{51}Ben Caplan has noted a similar but distinct problem: if one removed the remainder of C-Red from \(q\), then \(q\) is no longer in contact with C-Green.
be a point rather than some infinitesimally extended body. If one examines points of O-Red or O-Green, the same problem emerges. In a way, I have not answered this problem. However, I have tried to show that the problem is not unique to my view. Anyone whose view does not satisfy SUPREME CLOSENESS will have this problem in one form or another.\footnote{One way out of this is comparable to Zimmerman’s approach discussed at length below in §4.5.2. He denies that there are point-sized bodies. Analogously, I may deny that there are infinitesimally-sized bodies. All of this goes to show that there is nothing distinctively problematic about infinitesimals.}

**SIMPLE INDEPENDENCE:** Nothing about **Infinitesimal Contact** forces one to adopt a view on simples. There are some traditional arguments that suggest that if two objects are in contact at a point (as in the case of our peanut spheres), then the point of contact must be a part.\footnote{See Zimmerman’s discussion of Suarez in “Indivisible Parts and Extended Objects: Some Philosophical Episodes in Topology’s Prehistory,” [226].} The thoughts that motivate this idea implicitly rely on some version of DAUP. The fact that they are in contact is not really doing the work; rather, it is the newly recognized (from the contact) point-sized subregion of a body’s receptacle that encourages one to believe there is a point-sized part. Only if one accepts DAUP should one be forced to recognize the punctate parts of some body corresponding to the various points in its receptacle. If $x$ and $y$ are in contact, then there is indeed *some* part of each of $x$ and $y$ that is contact with the other—*Reflexivity* guarantees as much! It could turn out that there are point-simples, but not because of contact. This goes for any view of contact, in spite of the tradition suggesting otherwise.

**RECEPTACLE INDEPENDENCE:** Many see the need to limit the kinds of regions that an object can occupy because of issues of contact. For example, using the contact problem, Zimmerman argues that points cannot be receptacles.\footnote{Others expand this argument is developed further below. To engage this problem is very closely related to}
the number of receptacle-types because of contact. As mentioned in the discussion of SIMPLE INDEPENDENCE, points must be receptacles since a point is cited in the definition of contact.

By now it should be clear that these different desiderata constrain one another. As such, either limiting or expanding the types of receptacles have been answered to some extent already above. I answered the worry about limiting the receptacle types in the discussion of TOTAL CONTACT. The more shapes and sizes there are to accommodate, the more difficult it becomes to guarantee contact. Regardless, the considerations provided above show that we have good reason to find this picture of contact as compatible with the Liberal View as the next. Someone who is a firm believer in the Liberal View might be drawn to Infinitesimal Contact since all views on offer for such a person already give up TOTAL CONTACT. Personally, I would sooner surrender the Liberal View of Receptacles and with RECEPTACLE INDEPENDENCE than TOTAL CONTACT. For more on this, see §4.6. The only extant argument in favor of inflating the kinds of available receptacles for contact was found in the very last discussion about SIMPLE INDEPENDENCE. I argued there nothing about contact itself requires commitment to point-sized receptacles; rather, a surreptitious use of DAUP is in play.

Chapter 4.5

Alternative Solutions

In this section, I will provide a brief survey of a few extant solutions to the the puzzle given above found in the literature. In each case, I try to point out some of the difficulties remaining for that view and ultimately why I do not take that approach.
4.5.1 Smith’s Bite-the-Bullet View

The first kind of solution is one where one simply “bites the bullet” and violates both no interpenetration and supreme closeness. On this view, contact occurs between two bodies so long as they share a boundary point (whether or not that boundary point is contained in the receptacle of either object). Between open pairs, there is a non-extended spatial gap between them; between closed pairs, there is a non-extended spatial glut (interpenetration); open-closed pairs are self-connected as before. This is called the Bite-the-Bullet View because it does not resolve any of the mysteries described in §4.2, but simply embraces them. This means that any puzzle developed above for contact regarding gaps and interpenetration immediately reassert themselves as objections. Nevertheless, this view does have the advantage of satisfying total contact.

There are two problems for this view that I want to draw out explicitly. One problem for this view that is unique to it is that it embraces both spatial gaps and spatial gluts. Many of the other views reject at least one them. In the absence of further systematic encouragement to accept this view, I take these disadvantages as sufficient motivation to look elsewhere. Indeed, from my point of view, this solution’s major philosophical appeal is its greatest weakness in the present dialectical context—a point to which I will now turn.

