Illustrations of a Realist Methodology for the Philosophy of Mathematics

James V. Martin

A Dissertation Presented to the Faculty of Princeton University in Candidacy for the Degree of Doctor of Philosophy

Recommended for Acceptance by the Department of Philosophy

Adviser:
Professor John P. Burgess

September 2018
Abstract
This dissertation outlines and illustrates a methodological position most naturally described as realist. The kind of realism advanced and employed, however, has its most familiar analogue in legal and political philosophy.

The American legal realists are realist in the relevant sense. These philosophers argue that, when one looks at how courts really decide cases, intuitive judgments about a fair decision given the case’s facts play a more decisive role than reasoning based on distinctively legal rules. They note that judges may say otherwise, but that certainly does not imply this is the case. The political philosopher taking Raymond Geuss’s advice, “Don’t look just at what they say, think, believe, but at what they actually do, and what actually happens as a result,” also practices the relevant form of realism. These realists are alike in their attempts to replace a dependence on the unreliable self-reporting of political actors, whose own motivations may be opaque or self-deceived, with something more objective by appealing to the detailed examination of concrete cases and facts.

I argue that contemporary philosophy of mathematics needs to adopt a methodology realist in an analogous sense. Although the field has moved towards paying closer attention to mathematical practice, too much weight is still placed on the psychology of mathematicians and the analysis of the philosophical offerings found outside proofs and in prefaces.

In Chapter 1, I use Wittgenstein’s writings to aid in outlining the main tenets of the realist methodology illustrated in the following chapters. In these later chapters, I make use of this methodology to address the following: (1) The setting-dependence of mathematical questions, and problems with standard accounts of “mysterious” mathematical behaviors; (2) the role of geometric concepts and techniques in non-geometric contexts; (3) the puzz...
zling concept of mathematical coincidence; and (4) the explanatory potential of proofs by mathematical induction. In each of these cases, a realistic approach to subject brings clarity to the questions being asked and also provides the means for satisfying answers or dissolutions.
Es werden, in Gottes Namen, ja nicht geradezu sieben Jahre sein!

—Thomas Mann, Der Zauерberg
Contents

Abstract ........................................................................................................ iii

1 A Plea for Ordinary Realism ........................................................................ 1
   1 Wittgenstein as a Realist ................................................................. 4
      (a) Being Skeptical about Prose .................................................. 8
      (b) Paying Attention to the Complex Mixture of Mathematics .... 10
      (c) Focusing on Language in Use ................................................. 11
   2 Minimal Philosophical Assumptions .............................................. 14
   3 The Synthetic Approach ................................................................. 18
   4 ‘Use’ in Mathematics .................................................................... 21
   5 Example: \(-1/12 = 1 + 2 + 3 + \ldots\) ............................................ 25

2 On the Mystery of the Real Power Series ............................................ 30
   1 The Mystery ................................................................................. 32
   2 The Mystery, Solved ..................................................................... 34
   3 Naturalist Worries ....................................................................... 35
      (a) Expanding on the Worry ....................................................... 35
      (b) Responding to the Worry ..................................................... 36
   4 Reconsidering the Question ........................................................ 38
   5 Domain-crossing Questions .......................................................... 43
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Is This an Explanation?</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>(a) Mathematical Explanation</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>(b) Back to the Question</td>
<td>54</td>
</tr>
<tr>
<td>7</td>
<td>Steiner- and Kitcher-style Explanations</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>(a) Steiner’s View</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>(b) Kitcher’s View</td>
<td>58</td>
</tr>
<tr>
<td>8</td>
<td>Conclusions</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>(a) Why the Popularity?</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>(b) What Can Be Learned?</td>
<td>61</td>
</tr>
<tr>
<td>3</td>
<td>On the Role of Geometric Frames in Mathematics</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>1 Mere Convenience vs. Conceptual Abbreviation</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>(a) Example: Lüroth’s Theorem</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>2 A Wonderful Aid to the Mind</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>3 Geometric Scaffolding</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>4 Conclusions and Implications</td>
<td>94</td>
</tr>
<tr>
<td>4</td>
<td>Mathematical Coincidence and Care</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>1 A Budget of Coincidences</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>2 Initial Puzzles</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>3 The Dictionary Analysis</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>4 A Pascalian View</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>5 Does Anything Go?</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td>6 Conclusion</td>
<td>117</td>
</tr>
<tr>
<td>5</td>
<td>Mathematical Induction and Explanation</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>1 A Comment on ‘Explanation’</td>
<td>120</td>
</tr>
</tbody>
</table>
2  Lange’s Argument and Replies ........................................... 123
   (a) The Argument ...................................................... 123
   (b) Replies ............................................................... 124

3  Background Assumptions .................................................. 128
   (a) Asymmetry .......................................................... 128
   (b) Occam’s Razor and Simplicity .................................. 131
   (c) Instances and Generalizations ................................. 133

4  Justifying the Assumptions ............................................... 134
   (a) Optimism ........................................................... 135
   (b) Hyper-Physics .................................................... 136
   (c) Objective Explanation ........................................... 138

5  Proof and Explanation of Meaning ..................................... 139

6  Are Inductive Proofs Explanatory? ..................................... 141

Bibliography 145
Chapter 1

A Plea for Ordinary Realism

I am only trying to recommend a certain sort of investigation.
—Wittgenstein, Lectures on the Foundations of Mathematics

Philosophers of mathematics have contributed more than their fair share to the overworking of the term ‘realism.’ In the literature, describing an author as a realist may be taken to indicate something about her ontological commitments,\textsuperscript{1} an adherence to a semantic claim about the existence of certain recognition-independent truth-conditions,\textsuperscript{2} or even a decision to focus on exploring philosophical issues arising from topics more mathematically advanced than, say, basic set theory or arithmetic.\textsuperscript{3} In this dissertation, I shall, in part, attempt to combat this tradition of lexical overburdening by outlining and illustrating a methodological position that, while most naturally called realist, warrants this label simply by being realist in the everyday, non-philosophical sense of the word. I contend that this is the only kind of realism that philosophers of mathematics have any need to subscribe to, and shall strive to make the obviousness, as well as the fruitfulness, of the approach apparent in the chapters that follow.

\textsuperscript{1} See, for example, the “robust” or “thin” realism discussed in (Maddy 2007: 362-380).
\textsuperscript{2} This understanding of ‘realism’ is most commonly associated with Michael Dummett.
\textsuperscript{3} This is the sense (Corfield 2003) uses to urge more philosophical thinking about “real” mathematics.
In philosophical work, the most familiar advocates of the quotidian realism to be described in this opening chapter pursue practical questions in legal and political philosophy. By briefly considering the realist approach in settings where these sorts of philosophers work, I hope to ease the reorientation away from the philosophically-loaded senses of the term so commonly employed within the philosophy of mathematics and to better allow the term to resume its everyday meaning for the remainder of the dissertation.

In *Philosophy and Real Politics*, for instance, Raymond Geuss expounds a realist methodology that enjoins the political philosopher to inquire into political thought and action with the aim of discounting illusory motivations and goals by ignoring idealizations and rational reconstructions. Instead, he urges political philosophers to study the concrete realities that have actually motivated real human actors as they have pursued their definite social and political goals. Geuss’s realist offers the guidance, “Don’t look just at what they say, think, believe, but at what they actually do, and what actually happens as a result.”\(^4\) A philosopher following this advice will attempt to be a realist in the same fashion as Napoleon (according to Emerson): “He [will be] a realist, terrific to all talkers, and confused truth-obscurring persons.”\(^5\) By investigating real actions rather than the principles that purport to motivate and justify these actions, the realist strives to strip away obfuscation and ideology in order to get at the truth of things as they really are in the realm of politics. This is a paradigm case of taking a realist approach to a subject—realist, in the ordinary sense of the word.

The American legal realists can be seen as being realist in a similar way. These philosophers argued that when one looks at how courts really decide cases, intuitive judgments about a fair decision given the cases’s facts play a more decisive role than reasoning

\(^4\) See (Geuss 2008: 10, emphasis in the original). This kind of thinking is also prominent in the work of Max Weber; see, e.g., (Weber 1968: Part 2, Ch. X).
\(^5\) (Emerson 1850: 143)
based solely on distinctively legal rules. The legal realists note that judges may say (and may even be required to say) that their rulings merely apply a ready-made law (or ready-made laws) and do not depend on further moral or political considerations, but they remind us that this certainly does not imply that this is actually the case. In their theorizing about adjudication, the legal realists attempt to replace a dependence on the unreliable self-reporting of judges, whose own motivations and decision-procedures may be opaque or self-deceived, with something more objective by appealing instead to the detailed examination of cases and the concrete facts surrounding them. Again, this is a realist view of adjudication that has little to do with any of the common philosophical ways of understanding the term realism.

Contemporary philosophy of mathematics has rightly moved in the direction of attempting to pay closer attention to the ordinary practice of mathematics. This trend may suggest that the subject has also moved in the direction of the sort of ordinary realism described above. This has not really been the case, however. Too much weight is still placed on the psychology of mathematicians and the analysis of the often meager philosophical offerings found lurking around a mathematician’s proofs or tucked away in her prefaces. There remains, therefore, a need within the field for the adoption of a methodology that is realist in a sense analogous to the one employed in some of the more familiar arenas of realist thought discussed above. At any rate, the primary goal of this dissertation is to make the case for this methodological position by demonstrating the clearer view it affords of several problems of recent interest to philosophers of mathematics and science.

The philosopher who has previously been most adamant about practicing the kind of ordinary realism to be advocated here when thinking about mathematics is (the later) Wittgenstein. In the following section, I will briefly look to Wittgenstein for inspiration as

---

6 See, for example, (Leiter 2005: 50-53).

7 Cf. (Wittgenstein 1939/1989: 55, 103). It remains a matter of controversy, however, whether Wittgenstein really wanted nothing more than for us to look at the workings of mathematics “from close to” (1953/2009: §51, emphasis in the original).
I outline the main tenets of the realist methodology for the philosophy of mathematics that will be illustrated in the remaining chapters. I will then expand on what I take to be three key insights that can be gleaned from Wittgenstein’s writings on realism in the relevant sense in the three subsequent sections before closing the chapter with a simple, recently-popular mathematical example in relation to which these ideas come together nicely in playing a clarificatory role.

Wittgenstein as a Realist

Since I have been suggesting that the methodological realism advocated in this dissertation is the kind of realism everyone is already familiar with, taking the time to outline the main tenets of the approach may seem superfluous at best and immediately undermining of the claim to familiarity at worst. Furthermore, appealing to Wittgenstein of all people for what looks to be a list of theses about a kind of realism will likely add to the apparent strangeness of this portion of the chapter. However, each of these worries can be allayed by the following considerations. (I’ll work backwards through the potential concerns.)

First, although it’s true that at least since Dummett’s well-known review of the Remarks on the Foundations of Mathematics Wittgenstein has been associated with a form of anti-realism,⁸ that association—whether or not it is warranted—is orthogonal to the ordinary realism I will be using Wittgenstein’s ideas to aid in expounding: one can be a realist in the relevant sense while also defending some form of anti-realism in the Dummettian sense. The best recent work on Wittgenstein’s writings in the philosophy of mathematics has in fact moved away from trying to categorize Wittgenstein as an advocate of any particular “-ism” and has instead emphasized the methodological realism often on display in

---

⁸ See (Dummett 1959: 348).
his writings.\textsuperscript{9} It’s within this recent trend in Wittgenstein scholarship that the appeal to Wittgenstein’s thought in this chapter should be situated.

Second, the methodological suggestions to be outlined below are not intended to be theses in anything like the sense in which Wittgenstein is committed to disavowing theses. Wittgenstein’s objections to “theses” in philosophy arise from his general attempt to pursue philosophical clarity without any kind of dogmatism.\textsuperscript{10} He takes exception in this regard to philosophers putting forward propositions that must be true (e.g., that the world is the totality of facts, not of things) and that then serve as premises for further philosophical argumentation. Statements that have appeared to some as being in violation of the prohibition on theses in Wittgenstein’s writings, perhaps like those to be presented below, however, are in fact intended to be mundane “grammatical” assertions or summarizing comments that tie together a string of previous observations in a “perspicuous” way. Taken in this light, the “theses” advocated below don’t contradict the spirit of Wittgenstein’s philosophical method because they aren’t supposed to state necessary truths and don’t serve as the premises for further argumentation either.\textsuperscript{11}

Finally, while it’s true that the realist methodology being advocated here is purported to be of a familiar sort, what exactly this realism amounts to in the particular case of philosophy of mathematics is nevertheless worth attempting to make clear. Just as realism in political or legal philosophy requires some preliminary work that isolates certain sorts of actions and claims to which special attention should be paid or for which skepticism is especially warranted, realism in the philosophy of mathematics also needs to have guidelines in place that suggest where to focus attention and where to look for secure or insecure pieces of data. General concepts don’t simply apply themselves to particular cases.

\textsuperscript{9} See in particular the work of Felix Mühlhölzer and Juliet Floyd in the bibliography.

\textsuperscript{10} Cf. (Waismann 1929-32/1979: 183).

\textsuperscript{11} For more on the controversy about Wittgenstein and theses, see (Baker and Hacker 2005: Ch. XIV, §6).
Saying more about the realist approach in relation to the philosophy of mathematics is further necessitated by the fact that “practice-first” philosophers of mathematics already seem to think of their approach as being realist in the sense intended here. Philosophers taking this practice-oriented view of the subject have largely done so motivated by a fear of philosophical presuppositions infecting their observations of mathematics and its practices. By looking at mathematical practice with eyes unobstructed by theory, they hope to see the world of mathematics and of mathematicians as it really is. Obviously, there is something correct about the sorts of fears and the attempts to assuage them that come out of this line of thought. However, without the guidance of some amount of theorizing about methodology, appeals to practice tend to produce the kind of information that shows itself to be very hard to put to use: for example, mathematicians may say one thing and then do another; one mathematician or group of mathematicians may say the exact opposite as another; and so on.

Juliet Floyd has rightly called for more efforts to remedy this lack of theorizing in practice-oriented philosophy of mathematics, suggesting that “not enough has been done to pick apart the force, the character, and the scope of what an appeal to, or characterization of, mathematical practice should and can be.”\(^{12}\) This call to action is especially important in light of the fact, also emphasized by Floyd, that there is not obviously just one monolithic “practice of mathematics” to which philosophers can safely appeal. Floyd problematizes the current situation nicely when she asks—after noting that philosophers and mathematicians have recently tended to become more modest while working through philosophical issues in mathematics—“What is the best way to be modest?”\(^{13}\) The appeal to portions of Wittgenstein’s thought in this chapter is an attempt to begin answering this question.

The general character of Wittgenstein’s methodological thinking in relation to the philosophy of mathematics becomes clear already in the opening address to the students

\(^{12}\) (Floyd 2015: 17)  
\(^{13}\) Ibid., 17.
attending his *Lectures on the Foundations of Mathematics*. He begins these lectures by trying to answer the question of how he, a philosopher, can have anything at all to say about the foundations of mathematics. He answers this question by noting that he speaks ordinary language just as expertly as any other normal adult language-user and claiming that this ability is enough to qualify one to discuss many aspects of the subject.

I will only deal with puzzles which arise from the words of our ordinary everyday language, such as “proof,” “number,” “series,” “order,” etc.\(^\text{14}\) Knowing our everyday language—this is one reason why I can talk about them.\(^\text{15}\)

Many of the problems of recent interest in philosophy of mathematics also stem from the ways in which ordinary language interfaces with the practice of mathematics. Key terms here include “explanation,” “geometrical,” “surprise,” “coincidence,” “depth,” and “beauty.” Wittgenstein scatters remarks about explanation throughout his writings about mathematics—although he doesn’t focus on mathematical explanation within mathematics specifically—and a small portion of the *Remarks on the Foundations of Mathematics* deals with the surprising in mathematics.\(^\text{16}\) Clearly, however, the same sorts of methods that proved to be useful in investigating the sorts of problems Wittgenstein actually got around to addressing should be of use in investigating the current puzzles that arise from other pieces of everyday language employed within mathematics. The following three subsections will introduce what I take to be three of the most important of these methods and explain my view about how Wittgenstein understood each. The subsequent three sections will expand on these initial ideas and explain more precisely how they will be put to use in the remainder of this dissertation.