The Bite-the-Bullet view is most seriously defended by Sheldon Smith in his “Continuous Bodies, Impenetrability, and Contact Interactions: The View from the Ap-

55 As a solution that embraces gluts but eschews gaps, consider the analytic versions of Brentani-anism found in both Roderick Chisholm and Barry Smith’s work. I cannot go into much detail here because the framework proposed for dealing with boundaries, contact and so forth is rather alien to the overall discussion. The idea for them, though, is that boundaries are *sui generis* entities that are dependent upon the object(s) of which they are parts. This view takes interpenetration and contact as given and elaborates a very complex theory involving modal operators and more. I mention this view for the purposes of exhaustiveness, but cannot say much about it given that it deviates too significantly from the core conceptual picture provided in §4.1.
plied Mathematics of Continuum Mechanics,” [194]. There he employs the mathematical models beneath continuum mechanics to provide an answer to the contact problem. Continuum mechanics is a mathematical theory that models the mechanical behavior of continuous physical phenomena; fluid dynamics, static forces acting on continuous surfaces and, of course, contact and collisions interactions are all subsumed under this branch of mathematical applications. Smith’s philosophical method is in many ways different from the majority of the contenders in the debate over how to understand contact. I must make clear that Smith is conscious of this methodological divide: it is at the center of his arguments. He is challenging metaphysicians to pay closer attention to how contact is treated in contemporary physical theory—a fair request. If one is attempting to provide a way of theoretically managing real matter in the actual world, I think Smith’s approach is without equal. This is a well-developed mathematical model of the interactions of material bodies in continuous space with the highest empirical pedigree. In this regard, it has the edge on every other solution.

So what’s the worry? If one restricts oneself to the actual world, then the contact problem—at least, as presented here—vanishes on two scores. First, the theory of continuum mechanics takes contact for granted—as a datum of the physical world—and uses mathematical tools to describe it. To look to continuum mechanics for answers is to say there never was a problem. Smith follows a different path explaining how contact and the related concepts of body, boundary and impenetrability are deployed within state-of-the-art mathematical physics. Second, actual material bodies are not continuous: they are scattered arrays of molecules. In continuum mechanics, the

56 Along these lines, Smith writes:

In reality, the root argument [or, contact problem] starts with an overly rigid notion of impenetrability that is rejected in continuum mechanics, and it continues with an a priori notion of contact that is not relevant to the physics of continuous media. (Smith, [194], 535, emph. added)

Smith might construe the contact problem so that the puzzle amounts to how contact occurs rather than whether contact occurs. For my part, I am still upstream from Smith trying to figure out the question of whether first.
bodies are modeled as if they were continuous. Since the relevant data are actually silent on whether bodies can interpenetrate along the boundary, physicists do not begrudge violations of NO INTERPENETRATION. More specifically, Smith challenges **One-Place-One-Object** on the basis that it is casually violated within the theory of continuum mechanics. My concern is that interpenetration is acceptable only because we are treating the bodies as if they were continuous. To what does all this point? Since real bodies are not in fact continuous, asking in the pragmatic mood what to say when treating actual bodies as if they were continuous does not seem to illuminate what one would say when treating truly continuous bodies. If the actual world contains no truly continuous bodies, then one must look to other worlds for conceptual therapy. I want to know, “Could two closed bodies touch?”—even if nothing in the actual world is closed.

---

57 Within continuum mechanics, there is a companion principle to **One-Place-One-Object**, the Axiom of Impenetrability, which cannot be violated excepting sets of zero measure. Interpenetrations along sets of measure zero (one such set is any subset along the boundary of a material body) are treated as negligible.

58 Thanks to Nicholaos Jones for pointing out that there are alternative ways of understanding these theories regardless of Smith’s own perspective on them. In “How We Dapple the World,” [199], Paul Teller argues that theories like continuum mechanics are to be treated on a par with quantum theories, since they both involve a certain amount of idealization. In contrast with other philosophers of science, he argues that the truth is not a matter of reading the facts right off the consequences of the most foundational scientific theory—those theories that purport to have the widest applications. He encourages his readers to move away from the image of scientific theories as true or false simpliciter to an image where all (successful) scientific theories are “fallible veracities,” independent of their scope of application. Further discussions of these issues are found in Nancy Cartwright’s “Fundamentalism vs. the Patchwork of Laws,” [34] and another article by Sheldon Smith, “Models and the Unity of Classical Physics: Nancy Cartwright’s Dappled World,” [193].

59 I sincerely struggle to interpret the following:

However, continuum mechanics strays as little as possible from the features of actual matter while still remaining in the context of the root argument. If one wants to know what impenetrability and contact action would amount to were bodies continuous, it seems to me that there is no better place to look than modern continuum mechanics.