---

\(^{14}\) He later adds “foundation” to this list as well.

\(^{15}\) (Wittgenstein 1939/1989: 15)

\(^{16}\) (Wittgenstein 1956/1983: Part I, Appendix II)
Being Skeptical about Prose

I will only invent a new interpretation to put side by side with an old one and say, “Here, choose, take your pick.” I will only make gas to expel old gas.\footnote{(Wittgenstein 1939/1989: 14)}

One of Wittgenstein’s most well-known pieces of advice for taking a realist view of mathematics and mathematical practice is to take care to distinguish between prose and proof when reading mathematical writing. He suggests that mathematics expressed in prose is more likely to be misleading or to be used in the service of suspect metaphysical projects—and he thinks this to be true of \textit{any} prose interpretation. This is why Wittgenstein promises only to offer his own prose (or “gas”) in an effort to show that one has a choice about how to make the transition from something like a calculus to something more closely resembling ordinary language. With this promise in place, we shouldn’t expect to find Wittgenstein offering the “correct” interpretation of any mathematical results, and any methodological position inspired by Wittgenstein’s thinking will shy away from laying down the law about correct interpretations as well.

Wittgenstein’s general skepticism about prose is motivated by the fact that prose is supposed to be essentially deficient because “[t]he calculation illuminates the meaning of the expression in words. It is the finer instrument for determining the meaning.”\footnote{(Wittgenstein 1956/1983: II §7, emphasis in the original)} He often expresses the same idea when saying things along the lines of, “If you want to know what has been proved, look at the proof.”\footnote{See (Wittgenstein 1939/1989: 39). It’s interesting to note, however, that he immediately qualifies this claim, calling it an exaggeration and saying that it’s partly true and partly false.} The conceptual confusions and unclarities that arise out of prose are, in Wittgenstein’s view, the main objects of interest for philosophers of mathematics and can best be corrected by closely relying on the finer instruments available to us in proofs and what they show independent of their (re)interpretations.
Wittgenstein’s aim to challenge prose rather than proofs in his reflections on mathematics is also why he believes he can avoid becoming a ham-fisted director, who, instead of doing his own work and merely supervising his employees to see they do their work well, takes over their jobs until one day he finds himself overburdened with other people’s work while his employees criticize him.\(^{20}\)

The stated goal of Wittgenstein’s remarks on mathematics is not to interfere with any actual mathematics—it should only be (mis)interpretations that come into question. As he argues with respect to other topics, the philosopher must leave everything as it is.\(^{21}\) Of course, not literally everything is meant to remain unchanged since we’re supposed to come away with confusions removed or with a deeper understanding of the subject being addressed. However, it’s only the “names and allusions that occur in the calculus” that should be affected by the philosopher’s critique of mathematical texts.\(^{22}\)

This goal of only addressing prose and not actual mathematics makes the question of how the two can be clearly distinguished a pressing one. Many interpreters of Wittgenstein’s work have found drawing the distinction to be no simple matter, and for good reason.\(^{23}\) Consider, for instance, the heart of the critique of (Skolem 1923/1967) found in the Philosophical Remarks.\(^{24}\) Wittgenstein points to what he considers to be an odd feature of proofs by mathematical induction in elementary number theory: one actually proves the result in a base case and then in an induction step from which the result is then said to hold for all numbers. He is inclined to say that the move to the claim that the given property holds for all numbers is not really a part of the proof although the proof does present a pattern for demonstrating the result in regard to any particular number. If one takes the view that what is strictly in the proof determines the meaning of a result, “and so \(\varphi\) is true of all \(n\)”

\(^{20}\) (Wittgenstein 1933/1978: V §24)
\(^{21}\) (Wittgenstein 1953/2009: §124)
\(^{22}\) (Waismann 1929-32/1979: 149)
\(^{23}\) See, e.g., (Floyd 2001) and (Kienzler and Grève 2016: 81).
\(^{24}\) (Wittgenstein 1930/1975: §163)
could be seen as a prose interpretation of an otherwise unobjectionable proof. Since it’s not my intention to defend any of Wittgenstein’s particular observations about what is or isn’t prose, the question of whether or not he sees things aright here needn’t be settled. It is, however, important to note before moving on that prose and proof can’t be separated as easily as looking for one outside and the other inside of a “Proof: … Q.E.D.”-pair.

Paying Attention to the Complex Mixture of Mathematics

[M]athematics is a MULTICOLOURED mixture of techniques of proof.25

According to Wittgenstein, a realist view of mathematics will reveal—in its most commonly quoted form—that “mathematics is a motley.” I have, however, followed Felix Mühlhölzer in changing Anscombe’s more familiar rendering of Wittgenstein’s “BUNTES Gemisch” to “MULTICOLOURED mixture” in the above quotation. Mühlhölzer argues that the term ‘motley’ has negative connotations that don’t fit well with the general thrust of Wittgenstein’s remarks about the mixture of proof methods found in mathematics. I use his translation of this remark to signal my agreement on this point.26

On one way of understanding it, this “mixture”-perspective of mathematics is almost a truism. We clearly do use different proof methods with differing frequencies when doing, say, combinatorics as compared to real analysis or algebra or set theory and so on. And working with categories versus sets or the rationals versus the reals can feel like working with significantly different species of mathematical object. Despite these points in its favor however, this claim of Wittgenstein’s is often taken to be one of his most controversial mathematical remarks. Philosophers and mathematicians alike have taken exception to this line of thought, perhaps because they have taken him to be denying some kind of underlying

25 (Wittgenstein 1956/1983: III §46, emphasis in the original)
26 See, however, (Hacking 2014: 57) for a contrary view.
unity of the subject. Thus, John Burgess explicitly argues that “[m]athematics is no mot-
ley.”\(^{27}\) And the Fields medalist Jean-Pierre Serre exemplifies a common conception of the
unity of mathematics when he states that many important mathematical questions “are not
group theory, nor topology, nor number theory: They are just mathematics.”\(^{28}\) Although
there appears to be a conflict of opinions here, when I return to expand on this insight of
Wittgenstein’s in section 3, I hope to show that the conflict is indeed only apparent. That
Wittgenstein may not be saying anything stronger than that he wants to understand and ex-
plain “the colourfulness [die Buntheit] of mathematics”\(^{29}\) is at least suggested by the fact
that he prefices his remark about the multicoloured mixture of mathematics with the phrase
“I would like to say.” Wittgenstein often uses this kind of language to state views that are
tempting or that present useful perspectives, but which are not ultimately unquestionable.\(^{30}\)
Given this qualification, we shouldn’t rule out the possibility that one can be interested in
even a motley without denying that at bottom it’s all “just mathematics.”

**Focusing on Language in Use**

> [I]f we had to name anything which is the life of the sign, we should have to say it was
> its *use.*\(^{31}\)

Wittgenstein characterizes Frege as claiming that, if the formalists are correct that
“[a]rithmetic is concerned only with the rules governing the manipulation of the arithmeti-
cal signs, not, however, with the reference of the signs,”\(^{32}\) the signs used in mathematics
“would be dead and utterly uninteresting, whereas they obviously have a kind of life.”\(^{33}\)

\(^{27}\) (Burgess 2015: 60)
\(^{28}\) (Serre *et al.* 1999: 35)
\(^{29}\) (Wittgenstein 1956/1983: III §48)
\(^{30}\) This qualification to the quotation starting this subsection is also emphasized by (Säätelä 2011: 176).
\(^{31}\) (Wittgenstein 1958: 4, emphasis in the original)
\(^{32}\) See (Frege 1903/1960: §88). This is Frege’s way of restating the views of E. Heine and J. Thomae.
\(^{33}\) (Wittgenstein 1958: 4)
He returns to the metaphor of live and dead signs often in his writings. For example, the following exchange in the *Investigations* again connects the idea of a living sign with the concept of use: “Every sign by itself seems dead. What gives it life?—In use it is alive.”\(^{34}\) Evidently, understanding the importance of use in bringing the language of mathematics to life in Wittgenstein’s views about the subject requires at least a brief excursion into his views about meaning.

Two slogans are commonly associated with Wittgenstein’s general conception of meaning: (i) the meaning of a word is its use in language;\(^{35}\) and (ii) the meaning of a word is what the explanation of its meaning explains.\(^{36}\) These apparently distinct claims are ultimately tied together through the discussion of rule-following that takes place in the *Investigations*. In brief, Wittgenstein suggests that an explanation of meaning provides something like a rule for the use of a word, and this fact makes particular uses of the word applications of the rule for usage. Since Wittgenstein’s rule-following remarks are supposed to bring out the “internal relation” between rules and applications of rules—you understand a rule when you can apply it correctly—the slogans above are unified from this point of view.

The importance of use further enters this picture of meaning by way of the claim that we often get ourselves into philosophical tangles when we view language apart from its use—when language goes “on holiday.”\(^{37}\) Wittgenstein suggests that we often find questions about the meaning of specific words or propositions taken out of context difficult or impossible to answer, but, when in use, any hesitation about what a word or phrase means for the most part seems to instantly vanish. This is evidently the basic point of the following passage.

\(^{34}\) (Wittgenstein 1953/2009: §432, emphasis in the original)
\(^{35}\) (Wittgenstein 1953/2009: §43)
\(^{37}\) (Wittgenstein 1953/2009: §38)
If I am drowning and I shout “Help!”, how do I know what the word Help means? Well, that’s how I react in this situation. —Now that is how I know what “green” means as well and also know how I have to follow the rule in the particular case.\(^\text{38}\)

Wittgenstein’s general hope is that many, if not all, of our philosophical difficulties can be resolved if we can “bring words back from their metaphysical to their everyday use” where they function without fault.\(^\text{39, 40}\) It is also one of the primary hopes of the realist approach being advocated and illustrated in this dissertation that paying attention to the actual use to which a number of ordinary concepts are put in mathematical practice will dissolve some of the philosophical puzzlement that continues to infect stretches of our thinking about the subject.

Writing about the inspiration he takes from Nietzsche, Foucault once suggested that “[t]he only valid tribute to thought such as Nietzsche’s is precisely to use it, to deform it, to make it groan and protest.”\(^\text{41}\) Those taking inspiration from Wittgenstein are often seen as failing to make this kind of tribute to his thought. As a result, Ray Monk’s dim assessment of the current audience for Wittgenstein’s ideas in the philosophy of mathematics seems accurate.

Researchers in this field are now writing, not for logicians or mathematicians, nor even for philosophers of mathematics. They cannot even assume they will be read by other, more general Wittgenstein scholars. The only people they can assume will read their work are other researchers working on Wittgenstein’s philosophy of mathematics.\(^\text{42}\)

In the following sections, I will explain how I plan to use and deform the main ideas derived from Wittgenstein briefly outlined above. In my view, only through such deformation and

\(^{38}\) (Wittgenstein 1956/1983: VI §35, emphasis in the original)

\(^{39}\) (Wittgenstein 1953/2009: §116)

\(^{40}\) What exactly ‘metaphysical’ is supposed to mean in this statement is the matter of a debate that needn’t be settled here. For the record, I’m roughly in agreement with Gordon Baker, who suggests that metaphysical uses try to express essences or to pass themselves off as being scientific but are not. Cf. (Baker 2009: 96-100).

\(^{41}\) (Foucault 1975: 53-54)

\(^{42}\) (Monk 2007: 293, emphasis in the original)
decoupling from Wittgenstein-orthodoxy will these important methodological principles have the chance to gain the widespread acceptance they deserve.

Minimal *Philosophical* Assumptions

Wittgenstein’s suggestion, “If you want to know what has been proved, look at the proof,” is the most straightforward way to act on any skepticism about prose in mathematical writing. The suggestion may be a good one as far as it goes, and, despite its seeming to be an unorthodox perspective, it appears to have been shared by as unlikely a figure as G. H. Hardy. Wittgenstein often takes the views of Hardy as exemplifying precisely the kind of thinking about mathematics that he aims to deflate, yet the two appear to agree on this point at least: when discussing a pair of “real” mathematical theorems in *A Mathematician’s Apology*, Hardy says that “in the theorems, of course, I include the proofs.” Nevertheless, the advice calls for expansion and modification for at least the following reasons.

First, as already noted, it’s not an easy task to distinguish between prose and non-prose even within a selection of text demarcated as a proof. This being the case, simply saying, “Look at the proof,” amounts to little real guidance. Second, restricting attention to proofs alone could commit one to a kind of “proof chauvinism” that leads to many important and interesting aspects of the practice of mathematics being overlooked completely. Any skepticism about prose will suggest caution when examining non-proof material, but if we want a clear view of the subject, we had better be prepared to look everywhere we can.

---

43 (Wittgenstein 1939/1989: 39)
44 (Hardy 1940: 113)
45 Hegel may be another unexpected ally here: “[N]or is the result the actual whole, but rather the result together with the process through which it came about” (Hegel 1807/1997: 2).
46 The term proof chauvinism is from (D’Alessandro forthcoming), which argues that not every mathematical explanation is a proof. (Lange forthcoming) addresses this topic as well.
The primary way I will be acting on a skepticism about prose interpretations of mathematical results in this dissertation can be brought out most clearly by considering an analogy with the technical investigations of reverse mathematics. Reverse mathematics is a subfield of mathematical logic initiated by the work of Harvey Friedman in the 1970s that takes as one of its main goals the determination of minimal axiom systems required to prove standard mathematical theorems.\footnote{See, e.g., (Friedman 1975) and (Simpson 1999), which is the standard reference.} An axiom system is taken to be of the minimal strength required to prove a theorem if (i) the theorem can be proved from the axioms (over a weak base logic) and (ii) the axioms can be proved from the assumption of the theorem as well. The proof involved in demonstrating (ii) is where the ‘reverse’ comes from in the name: there’s a sense in which we’re going backwards if we start from a theorem and use it to prove an axiom. By finding the minimal setting in which a given theorem can be proven, we can hope to get some sense of that theorem’s real strength or content. Similarly, by making minimal philosophical assumptions about the correct way to think about a mathematical object or theorem, it may be possible to find the minimal thing or things that must be said in order to comprehend that mathematical object or construction, or to make sense of that particular result. The claims that look like they can’t be denied can at least be hoped to be free from the sorts of problems endemic to other more elaborate prose interpretations, and could arguably be considered to contain the real mathematical content involved in a given case.\footnote{The attempt to minimize philosophical background assumptions also helps to make room for the “pluralism in perspectives” suggested by Michelle Friend, another author that can be seen as attempting to find the best way to be modest when philosophizing about mathematics. See (Friend 2014: 25).} It might further be hoped that, just as the program of reverse mathematics has revealed a hierarchy of natural axiom systems based on increasingly powerful set-existence axioms, some sort of natural hierarchy of background philosophical assumptions might emerge to help organize swaths of interpretations of mathematics and its practice.\footnote{Steven Simpson (1999) has suggested correlations with the subsystems of second-order arithmetic considered in reverse mathematics and particular philosophical positions (e.g., that \( \text{RCA}_0 \) is roughly the system that
As a simple example of minimizing philosophical assumptions when viewing a familiar piece of mathematics, consider the question of whether 2 (the natural number) and +2 (the integer) are identical. \textit{Prima facie} the answer is no: 2 has no additive inverse, while +2 does have one; 2 has only a finite number of predecessors, while +2 has infinitely many; etc. We do often talk as if these numbers are the same though, and we \textit{can} do so by saying that the varying properties obtain when one and the same object is placed among the relations obtaining in different mathematical structures. But we certainly don’t \textit{have} to say this, and on the basis of purely mathematical facts, it seems as if the (logically) weaker position is to not say so: if there are good reasons to identify 2 and +2 in our thinking or our practices, they can be considered later.\textsuperscript{50} But if we can work without making this further—arguably, metaphysical—assumption, it \textit{may} prevent problems down the line.