(Smith, [194], 509)

Given the other disparaging remarks he offers on the root argument (‘contact problem,’ in my terms) throughout, this remark seems surprisingly conciliatory. Perhaps one should read “the context of the root argument” to mean that the relevant concepts are in still play, but they must stay closer to home.

60 Cf. the remark from Hawthorne as quoted in the Prologue above.

61 To see that Smith is not as interested in such a question, consider:

However, if [impenetrability] is the only property [these bodies] have, it is not clear
4.5.2 Gunky Solution

Zimmerman engages the contact problem in his essay “Could Extended Objects Be Made Out of Simple Parts?” [225]. In this paper, Zimmerman argues primarily for a mereological thesis: at least for the denial of Simples—in particular for the denial of the thesis that all objects are composed of point-sized simples—if not for Gunk. In order to do this, he employs a significant amount of topological apparatus, including a version of the puzzle above. The scapegoat for him is the existence of point-simples. A few passages indicate that Zimmerman prefers a Whitehead-Tarski view, according to which one takes open spheres to be fundamental and constructs points as the limit of nested concentric spheres. In other words, textual evidence suggests that Zimmerman’s favored view is gunky material bodies with three-dimensional, open receptacles.

In fairness, Zimmerman never offers a theory of contact. However, given other things he says, a natural choice for understanding contact would be Measure-Zero Contact from §4.2.2: two objects are in contact provided that there is no extended

---

62 In “Grit or Gunk: Implications of the Tarski-Banach Theorem,” [70], Peter Forrest gives reason to support Gunk, but for reasons more or less orthogonal to the contact debate. He does not present a solution to the contact problem, but rather offers independent reason to adopt for this particular solution. Essentially, he argues against the thesis of Simples (and that simples are points). Simples allows one to construct Tarski-Banach style paradoxes that violate basic intuitions about how measurability should work. The Tarski-Banach paradox is that if the Axiom of Choice is true, then one can divide up a solid sphere finitely many times, perform finitely many rigid motions to the pieces and then end up with two spheres of exactly the same size as the initial sphere. Forrest argues that the alternative is to adopt either a gunky view, or a gritty view (the latter is one where every object is simply the fusion of finitely many spatially scattered point-simples). He suggests that the Grit thesis might have the benefit of empirical support but that his interest is only in settling what is metaphysically possible as opposed what is actually the case.

63 The definition of ‘open’ depends on points, viz. every point of an open set X is an interior point of X. There are ways to elucidate these topological notions without appealing to points as in Tarski’s “Foundations of the Geometry of Solids,” [198].

64 See, in particular, Whitehead’s “The Anatomy of Some Scientific Ideas,” [216].

65 These remarks are found in the final portions of §6 in [225].
distance between them. Now that every thing is extended in three-dimensions, it is no longer possible to slip Skinny between the two open bodies in contact. Therefore, this solution avoids a central difficulty articulated in §4.2.2 about open pairs.66

My worry is that Zimmerman’s favored view is subject to arguments very similar to some that he himself offers. Zimmerman argues against open receptacles when combined with the claim that everything has a point-simple as a part. Unfortunately, with other principles Zimmerman accepts, problems emerge (described in detail below) analogous to those articulated by Zimmerman himself in his attack on point-simples. Below, I will make use of one of Zimmerman’s primary assumptions: a variant of the Doctrine of Arbitrary Undetached Parts (DAUP).67 Call Zimmerman’s variant ‘Z-DAUP.’ Z-DAUP states that any extended material body’s receptacle divides into a left-side and a right-side, each of which is the receptacle of an extended material body.68

In order to reveal some of the difficulties for open, gunky worlds, consider O-Red as before: an extended, colored body in an open world. By Z-DAUP, there is a distinct left part (L-Hemi) and a distinct right part (R-Hemi) of O-Red. Recall that L- and R-Hemi are both open because every receptacle is open. Define Redball− to be the fusion of the L- and R-Hemi, the existence of which is guaranteed by Finite Fusion. Now, clearly Redball− is a part of O-Red. Furthermore, they are distinct because Redball− does not occupy this central disc region but O-Red does. So, Redball− is a proper part of O-Red. By Weak Supplementation, there is something else, viz.

---

66 Zimmerman is also sympathetic to certain kinds of closed metaphysics, viz. those where boundaries are dependent entities. A contemporary example of this includes Barry Smith’s presentation of Brentano-style concepts of bodies and boundaries found in “Boundaries: An Essay in Mereotopology,” [192] among others.