Certainly, one natural objection to this kind of reasoning is to say that mathematical practice already does treat 2 and +2 as the same object, so we ought to do the same in our philosophical engagement with the subject. This appeal to practice may move too quickly, however. While it’s true that mathematicians often treat the natural numbers as a subset of the integers even when the integers are—by construction—usually taken to be equivalence classes of pairs of natural numbers, they often will note that it is an “abuse of language” to do so.\textsuperscript{51} Further, if we attempt to act on the realist slogan, “Don’t look at what they say, but at what they do,” the fact that 2 and +2 have two different constructions in \textit{ZFC}, which is still the most common foundation for these number systems, ought to be enough to conclude that the non-identity of these numbers should be taken to be the default, regardless of what a mathematician wants to say after the constructions are completed.

\textsuperscript{50} (Tsementzis and Halvorson forthcoming) offers some reasons for thinking that questions like, “Are 2 and +2 identical?” are not even particularly good or fruitful ones.

\textsuperscript{51} Cf. (Burgess 2015: 108).
The realist philosopher is likely to be charged with being pedantic here, but the analogy with reverse mathematics is again helpful in seeing why this need not be overly worrying. Reverse mathematics investigates theorems that have already been proved and their relations to logical systems unfamiliar to most ordinary mathematicians. These investigations could be deemed uninteresting or pedantic by more general practitioners as well, but the techniques employed are the ones necessary for the tasks the subject sets itself. Further, the reverse mathematician clearly doesn’t say that all mathematics should actually be done in the weak systems she studies. Her claim is only that it can be interesting and useful to study mathematics this way. The philosopher taking a realist approach to the subject can respond to charges of pedantry along similar lines. For the most part, distinguishing between 2 and +2 is not necessary or to be recommended. But, when thinking about certain philosophical claims (e.g., 2 always had an additive inverse, it was just waiting to be discovered among the structure of the integers), taking this extra care is potentially useful and illuminating.

One final objection should be addressed before moving forward. At least since Paul Benacerraf’s “Mathematical Truth” (1973), one of the primary goals for philosophers of mathematics has been to provide an ontology or a semantics for mathematics and its language. When taking this approach to the subject, the question, “Do ‘2’ and ‘+2’ refer to the same object?” looks like it must be answered before we can do much else. By saying as little as possible about 2 and +2 (e.g., “We needn’t identify them”), it may seem as if the philosopher taking the realist approach on offer here is shirking her main duties.

While there is certainly a long tradition that takes providing an ontology or semantics to be the (or if not the, at least a major) goal of the philosophy of mathematics, the question of what the ultimate goal of the subject should be is a kind of meta-question to which different answers can certainly be given. There is plenty of work to do in the philosophy

---

52 On most popular views about meaning, a semantics assumes a set of objects to serve as referents and so presupposes an ontology.
of mathematics that isn’t based on doing semantics or metaphysics. The realist approach advocated here suggests a set of methods that are intended to be tailored to a practice of philosophy of mathematics that pursues a program at least partially distinct from the one laid out in (Benacerraf 1973). For a philosopher of mathematics with different goals in mind, there may have to be some parting of ways.

The Synthetic Approach

The suggestion to see mathematics as a multicolored mixture of techniques of proof also requires some further explanation and expansion. Again, an analogy may prove useful here.

One can pursue the study of geometry by so-called analytic or synthetic means. When taking an analytic approach, one generally starts with a “space” of points (e.g., pairs of real numbers or equivalence classes of triples of real numbers in projective geometry) and then decides on ways to construct lines and other geometric objects out of these more basic entities. This approach is similar to the way in which set theory can be used to construct essentially all the objects and structures that the ordinary mathematician cares to deal with. A synthetic approach to geometry, on the other hand, begins with a set of axioms which can be developed independently of whether or not one finds a domain of points and lines that these axioms can be seen as describing. This was the approach to geometry taken by Euclid, and, in fact, every geometer before Descartes/Fermat came up with the idea of coordinate geometry. A mathematician taking a synthetic approach to a specific subject, e.g., topology, more generally will be likely to say, “Here are some axioms that define a topological space. Let me investigate them using the best means available,” without attempting to reduce the subject to something constructed out of sets. The realist philosopher of mathematics, who
wants to see the patchwork of mathematical techniques as it stands in all its multicolored
splendor, will tend towards a synthetic approach in this sense as well.\textsuperscript{53}

The realist is motivated to take a synthetic approach also in an attempt to respect the
actual ordinary practice of the subject. There’s no denying the importance of foundational
studies that show that everything we care about in mathematics can be done using only
sets or categories or types, as in the recent work on homotopy type theory (HoTT).\textsuperscript{54} But
the everyday practice of mathematics doesn’t generally concern itself with these kinds of
issues. “To say that the real numbers do not constitute an ‘organism’ of the same kind as
the rationals is to make a reasonable and defensible observation, soberly construed, about
what we do” in ordinary practice,\textsuperscript{55} and this is just the kind of thing that the realist position
advocated here recommends saying. When we look at the “normal science” of mathematics,
we undeniably see different goals and methods, and this fact should be taken seriously as
we approach the subject from a philosophical angle.

So, for example, number theory is out to investigate a potentially infinite sequence;
algebra is in the business of identifying and studying structural features shared by many
mathematical objects; analysis grew out of scientific applications and focuses on real
number spaces and their generalizations; at least one aim of set theory is to provide a
certain kind of foundation for classical mathematics. Each of these endeavors involves
different methods, different ways of thinking, and these appear to be differences ‘in the
math’ (Ernst \textit{et al.} 2015: 159).

If we want to understand mathematics as it really is, then, a synthetic approach to the subject
seems the best way to proceed.\textsuperscript{56}

\textsuperscript{53} Again there is some connection with Friend’s mathematical pluralism here. The “pluralist in methodology,”
according to that view, “is tolerant towards proof techniques, methods and results being imported from one
area of mathematics into another. He is also not averse to the suggestion that techniques in disciplines outside
mathematics can be useful to mathematics and the philosophy of mathematics” (Friend 2014: 26). See also
the “algebraic” approach discussed by (Leng 2009).
\textsuperscript{54} See, e.g., (Univalent Foundations Program 2013).
\textsuperscript{55} (Floyd 2015: 263)
\textsuperscript{56} There are some similarities between my use of ‘synthetic’ and the interesting, but somewhat idiosyncratic,
view presented in (Zalamea 2009). Zalamea stresses the importance of local interpretations of pieces of
mathematics and the role of category theory in axiomatizing regions of mathematical practice, as does the
realist as understood here. However, he also sees the term synthetic as pointing “to a veiled reality that

19
As I mentioned when I first broached the topic of the motley of mathematics, this view—and with it, presumably, the synthetic approach—is controversial among both philosophers and mathematicians. I suggested then that the concerns about Wittgenstein’s ideas regarding the mixture of proof methods in mathematical practice might come down to the thought that he is denying a fundamental unity to the subject—a unity these objectors believe must be respected. I suggest that the objection to Wittgenstein and the feeling that there is real conflict here may be based on giving too much metaphysical weight to both the idea that mathematics is a motley and the idea that at bottom the motley is “all just mathematics.”

To say that mathematics is a patchwork of proof methods and systems isn’t to deny that there are connections and similarities of methods and questions to be found all over the subject, often in surprising places. Nor is it to deny that some of the most important and deep problems of contemporary research can’t be nicely separated into group theory or analysis or topology alone. In fact, one of the things that most interests Wittgenstein about mathematics—and that inevitably interests anyone acquainted with the subject at all—is the myriad and unexpected ways in which what may be (or may appear to be) different areas of mathematics come to be linked together and integrated over time. Wittgenstein himself often makes the further claim that these links and connections are created rather than discovered, and this does imply that there isn’t a pre-established world of mathematics revealed by mathematical research. In my estimation, it’s this further claim that most rankles when it comes to this topic. However, one need not make such a strong further claim when being guided by the thought that viewing mathematics as a multicolored mixture of proof

methods is a useful perspective to take. The further, stronger claim is suggested, not just by Wittgenstein, but by the additional tenet of methodological realism that tries to make minimal philosophical assumptions whenever possible, but it’s not required or something that must be enforced for those who wish to take a realist view.

‘Use’ in Mathematics

I expect that the general principle of allowing the meaning of a term or proposition to be illuminated by how it’s used is clear enough and familiar enough from Wittgenstein’s writings to be fairly brief with it here. Additionally, the idea is one that seems to have already appeared broadly plausible to mathematicians and philosophers of mathematics. For example, William Thurston, another Fields medalist reflecting on his subject, suggests that the language of mathematics “is not alive except to those who use it,”\(^\text{59}\) and Stewart Shapiro and Hilary Putnam both subscribe to a “Use Thesis” that is Wittgensteinian in content.\(^\text{60}\) However, there are still a few points worth emphasizing about this tenet of the realist methodology before moving on.

First, there’s a common misconception about how appeals to use are supposed to function within the so-called ordinary language philosophy inspired by Wittgenstein’s work that should be avoided as we move forward. It’s often thought that appeals to use do nothing more than show what we would or wouldn’t say in a given situation and, further, that those kinds of facts have (or should have) no bearing on philosophical questions. Who cares, the thought goes, about whether or not it would be strange to say, e.g., “I know I have hands” at almost any given moment? This objection to appeals to use can be avoided by making clear the fact that such appeals are not intended to simply list what we happen to say. Rather, they

\(^{59}\) (Thurston 2006: 167)
\(^{60}\) See (Shapiro 1991: 212) and (Putnam 1980).
are intended to help us decide what it would make sense to say. This way of thinking about
attention to usage is especially clear in the work of Avner Baz, who expands on this view when he writes the following.

I should emphasize at the outset that, in this book, ‘use’ will be used as I believe Wittgenstein uses it, to refer to a certain kind of human achievement, however humble and everyday—one that contrasts not with mentioning the words, but with letting them idle, or failing to do any (real) work with them.61 One thing this means, and this is an important point to which I will return, is that whether certain uttered words are actually being used on the occasion of their utterance, inside or outside philosophy, and if so how, is never a straightforward empirical matter.62

Saying “I know I have hands” doesn’t achieve the kind of work that “knowledge”-language is supposed to achieve—e.g., assuring someone in doubt—when it’s uttered in most ordinary circumstances. It’s this fact that (on some views) make it something that doesn’t make sense to say in most cases and makes it not a report of real knowledge; not the fact that it seems like a strange thing to say most of the time.

The second thing to notice is the fact that the language of mathematics isn’t always used like an ordinary language is used, so there may be some question about what this language’s being in use is supposed to look like. This is a reasonable concern, but some ground can be gained towards understanding mathematical language in use by employing the via negativa.

When doing ordinary mathematics and logic, giving an “interpretation” of a language means setting up a map between the symbols of the language and some appropriate mathematical structure. Given this way of thinking about interpretations, it’s easy to think that questions about meaning and reference can be investigated using these simple model-theoretic methods. However, as (Mühlhölzer 2014) rightly points out, when we give an interpretation, we precisely aren’t using the language in question—we’re constructing a

---

61 I’ve omitted a long footnote from the passage at this point.
62 (Baz 2012: 2, emphasis in the original)
further mathematical object. If that’s correct and if meaning essentially has to do with use, then this kind of construction has nothing at all to do with meaning. One way to focus on the importance of use in mathematical practice, then, is to be careful not to allow talk of interpretations in the model-theoretic sense to be uncritically used to settle questions about meaning and other closely related issues. More generally, it may be easier to tell when a language is *not* being used than to give a full account of what exactly being in use requires or entails. Fortunately, in many cases this is enough.

Finally, it’s worth remembering that not every question about use needs to be put simply in terms of whether or not a sign or expression is being used. We can also ask about *how* these linguistic items are put to use. For example, perhaps upon close examination, the axioms that define a particular area of inquiry look as if they function like rules for the use of the terms involved, as is concluded in (Friederich 2011). Maybe certain theorems appear to function in this way as well when they’re put to use. For example, the Bolzano-Weierstraß theorem states that any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. This theorem is often used in a way that makes it seem like a rule that licenses one to conclude that such a convergent subsequence is already in hand. If we pay close attention to the ways in which theorems and axioms and so on are put to use in ordinary circumstances, we should be able to command a clearer view of the subject and its practices, again even without a fully worked out theory of mathematical use.

In fact, attention to use can motivate the realist to understand mathematical practice and what it reveals in much the same way Alasdair MacIntyre concerns himself with practices in his work.

By a ‘practice’ I am going to mean any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that
human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended.\textsuperscript{63}

The goods internal to an end mentioned in this quotation are supposed to be specifiable and achievable only in terms of the given practice. So, for example, one may achieve fame and fortune by becoming a mathematician, but those goods are also achievable by other means external to the practice of mathematics. However, the ability to appreciate the ingenuity of a particular idea employed in a proof is a good that can only be specified by reference to the practice mathematics and can only be achieved by those who are willing to enter into the practice at least to some extent.\textsuperscript{64} One comes to have the ability to appreciate various characteristics of important theorems or brilliant arguments or surprising results by spending time aiming to achieve the standards of excellence for significant or surprising theorems and beautiful proofs that partially define what it is to do mathematics. While it’s possible, as G.H. Hardy suggests, for “any intelligent reader, however slender his mathematical equipment” to understand Euclid’s proof of the infinitude of primes in a short amount of time, the ability to understand why a mathematician would judge it to be “first-rate” or “deep” is not so quickly attainable.\textsuperscript{65} By paying attention to the way things are really used and appreciated within mathematical practice, we can hope to have these internal goods and goals open up to us and thereby have a more complete and realistic understanding of the subject.

I will conclude this chapter by putting these three methodological recommendations to work in a simple example that fairly recently captured the popular imagination.

\textsuperscript{63} (MacIntyre 1981: 187)
\textsuperscript{64} Cf. (MacIntyre 1981: 188).
\textsuperscript{65} See (Hardy 1940: 18).
Example: \(-1/12 = 1 + 2 + 3 + \ldots\)

A few years ago, a video with the title “ASTOUNDING: 1+2+3+4+5+\ldots = -1/12” made its way around the Internet,\(^{66}\) and as of this writing, it’s been viewed over 5.5 million times. The video and the ideas illustrated in it provide a nice subject upon which to demonstrate how the methodological principles described in this chapter can work together to clear up confusion. It’s important to note, however, that I am not claiming that mathematicians would view the proof provided in this video as unexceptional—on the contrary, many quickly came forward to point out problems with it—nor am I unaware that the purpose of the video may have only been to get more people interested in mathematics in whatever way might work best. Nevertheless, “toy” examples can be illuminating, and, in this case, the kind of issues that the realist methodology elaborated in this chapter attempts to target are sufficiently highlighted by putting the methods to work here that the example is worth brief consideration.

Our video begins by raising the question, “What is the sum of the following series?"

\[
S = 1 + 2 + 3 + \ldots
\]

The viewer is supposed to have the intuition that adding larger and larger numbers together forever is going to make \(S = \infty\) or to make \(S\) greater than any given number. Yet, we’re quickly informed that the real answer is \(S = -1/12\). Any doubts about the veracity of this answer are then immediately scattered by the host’s pointing to the expression

\[
\sum_{i=1}^{\infty} i \rightarrow -1/12
\]

in a textbook on string theory.

To see why the sum of the series is \(-1/12\), consider the following two other series.