67 DAUP, also sometimes called ‘Arbitrary Partition’ or ‘Geometric Correspondence Principle,’ states that for every subregion S of a body M’s receptacle R, there is a part of M with S as a receptacle. A variety of critiques are found in Josh Parson’s “Theories of Location,” [147], Peter Simons’ “Extended Simples: A Third Way Between Atoms and Gunk,” [191] and van Inwagen’s locus classicus “The doctrine of arbitrary undetached parts,” [203]. It is also defended and discussed in the introduction to Hud Hudson’s The Metaphysics of Hyperspace, [94].

68 It is worth noting that Gunk is automatically true in open worlds with Z-DAUP.
this disc, that is part of Redball but does not overlap Redball\textsuperscript{−}. The problem is that this central disc is not open—nothing less than three-dimensions is.

Zimmerman, or like-minded Whiteheadians, might balk here. In particular, I am invoking Leibniz’s Law to establish the non-identity of O-Red and Redball\textsuperscript{−}. They might ask, “Is there really a property that they lack? To say that there is some region between L- and R-Hemi dodged by Redball\textsuperscript{−} but not by O-Red is to commit a petitio.” If we allow Leibniz’s law to apply to spatial location over non-fundamental objects (in this case, the Whiteheadian points constructed from the spheres), then this argument is vindicated. For the sake of argument, suppose for the moment that I cannot do this. I claim that Redball\textsuperscript{−} is disconnected. An alternative but equivalent definition of disconnected is as follows: a region $R$ is disconnected iff there are two disjoint (non-overlapping) open subregions of $R$, $A$ and $B$, such that $A$ and $B$ compose $R$. Because L- and R-Hemi are disjoint and open, Redball\textsuperscript{−} is disconnected. However, O-Red is a self-connected ball by stipulation. Therefore, O-Red and Redball\textsuperscript{−} are distinct. The point of all this is that Zimmerman’s preference for open worlds is at odds with MM\textsuperscript{+} and Z-DAUP.

4.5.3 Kilborn: Space-time is Discrete

William Kilborn in his article “Contact and Continuity,” [106], is primarily concerned with responding to an article by Theodore Sider, “Simply Possible,” [?]. Sider argues for a rejection of NO INTERPENETRATION by attempting to undermine the notion of repulsive forces discussed above in §4.5.2.\textsuperscript{69} Below I will rehearse Kilborn’s argument

\textsuperscript{69}Sider asks his readers to consider an exemplar of the concept of a permanent bachelor, some man who never marries. As one examines the permanent bachelor, she notices that he never gets married. No one is likely to be tempted to say that the permanent bachelor does not get married because of some kind of “anti-matrimonial forces.” He gets the name ‘permanent bachelor’ because he never gets married; it is not that he does not get married because of strange properties associated with being a permanent bachelor. Sider suggests that Zimmerman is committing the same mistake by insisting on repulsive forces between bodies. Zimmerman has (implicitly) stipulated that material bodies satisfy One-Place-One-Object but then infers that there must repulsive forces. According to Sider, the
against Sider, providing further undergirding for NO INTERPENETRATION. I will also examine Kilborn’s own solution to the contact problem, which is to deny that space is continuous.

Kilborn argues that interpenetration is impossible and that repulsive forces, though mysterious, are necessary. In order to do this, he appeals to the possibility of perforated space, i.e. that there are regions in space that cannot be occupied. These perforations are not simply non-receptacles: they are “anti-receptacles,” though not in virtue of their having or lacking certain mereotopological features. Allow to me to elaborate this distinction. On the one hand, a non-receptacle is a region \( R \) such that either \( R \) is not at present exactly occupied by some material body or \( R \) cannot be exactly occupied at all. In the latter case, parts of bodies can (non-exactly) occupy non-receptacle regions, but not exactly for whatever reason. On the other hand, a perforation is a region \( R \) such that no material object can occupy, neither exactly nor non-exactly—effectively, a hole or gap in space; hence, the “anti-receptacle.”

Now, returning to Kilborn’s argument, consider again C-Green but instead of C-Red, consider a closed spherical perforation in space, C-Hole. Following similar topological considerations from §4.2.3, necessarily, there is a non-zero distance between C-Green and any of the points included in the region of C-Hole. All the points in C-Hole are necessarily incapable of entering the occupation relation (exact or non-exact). Furthermore, Kilborn holds tight to Zimmerman’s presupposition that I dubbed, SUPREME CLOSENESS: “If two objects are in contact, then it is impossible for two distinct non-overlapping objects to be closer together than the two objects

---

permanent bachelor ends up unmarried by stipulation, allowing for the reason to be his bad hygiene, misogynistic tendencies or preference for solitude rather than because of anti-matrimonial forces. Likewise, the gaps between the bodies are due to stipulation, allowing for a host of ordinary reasons internal to the scenario to do the explaining rather than the presence of occult repulsive forces.