\(^{66}\)(Numberphile 2014)
\[ S_1 = 1 - 1 + 1 - 1 + \ldots \]
\[ S_2 = 1 - 2 + 3 - 4 + \ldots \]

If you take any of the partial sums of the series \( S_1 \), you'll either obtain the value 0 or 1, depending on whether you stop after an even or odd number of additions and subtractions. Since the infinite sum is of neither an even nor an odd length, however, it seems as if we ought to just take the average to these two values as the sum of the full series. That is,
\[ S_1 = 1/2 = 1 - 1 + 1 - 1 + \ldots \]

appears to be justified. Given this value, we can now calculate the value of \( S_2 \) by making the following observation.

\[ 2 \cdot S_2 = 1 - 2 + 3 - 4 + 5 - \ldots \]
\[ + 1 - 2 + 3 - 4 + \ldots \]
\[ = 1 - 1 + 1 - 1 + \ldots \]
\[ = S_1 \]

That is, \( S_2 = 1/4 \). Finally, the real value of \( S \) can be reached by using a similarly clever method of summing.

\[ S - S_2 = 1 + 2 + 3 + 4 + 5 + \ldots \]
\[ - 1 + 2 - 3 + 4 - 5 + \ldots \]
\[ = 0 + 4 + 0 + 8 + 0 + \ldots \]
\[ = 4 \cdot S \]

Solving the equation \( S - S_2 = 4 \cdot S \) for \( S \) using the fact that \( S_2 = 1/4 \), gives \( S = -1/12 \). This result surely is, if not astounding, at least surprising, but, as a science writer for the
New York Times commenting on the video sums it up, “This is what happens when you mess with infinity.”

The first thing to note about this video is the heavy burden that prose plays in the justification of the initial decision to say that $1 - 1 + 1 - \ldots$ sums to $1/2$. Clearly, no proof would involve an argument that ran, “Infinite sums aren’t of even or odd length, so we should take the average of the even- and odd-length sums to be the total sum.” A proof could begin with a step that looked like $1 - 1 + 1 - \ldots =_{df} 1/2$, but employing such a definition would make it immediately evident that the uses of $+$ and $-$ in the expression $1 - 1 + 1 - \ldots$ are not the same as the uses we are familiar with. Ordinarily, $+$ and $-$ only make sense in expressions of finite length, and when they are used in “infinitely-long expressions” they only have sense if the partial sums of the expressions converge. Neither of these conditions is applicable to the series $1 - 1 + 1 - \ldots$ though. Since a new use of these operations is obviously being defined at the outset in this rendering of the proof, there are no expectations that can be surprised by the revelation that $1 + 2 + 3 + \ldots = -1/12$. That is, doing away with the prose would quickly reveal that there’s actually nothing “astounding” going on in the video at all.

This initial observation about the video makes it tempting to suggest that the questionable move in the proof on offer is the decision to let $1 - 1 + 1 - \ldots = 1/2$. However, as is so often the case, it’s actually the initial question, “What is the sum $1 + 2 + 3 + \ldots$?” that leads us astray. As Wittgenstein notes in a different context, “The first step is the one that altogether escapes notice”: after taking this step, “[t]he decisive movement in the conjuring trick has been made, and it was the very one that we thought quite innocent.” From childhood onward, being asked for the sum of two numbers is a common enough occurrence that we’re naturally tempted to think that, just as questions about ordinary sums have a correct

---

67 (Overbye 2014)
68 (Wittgenstein 1953/2009: §308)
answer and need no further specification, so must the sum \( 1 - 1 + 1 - \ldots \). That is, we are tempted to overlook the fact that an infinite sum like the one under consideration has not been given any clear meaning as of yet. It was an important step forward in the history of mathematics for mathematicians to realize facts about the need for definitions like this, so it’s not surprising that we still make such mistakes. In Hardy’s estimation, for example, “it is broadly true to say that mathematicians before Cauchy asked not ‘How shall we define \( 1 - 1 + 1 - \ldots \)?’ but ‘What is \( 1 - 1 + 1 - \ldots \)?’”\(^{69}\) There are certainly more or less natural choices for how \( 1 - 1 + 1 - \ldots \) should be defined, but this doesn’t change the fact that it is defined and not a value simply to be found.\(^{70}\)

The decisions to focus on proof rather than prose and to pay special attention to how the signs + and − are being used, then, immediately clarify some of the issues raised by the video under discussion. Viewing mathematics as a multicolored mixture of proof methods brings further clarity. For certain purposes (for example, when working with sums and series in real analysis and the theory of integration), there may not be much point in defining a value for the sum of a series like \( 1 + 2 + 3 + \ldots \). However, subjects like analytic number theory and combinatorics make heavy use of so-called Dirichlet series, an important example of which appears below.

\[
L(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \ldots
\]

This series formally evaluated at \( s = -1 \) is

\[
L(-1) = \frac{1}{1^{-1}} + \frac{1}{2^{-1}} + \frac{1}{3^{-1}} + \ldots = 1 + 2 + 3 + \ldots,
\]

so the importance of Dirichlet series already goes some of the way towards explaining why coming up with values for series such as \( 1 + 2 + 3 + \ldots \) might be worth the effort in certain parts of mathematics.

\(^{69}\) (Hardy 1949: 6, emphasis in the original)

\(^{70}\) See, again, (Hardy 1949: 2, 6).
The Dirichlet series $L(s)$ is only convergent and, therefore, meaningful in the ordinary sense for complex numbers $s$ with $\Re(s) > 1$. For such values, $L(s)$ is just the famous Riemann $\zeta$-function.\(^7\)\(^1\) It’s a natural goal of mathematical practice to extend the domain of a function as important as the $\zeta$-function to as much of the complex plane as is possible. And it turns out that the process of “analytic continuation” allows the $\zeta$-function to be extended to all of the complex plane aside from $s = 1$. This extended $\zeta$-function has the value $-1/12$ at $s = -1$, and this fact provides a motivation for taking the sum of the series $1 + 2 + 3 + \ldots$ to be $-1/12$ since this function has the same value as $L(s)$ wherever both are defined.\(^7\)\(^2\) It’s important to note, though, that however many reasons there are for saying that $1 + 2 + 3 + \ldots = -1/12$, the question of what this infinite series sums to must first be given sense by defining what is meant by taking the sum of such a series. The sense that will end up being given to such sums will generally be determined by the particular goals and interests of one or more of the many multicolored pieces of mathematics and its practice.

The remainder of this dissertation is concerned with demonstrating that the clarity which the three main tenets of the realist methodology outlined in this chapter bring to this simple example can also be attained by applying these same methods to more complex problems and puzzles. To this project, I shall now turn.

---

\(^7\)\(^1\) See (Apostol 1976: Ch. 12).

\(^7\)\(^2\) Another argument for making this decision can be given using the methods of so-called Abel summation. See, e.g., (Stopple 2003: §8.2).
Chapter 2

On the Mystery of the Real Power Series

There is a mystery about this which stimulates the imagination; where there is no imagination there is no horror.
—Sherlock Holmes in *A Study in Scarlet*

Early in his popular book *Visual Complex Analysis*, Tristan Needham investigates what he calls “The Mystery of the Real Power Series,” which involves the relation between the function

\[ x \mapsto \frac{1}{1 + x^2} : \mathbb{R} \to \mathbb{R} \]

and the interval over which its Maclaurin series converges.\(^1\) Both because of its mathematical simplicity and because its solution is generally taken to offer a straightforward example of the mathematical explanation of a mathematical phenomenon, this so-called mystery has repeatedly drawn the attention of philosophers. Mathematicians have also generally shared the opinion that the facts of the case generate a mystery, and that the mystery’s resolution is a clear instance of mathematical explanation. This kind of widespread agreement makes the example an important one: for either we have before us a solid piece of data that can guide attempts to understand the nature of explanation in mathematics or a case that easily

\(^1\) (Needham 1997: 64-67)
misleads and may therefore serve as a stumbling block rather than a stepping stone on our way towards this understanding. This being the case, Mark Steiner is right to insist that “[t]his example deserves close consideration.”\(^2\) He is also right to suggest that this close consideration “yields many rich insights.”\(^3\) I hope to show in what follows, however, that the insights yielded differ significantly from what many—including Steiner himself—have supposed. In particular, I shall aim to illustrate how a calmer inspection of the facts of the case, the kind of inspection motivated by the realism outlined in Chapter 1, can rid us of the mystery and our horror by dissolving the aspects of the case most prone to stimulate our easily excitable imaginations.

I shall proceed as follows. Sections 1 and 2 respectively present the alleged mystery and its standard resolution. Then, in Section 3, I respond to a preliminary concern about the very project of the paper: namely, that it should not be undertaken by a philosopher with naturalist sympathies. The reexamination of our mystery begins in Section 4, where I explore the question of what exactly is wanting explanation in this case. It turns out that deciding whether or not we have a real mystery on our hands requires some consideration of the status of mathematical questions asked across varying contexts. Section 5, therefore, discusses some of the issues that arise regarding such questions. The theme of rethinking the standard assessment of the Mystery of the Real Power Series is then carried over into Section 6, where I question the status of the case’s standard resolution as an instance of mathematical explanation. After an overview and criticism of accounts of the mystery from the perspectives of the two most well-known theories of mathematical explanation in Section 7, I conclude in Section 8 by offering an explanation of the prevalence of the standard view of this case and making some suggestions about what conclusions can legitimately be drawn from our mystery.

\(^2\) (Steiner 1978b: 19)  
\(^3\) Ibid., 19.
The Mystery

Although Needham describes the Mystery of the Real Power Series more vividly than some, mathematicians and philosophers alike generally present the case in the same terms. Any real function has a Taylor series expansion about any point at which it has derivatives of all orders. In general, however, this power series will only converge to the original function within some finite interval containing the point of expansion. The real function

\[ f(x) := \frac{1}{1 + x^2} \]

is everywhere infinitely differentiable, so it has a Taylor series expansion in particular about zero. This series begins

\[ 1 - x^2 + x^4 - x^6 + \ldots, \]

and converges for \( |x| < 1 \); i.e.,

\[ \frac{1}{1 + x^2} = 1 - x^2 + x^4 - x^6 + \ldots \]

only for these values of \( x \).

But many have found this puzzling. Why is there convergence for only these values? After all, \( f \) “is a beautiful function of the real variable \( x \)” that is everywhere bounded, and the graph of the function below fails to suggest any obvious point at which the series should be expected to diverge. “There appears to be nothing in the nature of the function to account for” the convergence behavior of \( f \)’s Maclaurin series.  

---

\[ ^4 \text{For philosophers’ accounts, see, e.g., (Lange 2017: 290-292), (Leng 2011: 68), (Steiner 1978b: 18-19), (Waismann 1954/1982: 29-30), and (Wilson 2006: 313-314). For mathematicians’, see (Freitag and Busam 2009: 3), (Gamelin 2001: 146), (Markushevich 1965: 9), and (Spivak 1994: 508). The account I give here is essentially a composite of these sources.} \]

\[ ^5 \text{(Gamelin 2001: 146)} \]

\[ ^6 \text{(Ponnusamy and Silverman 2006: 188)} \]

\[ ^7 \text{“Maclaurin series” is just another name for the Taylor series expansion about zero.} \]
The Maclaurin series of a function like

\[ g(x) := \frac{1}{1 - x^2}, \]

on the other hand, is naturally expected to diverge outside of the interval \((-1, 1)\) because the function itself is unbounded as \(x\) approaches \(\pm 1\), as illustrated below.

If the series were to converge to \(g\) for any value of \(x\) greater than 1 or less than \(-1\), this would imply convergence at \(\pm 1\), which is impossible.

Thus, there seems to be a good explanation for why \(g\)’s Maclaurin series diverges where it does, but none at all for why \(f\)’s does. This is our mystery.
The Mystery, Solved

Jacques Hadamard famously noted that “the shortest and best way between two truths of the real domain often passes through the imaginary one.” Our mystery appears to be an instance in which Hadamard’s saying rings true since it is supposed to be solved by making an excursion into the complex plane.

If instead of considering $f$ as a function of the real variable $x$, we examine the function

$$f(z) := \frac{1}{1 + z^2}$$

with domain and codomain $\mathbb{C}$, the convergence behavior of $f$’s power series expansion seems to start making sense. This is because $f(z)$ has poles at $\pm i$. Just as the Maclaurin series of the function $g(x)$ was expected to diverge unless $|x| < 1$ because of the singularities at $\pm 1$, $f(z)$’s series is now expected to diverge outside the circle $|z| = 1$. This is because, again, convergence for any value of $z$ such that $|z| > 1$ would imply convergence at $\pm i$, which is impossible. So, if the Maclaurin series of our original function had converged to $1/(1 + x^2)$ for any $x$ with $|x| > 1$, this fact would have implied an impossible convergence in the complex plane. And this impossibility is thought to explain why the series does not converge to $f(x)$ for any such $x$.

Here then is the standard resolution of the Mystery of the Real Power Series. The behavior of the function’s Taylor series expansion about zero is explained “fully and without residue” by its behavior in the complex plane.\(^9\) It almost looks “as if the real function

---

\(^8\) This remark is often attributed to Hadamard (1949: 123), but he actually only endorses it, saying that “It has been written that […].” The original author is Paul Painlevé (1900: 1-2); cited, e.g., in (Lax and Zalcman 2012: x).

\(^9\) (Waismann 1954/1982: 30)
already knew that the complex numbers were there,”¹⁰ and that, in fact, the behavior of \( f \) is simply “trying to tell us about the existence of the complex plane.”¹¹

**Naturalist Worries**

Before moving on to raise some questions about this mystery and its resolution, it will be useful to pause and respond to a potential worry:

If there’s widespread agreement about how to think about this case among philosophers and mathematicians, what right does a philosopher have to question these conclusions? Doesn’t this commit the sin of offending against naturalism?

I contend that no offense against naturalism is committed in what follows, and it’s important to be clear about why that is so.

**Expanding on the Worry**

As has already been noted, one prominent trend in recent philosophy of mathematics is an increased attention to what mathematicians actually do. Rather than concerning themselves exclusively with worries about foundations or the existence of numbers and how we might come to know about them, philosophers working in this tradition think that “attention to mathematical practice is a necessary condition for a renewal of the philosophy of mathematics,”¹² and that “only detailed analysis and reconstruction of large and significant parts of mathematical practice can provide a philosophy of mathematics worth its name.”¹³ A primary goal of this school is, therefore, to engage in such analyses and reconstruction.

Those contributing to this kind of work also generally see themselves as staunch naturalists. Naturalism comes in many varieties, but the strain that seems to be most commonly

---

¹⁰ (Leng 2011: 68)  
¹¹ (Needham 1997: 65)  
¹² (Mancosu 2008a: 2)  
¹³ (Mancosu 2008a: 5)
cultivated among philosophers working within this tradition has been characterized especially well by John Burgess. This sort of naturalist denies “that philosophy has any access to exterior, ulterior, and superior sources of knowledge from which to “correct” science and scientifically informed common sense.”

Again according to Burgess, naturalists conceive of epistemology as an “inquiry conducted by citizens of the scientific community examining science from the inside, rather than an inquisition conducted by philosophers foreign to science judging science from the outside.” In other words, these naturalists are expected to take David Lewis’s Philosophical Humility Credo very seriously.

These two marked features of much work in contemporary philosophy of mathematics can suggest that the general schema for good work in the field is (i) do a case study looking at some area of mathematics and (ii) draw a few general lessons from what was observed. There’s no doubt that this kind of work is a good idea and can be important, but, without some subtlety, it can lead to an unjustified squeezing out of room for any criticism of the practices one observes. Criticism of the kind to be offered below in particular would seem to find no room.

Responding to the Worry

The subtlety that is needed, and which answers the worry raised above, is simply taking care to distinguish a critique of the mathematics involved in cases of interest from a critique of the language found surrounding it.