The above characterization of a perforation is sufficient for my purposes. One might also allow a perforation to actually move through space. In that case, a perforation would not be identical to some region, but—like matter—the perforation would occupy various regions. This latter characterization is closer to how Kilborn describes perforations. Again, the difference is minor for present purposes.
in question.” (Zimmerman, *loc. cit.*) So, there is a gap preventing contact of any kind between C-Green and C-Hole. This demands an explanation and yet none is satisfying: call it a repulsive force, call it corporeal gallantry,\(^{71}\) call it whatever—the presence of this gap is an undesirable consequence. Now, combine this undesirable consequence with the following plausible principle:

**Forces:** Only physical forces (mechanical, electro-chemical, etc.) can prevent contact\(^{72}\) between bodies or body-perforation pairs.

This principle when combined with the undesirable consequence yields absurdity.

In order to get from absurdity to a rejection of continuity, Kilborn says the following:

If Sider is right, though, and [spacetime is continuous], then either spacetime cannot be perforated, spacetime could be continuous and could be perforated but could not be both continuous and perforated or there would be interfering laws, forces or other contingent matters in every previously described possible collision course situation involving perforations. (Kilborn, [106], 278)

Let me boil it down. In light of absurdity, at least one of the following must hold: (1) space\(^{73}\) is not continuous, (2) there are no perforations or (3) **Forces** is false. Kilborn takes for granted the compossibility of continuity and perforations: if continuity is possible at all, then so are perforations. It follows that, for Kilborn, (2) is not a serious consideration. He also assumes that almost any explanation on offer for the gap between the perforation and ball is bizarre at best. Therefore, (3) is safe. Hence, according to Kilborn, what must go is the continuity of space.\(^{74}\)

I admit that I think Kilborn raises some difficulties for Sider, but I do not think that those difficulties merit the conclusion for which he uses them, *viz.* that space is

\(^{71}\) Zimmerman uses the phrase that the body would be “unaccountably deferential” to the hole, “managing to step out [of its] way.” (Zimmerman, [225], 12)

\(^{72}\) Again, note that Kilborn is assuming that any definition of contact must satisfy **SUPREME CLOSENESS ABOVE**.

\(^{73}\) I have tried to avoid tying space and time together in this chapter.

\(^{74}\) I take the quoted conditional to be false. On my view, space is continuous, perforations are treated the same as ordinary matter and **Forces** remains true—provided **TOTAL CONTACT** is satisfied (for more, see §4.6 below).
discrete instead of continuous. Kilborn does not pretend to offer a positive argument for discreteness, but simply urges that in light of absurdity, what must go is the assumption that space is continuous. I addressed why one might hesitate about discrete space at length above in §4.4.1. I think independent reasons must be given for this view to be adopted. In Quine’s parlance, the continuity of space lies pretty central in the web and cannot be thrown out without offering more on the detriments of continuous space and the benefits of discrete space.\footnote{To be sure, this is not something Quine has ever said—I am appealing only to the metaphor of the web. Something that so contravenes our ordinary conception of the structure of space requires more than a reductio. Although I do not want to go too far afield, Simons suggests that an advantage to discrete space is that it nicely complements his \textit{Extended Simples Principle}:}

\textit{Every physically basic item ([simple]) occupies at any time an extended region, called its \textit{locus}, but it has no physical proper parts. In particular it has no parts corresponding to subregions of its locus. (Simons, [191], 376)}

Although this provides further positive evidence for discrete space, the existence of extended simples is itself somewhat controversial.

\footnote{Kilborn does not provide the following definition, but it can be gleaned from an understanding of discrete spaces and some related discussion found in Simons' \textit{“Extended Simples,”} [191], and Forrest’s \textit{“Is Space-Time Discrete or Continuous?”} [69].}

As with any \textit{reductio}, one may conclude that the relevant assumptions are jointly inconsistent. Kilborn’s \textit{reductio} is effective to this end but I do not believe that abandoning the continuity of space is the best choice.