---

14 (Burgess 2008: 2)
15 (Burgess 2004a: 19-20)
16 I.e., that rejecting any mathematics for philosophical reasons borders on the absurd. See (Lewis 1991: 57-59).
17 As already noted in Chapter 1, it can’t be assumed that the mathematics and the surrounding language can be separated as simply as using Proof…Q.E.D. as the dividing line. Certainly, this approach to drawing the boundaries is a natural starting place, but it appears ultimately to be too naive.
It’s interesting to note that whenever we find talk of explanation in mathematics, it occurs in the context of what Hardy and Littlewood refer to as ‘gas’\(^\text{18}\) or what Wittgenstein would call ‘prose’.\(^\text{19}\) For example, in Johannes Hafner and Paolo Mancosu’s extensive sampling of the varieties of mathematical explanation,\(^\text{20}\) all the cited instances of talk about mathematical explanation occur outside the body of a proof. This is true even where a philosopher or mathematician happens to take a proof to offer an explanation of the proposition it proves.

While we can be confident that the norms of proof will do the work of keeping any observed mathematics in line, the mathematician’s prose has freer rein outside the confines of these norms. And because this type of musing is less strictly controlled, the prose and interpretations of the mathematician naturally may fall into the kinds of errors that lend themselves to philosophical discussion and critique. One often finds that the terminology and images mathematicians appeal to when considering the more philosophical implications of their results are transferred to mathematical settings by analogy with their uses in other sciences.\(^\text{21}\) Since these tools for interpretation tend not to be comfortably at home in their new setting, however, the prose couched in terms of them may very well be in need of the same kind of conceptual clarification that other potentially misleading uses of language demand. At least there is no good reason to presuppose this not to be the case. Of course, anyone making a critique of the interpretation of a piece of mathematics must take care to understand the work being surrounded by that prose, and, of course, the mathematician’s interpretation must itself be taken seriously. But to avoid critical discussion of the prose

\(^{18}\) Hardy (1929: 18) explains what he and Littlewood mean by ‘gas’: “[R]hetorical flourishes designed to affect psychology, pictures on the board in the lecture, devices to stimulate the imagination of pupils.”

\(^{19}\) Cf., for example, the following remark of Wittgenstein’s from a conversation with Waismann, “What’s needed is an analysis of mathematical discourse that separates the actual mathematics from the empty talk. [...] It is very important to distinguish as strictly as possible between the calculus and this kind of prose” (Waismann 1929-32/1979: 149).

\(^{20}\) (Hafner and Mancosu 2005)

\(^{21}\) See, e.g., (Shaheen 2017b) for more on this topic in other “non-causal” settings.
surrounding mathematical proofs would be to avoid doing, arguably, some of the main work philosophers of mathematics ought to be doing.

None of the criticism that follows attributes any mathematical error or ignorance to mathematicians or philosophers writing about the Mystery of the Real Power Series. To do so would be absurd (or, at best, out of place) since all of the mathematics involved is elementary. Nevertheless, my contention will be, in part, that the acceptance of a bad analogy between explanation in mathematics and explanation in the physical sciences has led mathematicians and philosophers to make philosophical errors in their thinking about the case under discussion here. That philosophers and mathematicians both quickly lapse into platonist imagery and metaphors when discussing the consequences of this mystery and its resolution shows that some straying from the domain of the purely mathematical has taken place. It turns out that this kind of straying is in need of reinsing in this case.\textsuperscript{22} Reexamining the agreed-upon mathematical facts of the case can show us the way to avoid the unnecessary importation of metaphysics and resolve our self-created philosophical puzzlement.

Reconsidering the Question

When faced with any difficulty, a typical philosopher’s move is to begin with a reexamination of the question posed. I’ll begin this section by making just such a move: What exactly is the mystery that needs solving and what is it that really wants explanation here?

The presentation of our mystery in Section 1 ended by asking for some explanation of what appeared to be puzzling convergence behavior for \( f \)’s Maclaurin series. Given that this series only converges to \( f \) within the interval \( |x| < 1 \), and given the fact that if the series converges at some point \( c \) it must also converge at any point \( x \) where \( |x| < c \), our

\[ \text{That is, we ought not think too much about a kind of pre-established harmony between the worlds of the real and the complex numbers, or about the imaginary numbers crying out in hopes of discovery.} \]
original question comes to, “Why doesn’t $f$’s Maclaurin series converge to $1/(1 + x^2)$ at ±1?” There’s good reason to think that this is not a mystery at all, however.

Recall that the Maclaurin expansion of $f$ about 0 is

$$1 - x^2 + x^4 - x^6 + \ldots .$$

Formally evaluated at $x = 1$, this expression is

$$1 - 1 + 1 - 1 + \ldots ,$$

which, of course, doesn’t converge in the sense of having converging partial sums.\(^{23}\) If $f$’s Maclaurin series were to converge anywhere outside the interval $(-1, 1)$, this would imply convergence at $x = 1$ or $x = -1$, which doesn’t occur, as was just very easily shown. But, to think that this presents us with a mystery is analogous to thinking that there’s a mystery about why $f$ has the Maclaurin series it does. There can be no mystery about that though since $f$ and its Maclaurin series are related by definition. This being the case, there’s no intelligible question to ask about why this series in particular is $f$’s.

What makes it look like we’re faced with a mystery here is the fact that it can easily seem as if it makes sense to ask regarding any real, infinitely differentiable function, “What is causing its Taylor expansion to fail to converge outside of its interval of convergence?” or as Michael Spivak puts it regarding the function $f$, “What unseen obstacle prevents the Taylor series from extending past 1 and −1?”\(^{24}\) That it’s something very close to this question that inspires the puzzlement in this case becomes clear when we take into consideration the fact that $1/(1 + x^2)$ is always compared to $1/(1 - x^2)$ when the mystery is presented. The most obvious difference between these two functions is that one has real singularities,

\(^{23}\) Although many years ago there were a variety of arguments given in favor of different convergent values for this series; e.g. 0, 1/2, or 1.

\(^{24}\) (Spivak 1994: 522)
while the other has none. “Well,” we’re tempted to ask, “if there’s no singularity to speak of in the $1/(1 + x^2)$ case, why the failure of convergence? What’s to blame?” But the failure of $f$’s Maclaurin series to converge outside of $(−1, 1)$ can be taken to show precisely that this question is not always a legitimate one in the domain of real functions despite its *prima facie* seeming to be: $f$ has no singularities and still fails to converge outside of $(−1, 1)$. It might be thought, therefore, that if $f$ and its convergence behavior are trying to teach us anything, it’s not about the existence of the complex plane, but about presupposition failure and the existence of pseudo-questions in relation to certain real functions.

Taking another tack then, perhaps the question that sets us off investigating our mystery has been posed too narrowly so far. I have only presented the Mystery of the Real Power Series as a mystery about the convergence behavior of $f$’s Maclaurin series, but $f$ has a Taylor series expansion centered around every other real number as well. We may, therefore, find a more serious mystery lurking outside of the confines of the expansion about zero.  

In order to see what other questions emerge in the more general case, we need to consider the general Taylor series expansion of $f$ around a point $c$, which is given by

$$
\sum_{n=0}^{\infty} a_n (x - c)^n, \text{ where } a_n = \frac{f^{(n)}(c)}{n!}.
$$

The Cauchy-Hadamard theorem, which will be discussed more below, says that the radius of convergence for $f$’s Taylor series expansion about $c$, denoted $R$, is given by

$$
\frac{1}{R} = \limsup_{n \to \infty} |a_n|^{1/n}.
$$

Using the following formula for the $n^{th}$ derivative of $f$ in this case

---

25 See (Needham 1997: 65-67) for this extended mystery.
\[
\frac{d^n}{dx^n} \left( \frac{1}{1 + x^2} \right) = \sum_{k=0}^{n} (-1)^k \frac{(2x)^{2k-n}k!(1 + 2k - n)^{2(n-k)}}{(n - k)!(1 + x^2)^{k+1}},
\]

where \(p^n\) is the rising factorial, or Pochhammer symbol,\(^{26}\) we can find the general coefficient of the Taylor expansion of \(f\) around \(c\):

\[
a_n = \frac{1}{n!} \sum_{k=0}^{n} (-1)^k \frac{(2c)^{2k-n}k!(1 + 2k - n)^{2(n-k)}}{(n - k)!(1 + c^2)^{k+1}}.
\]

Computing the values of \(|a_n|^{1/n}\) for \(n\), say, up to 1,000 shows that the sequence approaches

\[
\frac{1}{\sqrt{1 + c^2}}
\]

fairly quickly. This suggests that, in general, the radius of convergence of \(f\)'s Taylor series expansion about \(c\) is \(\sqrt{1 + c^2}\). We may, therefore, ask, “Why is the radius of convergence given by this particular number?” and hope to find a true mystery there.

This new question must be approached while taking into account two different possibilities regarding the calculability of the limit involved in the Cauchy-Hadamard theorem, however. First, suppose that this limit is calculable using only real methods.\(^{27}\) In such a case, there again appears to be no reason to be puzzled about where the number \(\sqrt{1 + c^2}\) has come from. If real methods are sufficient for proving that the radius of convergence is given by \(\sqrt{1 + c^2}\), then we simply had a function, managed to find an explicit formula for the coefficients of its Taylor expansion, and then found the radius of convergence for this

\(^{26}\) The Pochhammer symbol is sometimes also written as \((p)_n\) or \(p^{(n)}\), and is defined as follows.

\[
p^n = \begin{cases} 
1 & \text{if } n = 0 \\
(p+1) \cdots (p+n-1) & \text{if } n > 0,
\end{cases}
\]

\(^{27}\) Whether it really is or is not calculable with real methods is not especially of importance here since the considerations to be worked through are worthy of discussion for cases where real methods are known not to suffice. For example, (van der Waerden 1949: 179-180) shows that there is no algebraic formula for finding the roots of a cubic equation that doesn't give imaginary values for some values of the coefficients. At any rate, the difficulty in determining the limit without complex methods here stems from the fact that the usual tricks for doing the work, e.g., partial fraction decompositions or so-called Cauchy estimates, are not available.
series, all using only real methods. In such a case, therefore, it’s not at all clear that we are left with some kind of behavior that lacks an explanation. All that we were in search of is laid before us for examination.\(^\text{28}\)

Next, suppose that real methods are not sufficient for actually proving that \(\sqrt{1 + c^2}\) is the radius of convergence here. Is it necessarily the case that there’s a mystery remaining to be solved? That is, we ought to ask why it’s thought to be the case that “the asymptotics of the coefficients of \(f\)’s Taylor series” is not sufficient to explain the observed convergence behavior. When considering any sort of mathematical object or relation, we have to keep in mind that sometimes asking for an explanation can lead the disappointing discovery that none (or no very satisfying one) exists. Those philosophers of mathematics who have recently been interested in the phenomena of “mathematical coincidences” should be especially alive to this fact.\(^\text{29, 30}\) The mathematician Michael Spivak also makes note of this possibility when introducing the Mystery of the Real Power Series in his *Calculus* (1994: 508): “Asking this sort of \[ why\]-question is always dangerous, since we may have to settle for an unsympathetic answer: it happens because it happens—that’s the way things are!” If it turns out that this is just the way things are with this function and its Taylor series’s convergence behavior, then there is—by definition—nothing to be explained and no real mystery to be solved. But, of course, no one writing about the Mystery of the Real Power Series thinks that this \emph{is} just the way things are because all such authors already have the standard solution to the mystery in mind. Each of these authors points out the intended-to-be-surprising fact that, “although the question is about real numbers, it can be answered intelligently only when placed in a broader context.”\(^\text{31}\) This, then, naturally leads us to consider what happens to mathematical questions when they are placed in broader contexts.

---

\(^{28}\) Cf. (Shanker 1987: 337).

\(^{29}\) See, for example, (Baker 2009), (Davis 1981), (Lange 2010), and (Lange 2014).

\(^{30}\) The subject of coincidence and its role in mathematical practice will be addressed further in Chapter 4.

\(^{31}\) (Spivak 1994: 508)
Domain-crossing Questions

It’s the broader context of complex analysis which apparently allows the question that has the best chance of presenting us with a mystery to be asked in a domain where it has a very satisfying answer.

It’s a theorem of complex analysis, first noted by Cauchy, that a complex function’s power series expansion about a point converges within a radius determined by the nearest singularity to that point. So, in relation to the function \( f(z) \), we can ask the question that turns out to be illegitimate with respect to \( f(x) \) over the reals: “What singularity makes it the case that the power series doesn’t converge for all \( z \) (if it doesn’t) and how far is that singularity from our chosen expansion point?” The Taylor expansion of \( f(z) \) about a point \( c \) with no imaginary part will converge within a radius determined by the poles of the function at \( \pm i \), a radius given by the formula \( \sqrt{1 + c^2} \). This can be proved immediately and with very little trouble.

\[ \sqrt{1 + c^2} \]

\[ c \]

---


33 ‘Singularity’ here includes branch points to account for functions like \( \sqrt{1 + z^2} \) without poles, but whose Maclaurin series only converge within a finite radius.
The main question that must be raised in response to this determination of the radius of convergence of our function’s Taylor expansion about $c$ is whether or not this move to the complex domain answers the question that was originally posed. I’ll argue that it does not.

In order to have a convenient way to refer to the question(s) under consideration, I will call the original question about $f(x)$’s convergence behavior over the reals “the Real Question” and the most recent question about $f(z)$ “the Complex Question.” So, I am asking whether ‘the Real Question’ and ‘the Complex Question’ pick out the same question here.

Historically, mathematicians seem to have found this kind of “passage from the real to the imaginary,” most famously practiced in its early form by Laplace and Poisson, to be a means of addressing the very same question in a different context, although with techniques that could only be counted as useful for discovery or for suggesting truths inductively. That is, the imaginary results could provide evidence for what one should try to actually establish by direct proof, but could not yet be counted as being so established. It was not until Cauchy’s work providing rigorous foundations for complex function theory that the mathematical community appears to have started allowing complex methods to finally settle questions about real functions on their own. This initial discomfort at least suggests that there may be something funny going on when mathematical questions are transferred from one context to another. We needn’t accept that all that is questionable is so quickly resolvable through an increase in rigor of the kind Cauchy’s work has provided though.

Consider first the way most contemporary semanticists think about the meanings of questions. Starting with (Hamblin 1973) and (Karttunen 1977), understanding a question began to be thought of as being intimately related to understanding its possible answers. This approach to the semantics of questions is often inspired by an analogy to the semantics of indicative statements pursued by truth-conditional theories of semantics. According

---

34 See, e.g., (Bottazzini and Gray 2013: 95-98) and (Smithies 1997: 18-20).
35 See, e.g., (Cauchy 1814) and (Cauchy 1821/2009).
to such views, we understand indicative sentences when we know under what conditions they would be true. Interrogative sentences can then be taken to be understood when we know under what conditions they have been answered. Whatever one’s exact proposal for relating answers to questions in theorizing about the semantics of questions, this tradition of thinking about questions provides some reasons to think that the Real and Complex Questions are not in fact the same.