I have said enough about the tensions created in this discussion for both \textbf{NO INTERPENETRATION} and \textbf{CONTINUITY}. Now, I want to explicitly provide a characterization of how contact would work in discrete space.\footnote{\textbf{I} have said enough about the tensions created in this discussion for both \textbf{NO INTERPENETRATION} and \textbf{CONTINUITY}. \textbf{N}ow, \textbf{I} want to explicitly provide a characterization of how contact would work in discrete space.}

**Discrete Contact** For material bodies \(x\) and \(y\), \(x\) is in contact with \(y\) just in case,

\begin{itemize}
  \item[(i)] \(R_x\) and \(R_y\) are their respective receptacles in \(\mathbb{Z}^3\),
  \item[(ii)] there is some \(\vec{x} \in R_x\) and some \(\vec{y} \in R_y\) such that \(d(\vec{x}, \vec{y}) < 2\).
\end{itemize}

Ignoring the considerations given in §4.4.1 regarding the Pythagorean theorem, assume that distance between two points is what it would be in continuous space (even if no magnitude in the discrete space represents this distance.) The closest two distinct points in discrete space can be is 1, \textit{i.e.} when they are right next to each other.\footnote{Recall the definition of \(d : X^n \to \mathbb{R}, d(\vec{x}, \vec{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}. \text{ Here } n = 3.}
Any bodies that contain any points caddy-corner to one another will count as in contact. The largest such value would be $\sqrt{3} < 2$.\(^{78}\) An acknowledged value of **Discrete Contact** is its simplicity—there are no difficulties regarding spatial gaps or spatial gluts (interpenetration). Such an advantage must be weighed against the concerns raised throughout regarding discrete space.

4.5.4 **Hudson: Contact for Scattered Objects**

Neither in his *Metaphysics of Hyperspace*, [94], nor in the parallel account from his “Touching,” [92], does Hud Hudson offer a solution that handles all of the cases found in §4.2 completely. Hudson believes the **Liberal View of Receptacles**: absolutely any nonempty class of points is a receptacle. This is likely to go far in explaining why for him certain regions can’t enjoy contact; when it comes to tension between **TOTAL CONTACT** and **RECEPTACLE INDEPENDENCE**, Hudson favors the latter. In fact, some very *recherché* examples guide him towards his completed definition.

His final definition is called ‘(CONTACT-5).’ In his presentation, the language is three-sorted: capitals for regions, ‘p’ and ‘q’ for points in space and any other lower-case letter standing for a material object.

Necessarily, \(x\) touches \(y\) if and only if \(\exists R1, \exists R2, \exists w, \exists v, \exists p\) (i) \(w\) is a part of \(x\), whereas \(v\) is a part of \(y\), (ii) \(w\) exactly occupies R1, whereas \(v\) exactly occupies R2, (iii) \(p\) is a boundary point of both R1 and R2, (iv) \(p\) is a member of at least one of R1 and R2, and \(w \neq v\). (Hudson, [92], 126)

The above definition of contact has the advantage of putting distinct bodies into the same degree of contact that can be found within a single self-connected object, thereby perfectly satisfying **SUPREME CLOSENESS** and **NO INTERPENETRATION**. In addition, sense is made of contact between gritty bodies (any fusion of scattered

---

\(^{78}\)Imagine each part of space to be a block of 1 cubic unit in size. So, one block with one corner at \((0,0,0)\) and another at \((1,1,1)\) will be touching the block with one corner at \((1,1,1)\) and another at \((2,2,2)\). Orienting the axes in the usual way, distance will be determined by the bottom, front, left corner. Thus, the maximum distance between any two adjacent blocks is \(\sqrt{1+1+1} = \sqrt{3} < 2\).
point-simples such that none of them is a part of a self-connected whole). The idea is that there is contact if a gritty body converges towards the boundary point of some body. This picture has the advantage of enabling a broad-level account of contact that goes beyond simply evaluating ordinary self-connected extended material bodies. However, I think that the biggest weakness for this view is that, in spite of accounting for these *exotica*, it cannot account for contact between pairs of closed or pairs of open self-connected objects. A pair of open self-connected bodies fail criterion (iv) above and closed bodies are such that (iii) cannot be satisfied without interpenetration. Hence, although Hudson has made way for new, exciting material bodies to be in contact, he has laid aside the ones that everyone was worried about in the first place.80

---

Chapter 4.6

Reframing the Desiderata

To get a sense for where others’ views fall, I have provided the following table. Note that Hudson openly endorses *The Liberal View of Receptacles* and *Simples*. However, his theory of contact does not commit him to either of these. I flag this below in Table 4.1.

---

79Cf. the discussion of Forrest’s "Grit or Gunk," [70]. Forrest is slightly more restrictive about gritty bodies than Hudson.

80Hudson's view is essentially Bolzano's. Worries like the ones mentioned here are why Brentano uncharitably dubbed Bolzano's theory 'monstrous'.
<table>
<thead>
<tr>
<th>Desideratum/Theory</th>
<th>Bullet-Bite</th>
<th>Gunky</th>
<th>Kilborn</th>
<th>Hudson</th>
<th>Infinitesimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTINUITY</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>TOTAL CONTACT</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO INTERPENETRATION</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SUPREME CLOSENESS</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>SIMPLE INDEPENDENCE</td>
<td>✓</td>
<td></td>
<td>†</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>RECEPTACLE INDEPENDENCE</td>
<td>✓</td>
<td>✓</td>
<td>†</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key:
- ✓ – Satisfied
- † – Satisfied, but incompatible with author’s published view elsewhere
- ★ – Satisfied in exclusion of the other

Table 4.1. Comparisons of Theories of Contact on Desiderata.

Before concluding, I want to offer some refinements to some of the desiderata that might make my solution more appealing. In other words, some of the desiderata are too coarse-grained to properly distinguish the advantages provided by this view over another. Here, I try to argue that although this view does have disadvantages, it has fewer than its rivals.

The greatest tension exists within the trio of TOTAL CONTACT and RECEPTACLE INDEPENDENCE and NO INTERPENETRATION. It wants to begin by acknowledging that I have no interest in accommodating interpenetration. I have argued already for NO INTERPENETRATION and now I am simply going to refuse to countenance it as a flexible parameter. Therefore, I will suppose that any view that requires interpenetration is worse off than the present view. With that in mind, I want to draw the reader’s attention to the fact that no extant view satisfying TOTAL CONTACT has the ability to accommodate the Liberal View of Receptacles completely. The most notorious bodies are those that are extended in fewer than three dimensions. If the

---

81 Note that Kilborn satisfies all three of these. This costs him CONTINUITY, which is not a flexible parameter for me.
plane in which the contact would take place is perpendicular to the direction in which these bodies lack extension, then complete interpenetration would be required for contact. I believe Hudson (the Liberal View’s champion) is right to resist this possibility. In order to show that I have at least as much contact as Hudson and friends, I will adjust the desideratum to be slightly less stringent:

AMPLE CONTACT: All fully-extended bodies can come into contact with one another.

Casually, a fully-extended body is one that is extended in all three dimensions. It must be acknowledged that if one accepts AMPLE CONTACT over TOTAL CONTACT, an undeniable theoretical gain would have been made. With AMPLE CONTACT, Infinitesimal Contact allows a liberal-ized view of receptacles and contact at the same time. Many people who work on the issue of receptacles are often attracted to a more restricted class of receptacles than Hudson anyhow. For them, AMPLE and TOTAL CONTACT would be extensionally equivalent. On the other hand, for one who believes the Liberal View, Infinitesimal Contact satisfies AMPLE CONTACT but also puts into contact all those exotica in which he was interested. Perhaps Infinitesimal Contact could be a serious contender for one attracted to Hudson’s Liberal View but who is unsatisfied by the accompanying limited picture of contact.

This is not the only option. One could hold out for TOTAL CONTACT and simply restrict the choice of receptacles. Few are eager to defend the Liberal View. However, Infinitesimal Contact should not be incompatible with the Liberal View. What refinements are available that would capture the spirit—if not the letter—of RECEP-
TACLE INDEPENDENCE? Consider a more conservative view of receptacles, called the

Generous View of Receptacles: Any fully-extended region of space is a receptacle.\footnote{This allows gritty bodies, so long as they are not all in a single line or plane. One may also stipulate that there must be an accumulation point. See n. 3 in Chapter 2.}

\footnote{For discussion, see Gabriel Uzquiano’s, “Receptacles,” [201] and Richard Carwright’s, “Scattered Objects,” [35].}

\footnote{For the sake of brevity, I have eliminated a clause in this definition. Ideally, I would like to eliminate infinitesimally extended bodies as well. However, I do not have the space to develop the}
This view is attractive because it allows one to acknowledge his anxieties about the Liberal View while not being forced to view all receptacles as fulfilling a very rigid topological profile.\textsuperscript{85}

**Receptacle Accommodation:** *The Generous View of Receptacles is neither ruled out nor forced upon us.*

This new *desideratum* allows for **total contact** because it prevents the especially intransigent regions from being receptacles. One may put as much weight as he pleases on **total contact** and still hold a rather flexible view of receptacles.