Whether the set of propositions our questions denote contain all possible answers (following the Hamblin approach) or all true answers (following Karttunen), it looks like we ought to think of the Real and Complex questions as picking out different sets. Consider first the Real Question. If someone is capable of seeing this question as posing a mystery, he or she must be capable of understanding the question. If that requires understanding the conditions under which the question has been answered—cached out here as having some kind of access to the set of possible answers or the set of all true answers—and the answer, “Because of the poles at $\pm i$,” is included in this set, then the Real Question can’t be understood without knowledge of the complex numbers. But that seems like wrong conclusion to draw here. It’s precisely because the person being presented with the Mystery of the Real Power Series doesn’t have access to answers involving the complex numbers that there’s even a chance of there being a mystery to investigate in the first place. To someone who has had even a basic introduction to the complex numbers, $1/(1 + z^2)$ and $1/(1 - z^2)$ would immediately be seen to have their Maclaurin series fail to converge for $|z| > 1$ for exactly the same reason: because of the function’s poles. These considerations suggest that we ought to take the Real Question to be looking for real (number) answers and the Complex Question to be looking for complex (number) answers. This being the case, if there’s something correct about the general view about the semantics of questions

36 Cf. (Groenendijk and Stokhof 1997: 1067).
employed here, we ought to conclude that the questions aren’t the same. In other words, the Real Question isn’t answered by pointing to facts about the complex plane.\textsuperscript{37}

The second most popular approach to the semantics of questions, the so-called force-radical view, suggests that interrogatives are formed by combining a “sentence radical” and a “force indicator.”\textsuperscript{38} According to this kind of theory, the radical of both the Real and Complex question can be taken to be something like the following.

$$[\lambda x.\text{reason}(x) : x \text{ makes the radius of convergence of } f \text{’s Taylor expansion what it is}]$$\textsuperscript{39}

A force indicator would then turn this radical into a question asking for the reason the Taylor expansion about \(c\) fails to converge outside the radius given by \(\sqrt{1 + c^2}\). Here the problem with identifying the Real and Complex Questions doesn’t have to do with a particular set of answers to the questions, but with the domain of reasons which can combine with the sentence radical to form a proposition. For reasons similar to those presented in relation to the previous semantic view about questions, we shouldn’t want to include reasons involving the complex numbers in this domain of reasons for just anyone able to comprehend the Real Question. This domain is supposed to include the reasons that the person who understands the question could combine with the sentence radical above to give an answer. If we want to allow a speaker unfamiliar with the complex numbers to be able to understand the Real Question and reasons about the complex numbers are included in this domain, such a speaker would somehow be in the position of being able to give answers she literally couldn’t understand to a question like this one. Thus, on this view of the semantics of ques-

\textsuperscript{37} For reasons unrelated to the ones canvassed in this paragraph, it seems like it’d also be a mistake to take answers to the original question to include entire lessons about the complex numbers which would put the original hearer into a position of being able to understand the facts about the function’s imaginary poles. Facts about multiplying complex numbers and so on are presupposed by, but are not part of answers involving the complex numbers in questions like this one.

\textsuperscript{38} See, for example, (Stainton 1999).

\textsuperscript{39} See (Stanley 2011: 45) for this suggestion of extending Kartunen’s theory to why-questions.
tions, it once again seems as if we should distinguish between the Real and the Complex Questions.

Of course, there’s a close connection between the Real and the Complex Questions, and it would be foolish to deny that. It also would be wrong to deny that in some sense the Complex Question is the right one to be asking in relation to \( f \) and its Taylor series expansion. The benefits of working in larger domains with special nice properties have been usefully discussed in (Manders 1989), and these kinds of benefits are certainly to be found here. E.g., there’s the simplifying effect of working with the complex numbers rather than the reals where we no longer have to distinguish between quadratic equations with solutions and those without. These mathematical facts are not being disputed. The only points at which disagreement might arise are regarding Manders’s claim that “we extend domains because that enhances understandibility of the original setting,”\(^{40}\) and the claim made by the standard solution of the Mystery of the Real Power Series to have answered the original question we were faced with. If we take the realist methodology argued for in Chapter 1 seriously, we will find reason to object to both of these claims.

What actually happens in practice in relation to this case is that a new domain of objects with an isomorphic copy of the old domain is created and a new question that very closely resembles the Real Question is then formulated.\(^{41}\) Many have wanted to identify the complex numbers with no imaginary part with the ordinary real numbers, but this is an additional philosophical/metaphysical move that needn’t be made. Many have also wanted to identify the Real and the Complex Questions because they can be formulated using the very same words, but if we look at how these questions are used, we can see that there are differences worth maintaining between them.

\(^{40}\) (Manders 1989: 561, emphasis added)

\(^{41}\) (Shanker 1988: 189) aptly calls the problems that are addressed with similar looking questions “family-resemblance problems,” recalling Wittgenstein’s notion of a family-resemblance concept. See (Wittgenstein 1953/2009: §67).
Since I have argued that the Real and Complex Questions are not the same, it’s worth spending some time considering how the questions are related. Certainly, the Real Question must minimally be doing some work in directing the activities of any mathematician attempting to answer it. A story told by Augustine in his *Confessions* can actually help clarify the role such a question can often play. Augustine recounts the story of being part of a competition as a boy that asked a group of students to rewrite the angry speech of Juno in Book I of *The Aeneid* in prose. “The contest was to be won by the boy who found the best words to suit the meaning and best expressed the feelings of the sorrow and anger appropriate to the majesty of” Juno.\(^42\) If we take a question such as, “What is the best way to render Juno’s speech into prose?” as being posed to these students and again relate understanding a question to understanding the conditions under which it has been answered, it seems as if a compelling case can be made that this question was not fully understood either by those posing or trying to answer it. Surely the judges would have had some appropriate wordings in mind as they posed the question and would have had some sense of what a winning submission might look like, but the contest would not be the type of contest it was if they had fully specified the conditions under which an entry would win in advance. Part of what makes such a competition interesting for all those involved must be the learning about what good answers look like through pursuing an answer and through judging both good and bad attempts at providing one. The question, “What is the best way to render Juno’s speech into prose?” can be seen as guiding the inquiry of all those involved in the contest without being answered by the outcome, and perhaps without even being completely understood at the outset. What is settled in the course of the competition is which of those renderings presented this time did the best job of attaining the goal of finding the best prose version of Juno’s speech as we currently understand that goal.

\(^{42}\) See (*Augustine 1961: I.xvii*). *Cf*. the fairy tale discussed by Wittgenstein and cited in (*Säätelä 2011: 173*), where the prince asks a certain smith to bring him a “hubbub” without knowing what such a thing might be.
A question like the Real Question asked above can similarly be seen as an impetus for a certain kind of search: one which begins by considering answers in the real numbers, but then perhaps shifts to a search among the complex numbers and a corresponding attempt to find the right way to formulate a new and analogous question there. In response to more difficult questions, this back-and-forth process of looking for answers and relevantly similar questions is a common one in this history of mathematics. For example, the question of whether or not a regular heptagon could be constructed using a straightedge and compass came over time to be transformed into the distinct question of whether or not $2 \cos \left(\frac{2\pi}{7}\right)$ has a minimal polynomial whose degree is a power of 2. The answer to this question has now been accepted as settling the question of the constructibility of the regular heptagon in the sense that no one would consider spending time trying to find such a construction. But it’s important to recognize that it’s at least possible to see the original question as not being the one answered by the question about a certain polynomial connected with $2 \cos \left(\frac{2\pi}{7}\right)$. If we look strictly at what happens at the level of the mathematics involved, this is perhaps the most natural view to take in fact.

Similar sorts of questions that motivate research, but are ultimately abandoned for different analogous questions, were also often commented upon by Wittgenstein. Regarding an unanswered mathematical question for which the means of approaching an answer aren’t even close to being known, Wittgenstein would recommend the following.

Ask yourself, What uses does one make of the question? It does stand for a certain activity by the mathematician, of trying, of messing about. If the question did not stand for something, one would expect any sort of activity. The question has then that meaning—as much meaning as that messing about has.

---

43 A regular heptagon with radius 1 has sides of length $2 \cos \left(\frac{2\pi}{7}\right)$, which is a root of the irreducible cubic $x^3 + x^2 - 2x - 1$. Since this polynomial is irreducible, this is also the minimal polynomial of $2 \cos \left(\frac{2\pi}{7}\right)$. The degree of the minimal polynomial for any constructible number must be a power of 2.

44 (Wittgenstein 1932-33/2001: 221)
Again, seen as an attempt to describe in the most basic terms what often happens in response to a question posing an unsolved problem, this is evidently a reasonably accurate description.

Let me summarize what I take to have been established in this section before moving on. By considering a bit of history, some semantics, and a few facts about mathematical practice viewed realistically, I have argued that the Real and Complex Questions are not the same. I’ve further claimed that the Real Question isn’t eligible to be answered by an answer to the Complex Question. So, where does this leave our original mystery? First, note that dialectically the poser of the Mystery of the Real Power Series is in the position of being unable to claim that the Complex Question presents us with a mystery. This is because over the complex numbers $1/(1 + x^2)$ and $1/(1 − x^2)$ share the property of having a pole that can be seen as stopping the convergence of any Taylor expansion outside of its radius of convergence, and $1/(1 − x^2)$ was taken to offer us a paradigm case of well-understood convergence behavior over the reals. To deny that the Complex Question is just as completely well-understood, would be flatly inconsistent. Therefore, we are left with the Real Question as either being (i) unanswered, but uninteresting now that we have answered the Complex Question; (ii) unanswered, but answerable by real methods, and so unmysterious; or (iii) unanswered, and unanswerable by real methods, and so either (a) a mystery or (b) something that just is the way it is, as Spivak might put it. I suspect that the mystery option is the least palatable of these alternatives.

Is This an Explanation?

Since I have argued that the Mystery of the Real Power Series is not in fact a mystery, it should be clear that I won’t suggest that facts about the complex plane provide an explanation of any mysterious behavior. However, it is still worth asking whether or not this
case provides an example of a mathematical explanation in some sense. Again, in order to determine what to say here, the best path is to ignore the prose and begin with a direct examination of the facts.

Mathematical Explanation

Before moving on to discuss whether or not this case offers an example of a mathematical explanation, however, it will be useful to pause briefly and attempt to shed some light on the notion a mathematical explanation in general.

Every explanation of ‘mathematical explanation’ begins with the perfunctory drawing of a distinction between proofs that merely show that a proposition is true and those that show why it is. Unfortunately, because there seem to be serious problems with the two most worked-out accounts of explanation in mathematics,45 this is also essentially the extent of the explanation that can safely be given. Instead of trying to say anything more, then, I will simply offer two quick examples that should help demonstrate what this distinction is trying to get at.46

Example 1. Anyone who has ever heard the story about a young Gauss outsmarting his lazy schoolteacher is aware of the following identity.

\[1 + 2 + \ldots + n = \frac{n(n + 1)}{2}\]

This identity can easily be proved by mathematical induction, but (as a number of philosophers and mathematicians have agreed) a proof by induction does not do a very good job

---

45 Mark Steiner (1978a) and Philip Kitcher’s (see e.g., (Kitcher 1989)) are the two main accounts of mathematical explanation on offer. They are persuasively criticized in (Hafner and Mancosu 2005) and (Hafner and Mancosu 2008) respectively.

46 It is often controversial whether or not a proof really is explanatory, so these examples may be questioned. However, that there really is such a distinction to be drawn and that mathematicians really are often motivated to find more explanatory proofs has been extensively argued in (Mancosu 2001) and (Mancosu 2008b). It is possible to remain skeptical though.
of explaining why a given result holds. A better proof of this equality, according to, e.g., Steiner (1978a: 136), follows more closely the lines along which Gauss’s reasoning is usually explained to those first learning of this identity. That is, we write the series twice—once in ascending and once in descending order—and then sum the terms one at a time.

\[
\begin{array}{cccccc}
1 & + & 2 & + & \ldots & + & n \\
n & + & n - 1 & + & \ldots & + & 1 \\
\hline \\
(n + 1) & + & (n + 1) & + & \ldots & + & (n + 1)
\end{array}
\]

It’s clear now that \(1 + 2 + \ldots + n\) must be equal to half of \(n(n + 1)\), and at least according to some this proof explains why.

**Example 2.** This second example is drawn from Carlo Cellucci (2008: 203-204). The equality,

\[
\frac{1}{3} = \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \ldots,
\]

can be proved by making use of the well-known fact that

\[
\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \ldots = \frac{1}{1-x}, \text{ for } |x| < 1.
\]

A simple substitution of \(x = 1/4\) and a subtraction of the extra 1, yields the sum of 1/3 straightaway. But, once again, there may be a feeling that this proof has not provided any insight into why the identity holds.

An allegedly more explanatory proof makes use of the figure below. The largest shaded triangle is 1/4 the area of the whole; the next largest is \((1/4)^2\); and so on. So, the sum of the shaded triangles is equal to the sum of the series

---

\(47\) See Marc Lange’s (2009) for an attempt to explain why this is the case in general.
\[
\frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \ldots.
\]

But each shaded triangle only fills 1/3 of the horizontal row of equal-sized triangles on which it lies. So, the shaded triangles represent 1/3 of the triangle’s total area. Therefore,

\[
\frac{1}{3} = \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \ldots.
\]

Again, we are now supposed to understand why this equality holds, where before we only knew that it did.

These simple examples should be enough for present purposes to illustrate the basic distinction between an explanatory and a non-explanatory proof.
NB: These examples are not wholly satisfactory for at least the following reasons: (i) they both make use of visualization, which should not be built into the concept of an explanatory proof; and (ii) the second example makes use of reasoning about a finite sequence of triangles to conclude something about an infinite sum. Nevertheless, examples of explanatory proofs that everyone agrees on are rare, and those that can be presented quickly without requiring much background are even rarer, so we will have to make do. Again, I take it that these examples suffice for the present.

Back to the Question

With these methodological considerations in mind, we can move on to reconsider our so-called mystery. We should first review the basic facts of the case.

We began with the function \( f : \mathbb{R} \to \mathbb{R} \) that maps
\[
x \mapsto \frac{1}{1 + x^2},
\]
and noted that the interval of convergence of its Maclaurin series is \((-1, 1)\). We then moved on to look at the function \( f : \mathbb{C} \to \mathbb{C} \) that maps
\[
z \mapsto \frac{1}{1 + z^2},
\]
and saw that its radius of convergence was \(|z| < 1\). Can we speak of an explanation here?

Since these functions have different domains and codomains, they are different functions. This being the case, giving them different names will help avoid confusion. I will label the real function ‘\( f_R \)’ and the complex function ‘\( f_C \)’. Now, the bare facts of the case only tell us that the Maclaurin series of an analogue of \( f_R - f_C \) restricted to \( z \) with \( \Re(m(z)) = 0 \)— in the complex plane cannot converge for any \( z \) with \(|z| > 1\). This is because, if it did,  

\[48\text{ This is a classic example of the kind of inference that is not generally warranted.}\]
it would imply convergence for $f_C$’s Maclaurin series at $\pm i$, singularities of $f_C$. Although this is, perhaps, interesting information about $f_C$, this kind of information does not explain anything about $f_R$ itself: $f_R$ and $f_C$ are independent functions that live in different spaces. I claim, therefore, that we will have to look elsewhere to find an explanation, if there is one to be found in the vicinity. Further, since the convergence behavior of $f_C$ is immediately established by Cauchy’s theorem about the convergence of a power series expansion being limited by the nearest singularity, this behavior isn’t something that stands in need of an explanation. In other cases, where, perhaps, the location of singularities is not so apparent, there still seems to be no real sense of mystery to be found. That is, finding the singularities of a function is routine, and there is no further puzzling question to ask along the lines of, “Why does this function have a singularity where it does?” Clearly, if it did not have those singularities in just those places, it would be a different function. If this is the only mathematical explanation the example offers, the case is not a particularly interesting one.

The general lesson that emerges when we look closely at the Mystery of the Real Power Series is that as long as we confine ourselves to the simple mathematical facts of the case, there is no room for puzzlement to be generated. The facts about $f_R$ and the radius of convergence of its Maclaurin series are easily explained by the real Cauchy-Hadamard Theorem; likewise for $f_C$ using the complex version of this theorem. We only get into trouble when we begin asking illegitimate questions; i.e., by asking questions of $f_R$ that only make sense in relation to $f_C$.

We might similarly be tempted to make philosophical mistakes if we were to ask, about the natural number 2, “Why can’t I add something to 2 to get zero?” The reply, “You’d have to add $-2$ to 2 to get zero, and there are no negative natural numbers,” may be tempting. Yet, ‘add $-2$ to 2’ is either ill-formed since it attempts to add numbers from two different number systems, or is a statement about an analogue of 2, $+2$ in the integers, so it
would not explain why such an addition could not occur among the natural numbers. The right response to the initial question here seems to be simply, “There’s just no such thing as adding two natural numbers to get a smaller natural number.” In both cases, the assumption that questions about one domain of mathematics can meaningfully be asked again without modification in another domain is what makes room for this kind of confusion to arise.