Now for an adjusted table:

<table>
<thead>
<tr>
<th>Desideratum/Theory</th>
<th>Bullet-bite</th>
<th>Gunky</th>
<th>Kilborn</th>
<th>Hudson</th>
<th>Infinitesimal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONTINUITY</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>AMPLE CONTACT</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>NO INTERPNENTRATION</strong></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>SUPREME CLOSENESS</strong></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td><strong>SIMPLE INDEPENDENCE</strong></td>
<td>✓</td>
<td></td>
<td></td>
<td>†</td>
<td>✓</td>
</tr>
<tr>
<td><strong>RECEPTACLE ACCOMMODATION</strong></td>
<td>✓</td>
<td>✓</td>
<td>†</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Key:

✓ - Satisfied
† - Satisfied, but incompatible with author’s published view elsewhere

Table 4.2. Comparisons of Theories of Contact on Refined *Desiderata.*

**Chapter 4.7**

**Conclusion**

Here I have examined numerous attempts to solve the contact problem: pairs of bodies with certain topological profiles—regardless of their *physical* profiles—cannot apparatus for satisfactorily excluding such bodies. In any case, even without such an adjustment, a substantial measure of the *desiderata* are satisfied.

\textsuperscript{85}For such an example, see Cartwright’s “Scattered Objects,” [35].
touch without counterintuitive consequences. Many of these attempts are at odds with some intuitive desiderata, especially those desiderata that I prioritized above: CONTINUITY, TOTAL CONTACT and NO INTERPENETRATION. I have proposed a solution to this problem that satisfies these three prioritized desiderata and more with suitable refinements.
Chapter 5

Epilogue

The maxim, “By their fruits ye shall know them,”
applies also to theories.
—Hilbert, Über das Unendliche

You will publicly request Jacobi or Gauss
to give their opinions, not on the truth but
on the importance of these theorems.
—Evariste Galois, Letter to Chevalier
on the eve of Galois’ death

I recognize that many people will not agree with the conclusions I have drawn
above. Nevertheless, I would consider this study an extraordinary success if others
attempt to use infinitesimals in other metaphysical contexts. Even if others use
infinitesimals to draw opposing conclusions in metaphysics, I would consider this
work a success. At the very least, my hope is that people will find it necessary to
disagree with me: even if infinitesimals do not provide the best solutions to various
philosophical puzzles, they deserve a place in conceptual space.

Nearly fifty years ago, José Benardete wrote Infinity: An Essay in Metaphysics.
His study was in many ways similar to my own. One of his major objectives was
to see what follows in metaphysics when taking contemporary mathematicians’ con-
cept of infinity very seriously. His study is very different from mine in the following
three ways. First, much of Benardete’s work involves the creation of puzzles by using

\footnote{1Quoted from the English translation, [91], 200.}
\footnote{2Quoted in Galois Theory, [62], 223.}
contemporary concepts of infinity; my study focuses on providing solutions. Second, Benardete is most interested in seeing contemporary mathematicians’ concept of infinity put to work in metaphysics, whereas I am putting contemporary mathematicians’ concepts of continuity (hence, infinitesimals) to work.\footnote{Benardete mentions infinitesimals in a few places. They do not play a central part in his study. Through no fault of his own, Benardete’s comprehension of infinitesimals is a bit simplistic and exhibits the limits of what had been accomplished in the well-known mathematics of the time. Robinson was writing as early as 1961 and others, like Curt Schmieden, in the 1950’s. However, infinitesimals did not gain a large audience until the 1970’s.} Finally, unlike the continuum, the concept of infinity does not exhibit as much mathematical diversity as the concept of continuity. This latter fact is reflected in Benardete’s text.

Before concluding, I want to echo some of his final sentiments:

I freely admit that the metaphysical adventure very much smacks of the quixotic. In a spirit of metaphysical hyperbole, one finds oneself engaged in an investigation of utopian concepts that transcend the empirical. There have always been, there always will be, philosophers of a severe and sober cast of mind who set their faces against all such utopian transports. However much I may inveigh against them, these philosophers are my friends. For I am persuaded that we are engaged in a common enterprise and that it belongs to the very rhythm of philosophy itself that the downward dialectic should alternate with the upward. Nor should I wish to deny that it is the downward dialectic which possesses a power and an authority to which the upward dialectic can only fitfully aspire. (Benardete, [13], 285)

With Benardete, I hope that more sober philosophers will forgive any flights of fancy and fraternally grasp the extended hand of friendship.
Bibliography


[57] ———— “All Numbers Great and Small” in [61], pp. 239-258.


[111] ————– “Cauchy and the Continuum: The Significance of Non-standard analysis for the History and Philosophy of Mathematics” in [110], pp. 43-60.


160