Despite not having found explanations where they might have been expected, I do think there are still two things that might legitimately be counted as explanations here. First, it seems reasonable to claim that the convergence of the Maclaurin series of \( f_C \) within the circle \(|z| = 1\) is explained by the fact that \( f_R \) converges for \(|x| < 1.\)\(^{49}\) However, this is not because the real function somehow causes the complex function to have this radius of convergence, but, rather, because the complex numbers are constructed with an eye to allowing the truths about the real numbers that find analogues there to remain true of their analogues wherever possible.

Second, it may also be reasonable to call Cauchy’s theorem that the power series expansion of a complex function converges up to the singularity of the function nearest the point of expansion an explanation of the meaning of the complex version of the Cauchy-Hadamard Theorem—that a power series’s radius of convergence \( R \) is determined by

\[
\frac{1}{R} = \lim_{n \to \infty} \sup \sqrt[n]{|a_n|}.
\]

This claim will be discussed in more detail in the following section.

These explanations are very different than the ones usually thought to be found in this case (and the first is arguably not even a mathematical explanation of a mathematical fact), but I contend that they are the only kinds of explanations that bear scrutiny in relation to this familiar example.

\(^{49}\) This is the direction of explanation Cauchy suggests in his early *Cours d’analyse* (1821/2009: 198).
Steiner- and Kitcher-style Explanations

Since the analyses of mathematical explanation presented by Mark Steiner and by Philip Kitcher remain the most familiar and popular general accounts of the phenomenon, it’s worth taking some time to consider what their theories would say about the Mystery of the Real Power Series and to at least very briefly investigate how their theories fare here. Although, as already mentioned, there is warranted skepticism about whether either of these theories can offer the full story about mathematical explanation within mathematics itself, they may each still provide some insight into this specific case.

Steiner’s View

The key ideas that constitute Steiner’s notion of a proof that explains why a theorem holds are given in the following paragraph.

My proposal is that an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property. It must be evident, that is, that if we substitute in the proof a different object of the same domain, the theorem collapses; more, we should be able to see as we vary the object how the theorem changes in response. In effect, then, explanation is not simply a relation between a proof and a theorem; rather, a relation between an array of proofs and an array of theorems, where the proofs are obtained from one another by the ‘deformation’ prescribed above.50

Because it’s not essential to the discussion and because the idea is intuitive enough, I won’t go into more detail about what precisely Steiner means by a “characterizing property.” One important feature of characteristic properties that must be noted, however, is that these properties are meant to be domain relative. That is, a characterizing property is meant to be one that uniquely picks out a given object or structure in some particular domain. Steiner explicitly wants this type of property to be domain-relative “since a given entity can be a

50 (Steiner 1978a: 143)
part of a number of differing domains.” Note, however, that this begs one of the primary questions that has been under consideration in this paper: namely, whether or not the same mathematical object can be moved around from domain to domain without change. Nevertheless, if it’s a characterizing property of $f_C$ that does the explanatory work, it looks like it can only be in relation to the Complex Question, which didn’t present any mystery at all, and as of yet, there’s no characterizing property of $f_R$ on offer to evaluate as to whether or not it does the necessary explanatory work. I conclude, therefore, that Steiner’s theory of explanation doesn’t contribute to a clearer view of the Mystery of the Real Power Series.

Kitcher’s View

The basic thought behind Kitcher’s modelling of scientific explanation is that explanations help us understand the world better by “showing us how to derive descriptions of many phenomena, using the same patterns of derivation again and again.” A proposition that is derivable from the set of argument patterns that best systematize the facts within a given domain is considered to have been explained by that argument pattern.

The following example of how this is meant to apply in the case of mathematical explanation is provided in (Hafner and Mancosu 2008: 154). Facts about the tangent lines to differentiable curves at various points are important pieces of information that calculus affords us. The argument pattern below should, therefore, be expected to be a part of the best systematization of our knowledge of calculus.

1. Let $f(x)$ be a differentiable function and let $(x_0, y_0)$ be the point at which a tangent line is sought.

2. Define $g(x) = f(x)'$.

3. Calculate $g(x)|_{x=x_0}$ and set this equal to $m.$

---

51 (Steiner 1978a: 143)
52 (Kitcher 1989: 432)
4. The equation of the line tangent to $f(x)$ and $(x_0, y_0)$ is then given by $y - y_0 = m(x-x_0)$.

According to Kitcher’s account, the argument pattern above can be said to provide explanation that answers any particular question about why some polynomial $h(x)$ gives the tangent to the curve $f(x)$ at a given point $(x_0, y_0)$.

Obviously, there are more details to Kitcher’s account, and (Hafner and Mancosu 2008) does the best job of trying to fill them all in. Nevertheless, this view of explanation at least prima facie looks like it might be of help. Consider the following argument pattern.

1. Let $f$ be any function that can be given locally about $c$ by a convergent power series.
2. Find the singularity of $f$ closest to $c$.
3. The radius of convergence of $f$’s power series expansion is given by the distance from $c$ to this singularity.

It looks like this argument pattern allows us to answer any given question about the radius of convergence $f$’s power series expansion about any point. However, we need to be clear about which function ‘$f$’ is supposed to refer to and about which points $c$ can range over. If we’re talking about $f_C$, then this argument pattern does answer all our questions, but it doesn’t solve a mystery since the Complex Question isn’t mysterious. If we’re talking about $f_R$, then this argument pattern simply breaks down: $f_R$ has no singularities. Again, I conclude that one of the popular theories of mathematical explanation in mathematics fails to provide insight into this case.

The evaluation of each of these views does depend on the success of my defense of the realist methodology in Chapter 1 and the arguments about distinguishing between the Real and Complex Question. However, if this defense is successful and the arguments effective, these responses to Steiner’s and Kitcher’s accounts of mathematical explanation should appear to be fairly straightforward.
Conclusions

I have spent the majority of this paper suggesting that the standard line about the Mystery of the Real Power Series is misguided. This naturally raises the following questions: (i) Why, then, has this line been so popular? and (ii) What lessons, if any, should we take from the case?

Why the Popularity?

Surely, one reason it’s been so easy to think of this case as presenting us with a mystery and its resolution is that it’s enjoyable to see a mystery being resolved, and it’s very easy to present the Mystery of the Real Power Series in such a way that it looks like this is what’s going on. The Complex Question on its own isn’t mysterious, but there’s a very nice theory that answers it completely. The Real Question seems mysterious if we take it to be asking for a barrier to convergence for \( f_R \)’s Taylor expansions, and there’s no easy theory in place to answer this question. By posing the Real Question and giving a complex answer, we can rearrange the mathematical facts in such a way that they appear more interesting than they might be otherwise.\(^{53}\)

I would also suggest that the assumption that mathematical explanation will closely mirror other forms of scientific explanation is another root of the popularity of what I have argued is a misguided approach to this case. When engaged in other sciences, a question like, “Why does metal conduct electricity?” continues to make sense even when the underlying theory of matter is modified. However, a mathematical question like, “Where is the singularity that’s causing this divergence?” which makes sense when dealing with complex

\(^{53}\) Cf. (Wittgenstein 1956/1983: Appendix II, §3) and (Mühlhölzer 2002: 310) for discussions of surprises that arise out of new arrangements of familiar facts.
functions, turns out ultimately not to make sense when dealing with real ones. It seems to be the expectation that a mathematical question will function just like a scientific one that leads to this difference’s being overlooked. While it remains true that questions of this sort can be useful heuristics in guiding later mathematical developments, it is important to remember that the developments that lead to the question making mathematical sense are the very same ones that endow it with its precise meaning.

Furthermore, when looking for an explanation of a mathematical fact, taking mathematical explanation to be very much like scientific explanation seems to suggest that we are looking for something almost like causal explanations in an area where causation has no place. Philosophers tend to respond to the Mystery of the Real Power Series with wonderment: “How did the real function know about the complex numbers?”; “Look, the complex numbers were there all along stopping the convergence of the power series but we didn’t know it.”; and so on. These imagined-responses are exaggerations, of course, but they are only slight exaggerations of the pseudo-causal thinking that can occur in cases like this. It would undoubtedly be nice if all the work that philosophers and scientists have put into to coming to an understanding of scientific explanation, where causation often plays a key role, could simply be transferred over to the mathematical case, but this appears to be no more than wishful thinking.

What Can Be Learned?

Mathematicians do often speak in terms of one fact explaining another or a proof explaining a proposition, and there is clearly nothing wrong with this. We should just not blindly take this kind of explaining to be of the same sort that goes on in other sciences because, as I have argued, that picture of mathematical explanation can be misleading.

54 Again, this can be seen to be due to something like a presupposition failure.
A perfectly good use of ‘explain’ and its cognates occurs in contexts in which what is given is an explanation of meaning. This sense of explanation actually occurs quite frequently in mathematical writing; e.g.,

Proofs are important in mathematics for several reasons, not the least of which is that a proof deepens our insight into the meaning of the theorems and gives a natural delineation of the extent of the theorem’s validity.\(^{55}\)

The next theorem explains the geometric meaning of the case when the equality occurs in the triangle inequality in the standard Euclidean metric on \(\mathbb{R}^n\).\(^{56}\)

Further, given the fact that talk of explanation and understanding are also practically interchangeable in the context of mathematical uses—an explanatory proof often is said to provide better intuition, which is closely linked with understanding; proofs that are non-explanatory are criticized for not letting us fully understand why the result holds; etc.—it would be wise to make efforts to see how far this alternative analogy helps us to understand explanation in the mathematical case, where the scientific analogy seems to fall short.

In the case examined here, we saw that constructing a function \(f_C\) that is an analogue of \(f_R\) over the complex numbers gives mathematical meaning to the previously illegitimate question about the relation between a function and the convergence of its Maclaurin series, and additionally provides the means for answering the question. Here, then, is a first instance of a mathematical construction that has been thought to do explanatory work doing work that may be better characterized as explaining the meaning of a mathematical question.

Similarly, we have seen that the Cauchy-Hadamard Theorem tells us about the radius of convergence of both \(f_R\) and \(f_C\). Cauchy first published this result in his 1821 *Cours d’analyse*. His later discovery that the power series expansion of a complex function converges up to its singularity nearest the origin—published in (Cauchy 1841)—gives new

\(^{55}\) (Conway 1978: 90)
\(^{56}\) (Kumaresan 2005: 8)
meaning to the formula

\[
\frac{1}{R} = \limsup_{n \to \infty} \sqrt[n]{|a_n|}.
\]

With good reason, we might say that Cauchy’s later theorem shows us that the real meaning of the complex Cauchy-Hadamard Theorem is that a function with the complex power series

\[
\sum_{n=0}^{\infty} a_n z^n
\]

has its first singularity at distance \( R \) from the origin.

These are clearly only preliminary remarks about this alternative picture of explanation in mathematics, and much more needs to be done to justify this claim about mathematical explanation. That task will have to be carried out in other work.\(^{57}\) However, I believe that enough has been said here to establish that there are strong reasons to look to models of explanation other than the model of explanation used in the physical sciences when thinking about explanation in mathematics, and that thinking of mathematical explanation in terms of explanation of meaning is one particular model worthy of further exploration.

\(^{57}\) Drawing on something like John Burgess’s notion of the pragmatic analytic (2004b: 52-55) might be of use in this direction.
Chapter 3
On the Role of Geometric Frames in Mathematics

Numquam hodie effugies; veniam, quocumque vocaris.¹
—Virgil, Eclogue III

Geometry seems to overstep its bounds repeatedly in contemporary mathematics. Its role in the proof of Fermat’s Last Theorem is perhaps the most famous instance of the phenomenon,² but much simpler examples illustrate the occurrence as well. Consider the problem of finding a pair of rational numbers such that the difference between the square of one and the cube of the other is ten. This problem is equivalent to finding solutions to the Diophantine equation

\[ y^2 - x^3 = 10. \]

One solution—\( x = -1 \) and \( y = 3 \)—can be found immediately, but it’s a natural question to ask whether others can be found and, if so, how. In his commentary on Diophantus’s *Arithmeticorum*, Claude Bachet found that he could discover further solutions to an equation like this one using clever substitutions.³ For example, setting \( y = 3 - N \) and \( x = -1 - 2N \) reduces our original equation to an equation linear

¹ “You shall not escape me; wherever you call me, I’ll be there.”
² Fermat’s Last Theorem is ostensibly a proposition about the natural numbers. Yet, Andrew Wiles’s key contribution to its proof consisted in showing a certain class of *elliptic curves* to be modular. (Wiles 1995) and (Taylor and Wiles 1995) proved that so-called semistable elliptic curves over the rational numbers are modular. With (Breuil et al. 2001), the modularity of all elliptic curves over \( \mathbb{Q} \) has now been established.
³ (Diophantus and Bachet 1621)
in $N$: $8N + 13 = 0$. Solving for $N$ yields another solution to the problem—$x = 9/4$ and $y = 37/8$. The process can be repeated, and presumably something like the way Bachet discovered his so-called duplication formula.\(^4\) Given a Diophantine equation of the form $y^2 - x^3 = c$ and a rational solution $(x_0, y_0)$, another solution can be found using the following formula.\(^5\)

\[
\left( \frac{x_0^4 - 8cx_0}{4y_0^2}, \frac{-x_0^6 - 20cx_0^3 + 8c^2}{8y_0^3} \right)
\]

This equation is rather complicated, and it’s natural to ask where it came from and why it holds. “The answer,” according to at least one presentation of the formula, “is that it comes from geometry!”\(^6\) If we consider the elliptic curve given by the equation $y^2 - x^3 = c$ and find a line tangent to a given rational solution, the point at which this line intersects the curve will give an additional rational solution. The procedure for the case of the equation $y^2 - x^3 = 10$ is illustrated below. Calculating how this second solution can be found from the original one yields the Bachet duplication formula. But Bachet discovered his formula before either calculus or coordinate geometry had been invented. How did he manage this? And how did geometry get involved in this apparently algebraic problem?

\(^4\) Cf. (Mordell 1923: 43-44) and (Silverman and Tate 2015: xviii).

\(^5\) This equation doesn’t work when $c = 1$ or $-432$ however. See (Silverman and Tate 2015: xvi).

\(^6\) Ibid., xviii.
More generally, we might ask, what explains the pervasive effectiveness of geometric methods and language in mathematical settings beyond geometry proper?

Stock answers to this question often appeal to a deep unity of mathematics, e.g., “Well, geometry and number theory are really the same thing anyway,”7 or to forms of platonism that postulate shared inherent structure.8 These answers, however, amount to little more than unilluminating renamings of the original problem. Further, more deflationary answers to the question suggesting that virtually any tool can be made to work in virtually any setting also seem too quick to give up on the possibility of there being any insight to be had into this general phenomenon.9

This situation should be seen as deeply dissatisfying in light of the fact that the use of geometry in ostensibly non-geometric subjects has come to be such a common and prominent feature of mathematical practice: Jean Dieudonné (1981: 231), for instance, estimates that 90% of modern mathematics involves some form of geometry.10 In fact, given the recent increased interest in understanding and analyzing mathematical practice within the philosophy of mathematics, providing careful, demystifying answers to this question ought to be taken to be one of the area’s major goals.11 It is certainly a goal that ought to be appealing to anyone motivated by the methodological realism discussed in the opening chapter.

Answering this overarching question is a daunting task, but progress can be made by pursuing answers to more focused and manageable questions. In particular, if we begin by

---

7 See (Neukirch 1999: xi).
8 Cf. (Corfield 2003: 81).
9 See (Ruelle 1988: 265-266) for a perspective of this sort on why concepts from the physical sciences often find uses in unexpected mathematical subjects.
10 N.B: I do not assume that the line between between algebra and geometry is, or could be, sharply defined. It is generally agreed, however, as Timothy Gowers writes, still “there is a definite difference between algebraic and geometric methods of thinking—one more symbolic and one more pictorial” (2008: 2). These ways of thinking tend to be aligned with problems that appear to be naturally algebraic or geometric. On this distinction, see also (Newton 1769: 470) and (Baldwin 2018: 281).
11 See (Rota 1997b: 114) for a recommendation to this effect. Rota goes so far as to claim that “[n]o philosophy of mathematics shall be excused from explaining such occurrences”; viz. occurrences that illustrate the “unreasonable effectiveness” of mathematics within mathematics itself.
asking what reasons actually motivate the introduction of geometric language, we may get a better sense of why pressure to use this language wherever possible should exist. Further, if we find that reformulating a problem in geometric language alone brings some benefits or can play an important mathematical role, this discovery should go some of the way towards explaining the consistent appearance and efficacy of geometric methods in such a wide range of circumstances.\footnote{Another way of gaining some insight into this guiding question may be through historical investigation. Many of the tools and concepts of algebra were developed to allow for computations to be carried out in geometric settings. It may, therefore, not be completely surprising that geometry often seems to “fit” these algebraic environments as well. This pursuit will not be undertaken here, however.} This more specific question will be the object of exploration here: Why might one use geometric language outside of geometry proper? And what role does this language play that makes its use important?

An examination of some of the basic features of mathematical practice reveals a number of roles that any language or set of concepts must play within the mathematical subject in which it is employed. In what follows, I shall argue that geometric language and concepts can fill these roles in an especially effective way, and that this fact at least partially explains their continual appearance in unexpected settings. More specifically, I shall offer an account on which the repeated appearance of geometric language outside of geometry is explained by (i) the convenience provided by geometrically abbreviated formulations of theorems and proofs; (ii) our special facility with geometric language, which makes geometric abbreviations more useful than others; and (iii) the fact that geometric concepts, due to their close connection with intuition as well as the web of conceptual connections among them, meet the need for mathematical concepts that play several fundamentally guiding roles. These roles include grounding particularly complex proofs, steering development, and providing a frame onto which the content of new mathematics can be grafted. In cases for which the introduction of geometric language has been most fruitful, there has been a need to introduce concepts to comprehend great complexity and to enable and guide new
research. Geometric language gains its specifically mathematical significance by being especially well-suited to playing these roles.

**Mere Convenience vs. Conceptual Abbreviation**

Because abbreviations provide the advantage of convenience and because geometric abbreviations are just a particular kind of abbreviation, it might be thought that no more needs to be said about this language’s specific advantages regarding brevity. Moving this quickly, however, would overlook many of the benefits that accrue with the use of abbreviating, geometric language—benefits that go beyond mere convenience.

To see vividly the (minimally) psychological need for abbreviated formulations of theorems and proofs, consider Bourbaki’s—still significantly shortened—expression for the number one below:¹³

\[
\tau_Z(\exists_u,U(u = (U, \{\emptyset\}, Z) \land U \subseteq \{\emptyset\} \times Z \land \forall_x(x \in \{\emptyset\} \rightarrow \exists_y((x, y) \in U)) \land \\
\forall_{x,y,y'}(((x, y) \in U \land (x, y') \in U) \rightarrow y = y') \land \forall_y(y \in Z \rightarrow \exists_x((x, y) \in U)))
\]

Adrian Mathias has calculated that Bourbaki’s completely-unabbreviated term for the number one contains approximately 4.5 trillion symbols along with around 1 trillion “links” between these symbols.¹⁴ Mathematics carried out in this form is literally unimaginable. This being the case, the possibility of writing ‘1’ in place of several trillion symbols, whether by means of letting it abbreviate these symbols or by means of introducing a new defined constant,¹⁵ is certainly a major advantage for human mathematicians.¹⁶

---

¹³ ([Bourbaki 1968: 158n])

¹⁴ See (Mathias 2002). ([Bourbaki 1968: 15-18]) explains Bourbaki’s use of so-called links and boxes to avoid bound variables.

¹⁵ The first method of making a language more user-friendly is the informal counterpart of the latter. Both means are often used simultaneously. See, e.g., ([Shoenfield 1967: 6, 57-61]).

¹⁶ This is not to say that normal mathematical proofs just are abbreviations of proofs in some formal, foundational language. (See (Azzouni 2013) for a recent argument against such a view.) If one takes the view that mathematics must be built up from, say, sets, abbreviations of the language of set theory of this sort will
Furthermore, if we are interested in maintaining a realist view of mathematical practice of the sort recommended in Chapter 1, we ought to be skeptical of any claim that 1 “really” is defined by this complicated Bourbaki expression, when there is no sign of this fact showing up in practice. In fact, we might question whether this is actually a case of abbreviation at all instead of a form of the “indifference in practice” discussed by John Burgess. When we’re doing ordinary mathematics, we don’t care what definition of 1 is chosen so long as there is an entity having the appropriate properties available to us. The claim that ‘1’ abbreviates this long enormously long expression might, therefore, be better seen as indicating that the object it denotes is eligible for playing the required role.

While it is true that this example has nothing in particular to do with the introduction of geometric abbreviations, such abbreviations offer the same kind of shortening convenience. For instance, while bemoaning the overuse of matrix methods in linear algebra, Emil Artin notes, “It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out” and replaces them with the more geometric language of linear transformations. He suggests, further, that “[m]athematical education is still suffering” from the failure to adopt this geometric approach wherever it is possible to do so.

As this example and others like it illustrate, the need for, and the advantage of, the convenience offered by geometric abbreviation is well-established and is universally agreed upon. It is sometimes thought, though, that this is the only role these tools are needed to play in mathematics. That is, it sometimes suggested that, while we may need these types clearly be required. But also, if one begins with nearly any axiomatized area of the subject, things quickly become complex to the point of needing tools for abbreviation in exactly the same fashion.

---

17 See, e.g., (Burgess 2015: 147). See also (Pettigrew 2008).
18 Cf. (Axler 1997: 81): “Compare th[is] simple proof [that every operator on a (nontrivial) n-dimensional complex vector space has an eigenvalue] with the standard proof using determinants. With the standard proof, first the difficult concept of determinants must be defined, then an operator with 0 determinant must be shown to be not invertible, then the characteristic polynomial needs to be defined, and by the time the proof of this theorem is reached, no insight remains about why it is true.”
19 See (Artin 1957: 13-14).
of shortening devices, in no sense does mathematics need them. This kind of view is expressed, for instance, by Alonzo Church when he writes that abbreviating definitions are nothing but “concessions in practice to the shortness of human life and patience.”

Looking further into the passage from Artin cited above, however, suggests one reason why this view is mistaken. It will also provide us with a first example of the important guiding roles geometric language can often play.

The Artin quotation continues by noting that the non-geometric language of matrices not only forces longer proofs, but when matrix methods are used, the more “intuitive” methods of “geometry [are] eliminated and replaced by computations.” Here, we see the content of the newly introduced language being highlighted for the first time. More specifically, Artin points out that the abbreviated geometric language does more than simply shorten statements and proofs; it also introduces contentful concepts that can fruitfully be employed in reasoning.

A definition’s ability to make for shorter statements of theorems is apparent, but the suggestion that a definition can introduce new content may initially seem curious. There are, however, two familiar ways in which definitions manage to do just that. Typically, referring to a formal definition as a piece of geometric language or describing it as isolating a geometric concept conveys the fact that the introduced term is already associated with something intuitively geometric. This is the case, for instance, when a certain collection of objects is defined as a ‘space’ or a subset of these objects as a ‘hyperplane.’ These choices of language will naturally encourage certain ways of thinking about the entities defined, which would not be present if, say, ‘tea cup’ or ‘L’ had been chosen instead. They will also suggest

---

20 Cf. (Manin 2010: 47) for example.
21 (Avigad 2003: 275-278) raises other similar and interesting questions about the sense in which mathematics needs objects such as topological spaces, manifolds, or group characters in light of the successes of simulating and studying large portions of mathematics within fairly weak systems of arithmetic à la (Simpson 1999).
22 See (Church 1956: 76).
23 (Artin 1957: 13-14)
other areas of mathematics to which these concepts might naturally lend themselves, and, therefore, contribute to the extension of their range of application. Formal definitions with this kind of attached content also occur naturally whenever a definition is introduced in an attempt to make more precise, or to abstract from, an informal concept already in use; e.g., in the rigorous definition of continuity or the abstract notion of distance. In cases like these, the formal term is counted as being geometric in virtue of its relation to some associated geometrical content.

A second common way a formally introduced term can come to have intuitive content is through a mathematician’s eventual familiarity with it. ‘Intuition’ in everyday mathematics is most often used in reference to this sort of acquaintance, which is developed after working closely with an object for an extended period of time. It is the kind of intimacy that informs judgments about how an object “should” behave or be thought about. It is also the sense of ‘intuition’ generally employed when mathematicians ask, “What’s the intuition for X?” where X is a certain construction, proof, or method, and where it may be taken in the sense of “the idea behind.” Each of these ways of imbuing an introduced term with real content will be seen to put the term in position to play new and important guiding roles. These different sorts of intuitions also ground much of the “multicolored”-ness we experience when considering the subject of mathematics as a whole.

One familiar guiding role that contentful concepts play in the practice of mathematics is in helping determine, in a sense, what a subfield is about and what the interesting questions of study are. The following quotation is taken from a standard textbook on harmonic analysis, but similar sentiments can be found in virtually any textbook chosen at random. It illustrates both aspects of this first guiding role for contentful concepts.

---

24 See (Cheng m.s.) for a discussion of the common mathematical use of ‘morally’ in this kind of context: e.g., “Morally, this sequence should converge,” etc.

25 E.g., a normal subgroup should be thought of as the kernel of some homomorphism.
Modern analysis is unthinkable without this concept [sc. distributions] that generalizes classical function spaces.\textsuperscript{26}

There is a mundane, practical sense in which a claim of this sort is often true. Without, for example, the concept of a group, it would not be possible to recognize that an object being studied had a group structure and by that recognition have a plethora of results that apply to it near to hand. Instead, one would have to reprove all the needed facts in each case.\textsuperscript{27} It might be unthinkable in such a situation that much progress could ever be made by these means. Yet, authors who make claims like the one above do not seem to have this pragmatic use of the concepts they refer to in mind. Rather, the claim that modern analysis is unthinkable without distributions indicates the author’s view that modern analysis could not exist without this concept’s being available.\textsuperscript{28} This concept is, then, at least partially constitutive of a whole area of study and questions whose solutions might lead to a better understanding of its properties and to more effective ways of using it will have a good chance of being deemed interesting or important. The content of such a concept, which stands in for and helps in the digestion of a complicated definition, therefore, plays a role that is a crucial guide to this whole portion of mathematical research and practice.

A further fundamentally guiding role for contentful concepts can be seen in relation to an important point made by Wittgenstein in his description of mathematical practice in Remarks on the Foundations of Mathematics: “A shortened procedure tells me what ought to come out with the unshortened one. (Instead of the other way round.)”\textsuperscript{29} The idea behind this passage from Wittgenstein echoes a sentiment expressed by Richard Dedekind in his text on algebraic integers.

\textsuperscript{26} See (Deitmar 2005: iii). Cf. “[A]lgebraic topology is unthinkable without homotopy theory” (Gel’fand and Manin 2003: x).
\textsuperscript{27} Cf. (Reid 1995: 152).
\textsuperscript{28} That is, without the formally defined concept, which depends on the intuitive notion that motivates the definition to fill in its mathematical content.
\textsuperscript{29} (Wittgenstein 1956/1983: III.§18)
It is preferable, as in the modern theory of functions to seek proofs based immediately on fundamental characteristics, rather than on calculation, and indeed to construct the theory in such a way that it is able to predict the results of calculation.\(^{30}\)

The kind of case we are meant to imagine in relation to each of these suggestions is one in which there is a purported proof in, say, the system of *Principia Mathematica* which is supposed to prove the translation there of \(10^{100} = 10^{100} + 1\).\(^{31}\) Now consider the question, “Is this a proposition of mathematics?” Wittgenstein’s thought is based on the natural reaction that, of course it is not because \(10^{100} \neq 10^{100} + 1\). That is, even if painstaking checks of the very long Russellian proof, repeatedly revealed no errors, the response of everyone considering the matter is expected to be, “Well, there must be a mistake somewhere.”\(^{32}\) This is because we can see conceptually that these are different numbers.\(^{33}\) It is in this sense that the shortened procedure, because it is more conceptual, can tell one what ought to come out with the unshortened one.\(^{34}\) Seen in this light, abbreviations may often play a significant mathematical role. The shortened proofs and statements of theorems that this briefer, and more conceptual, language makes possible play a role within mathematics proper of settling—for example, by allowing us to predict the results that a calculation should yield—issues of correctness and incorrectness, and, in a sense, ensuring that complicated calculation-based proofs do not go astray. In other words, the shortened proofs may play a foundational role with respect the unshortened ones.

\(^{30}\)(Dedekind 1877/1996: 102)  
\(^{31}\) I will ignore for present purposes the difficulty of even deciding what such an unsurveyable proof is supposed to be a proof of.  
\(^{32}\) That is, if the system’s consistency is not called into question in the process.  
\(^{33}\) Given that our system of naming numbers is set up appropriately.  
\(^{34}\) This is, of course, not to say that there will never be perfectly good surprising results or violations of what seem to be intuitively clear facts. A fallible guide can still be an essential one though. It is also not to say that intuitive results must always play this guiding role. The more intuitive methods of the Italian school of algebraic geometry, for example, of Severi or Segre, eventually only came to be accepted when they could be reformulated to meet the more rigorous standards introduced by Zariski and his successors. See, e.g., (Brigaglia and Ciliberto 1995) for some of this historical development.
Although it is clear that these two guiding roles played by the content of defined concepts are important to mathematics, describing them as *mathematical* roles may be controversial. I think it should not be, however. For, while it is true that one of a mathematician’s main jobs is proving theorems, it is also true that creating concepts, formulating definitions, and constructing languages are generally, and rightly, seen to be activities of no less importance. This conception of mathematics, as (at least in part) essentially a concept-forming enterprise, has roots in the writings of mathematicians such as Riemann, Poincaré, Brouwer, and Federigo Enriques; it is emphasized by Frege in the conclusion of his *Grundlagen*, and plays a central role in Wittgenstein’s later reflections on mathematics.

As an illustration of the perceived mathematical importance of finding the right definitions (which often involves finding the right way of thinking about the objects defined), consider the following examples from algebraic geometry. In his history of the subject’s development, Jean Dieudonné judges that the successful definition of the concepts of a ‘generic point’ and ‘intersection multiplicity’ constitute “the most conspicuous progress realized” from 1920 to 1950. Further, by anyone’s standards, Grothendieck’s definition—building on the work of, e.g., Zariski, Serre, and Chevalley—of the scheme and the subsequent recasting of the subject into this language was one of the most important mathematical developments of the latter-half of the twentieth century. In each of these cases, the introduction

35 This search for definitions in mathematics has recently received renewed attention in the philosophy of mathematics. (Lakatos 1976) is the classic study of this sort, but interesting, more recent work has been done in, for example, (Corfield 2001) and (Tappenden 2008). See also (Longo 2005: 366-368) and (Cellucci 2009: 218).

36 Cf. (Laugwitz 1999: 305-307), which argues that concepts were the fundamental objects of mathematics for Riemann.


38 Brouwer’s views on the (lack of a) role for language in mathematics, however, were not shared by many.

39 (Enriques 1914: 184-185)

40 (Frege 1884/2007: §88)


42 See (Dieudonné 1972: 849). These concepts had been in use prior to this time. The ‘conspicuous progress,’ however, consisted in giving them a precise formulation. This distinguishes the case from the next example where the concept of a scheme is newly introduced.
of a new concept or definition is rightly taken to be a major mathematical advance, not merely a psychological aid.\textsuperscript{43} If it is correct that mathematics is essentially in the business of forming concepts, this is presumably at least in part because of the sorts of guiding roles that concepts can play which are not available in the language without them. If so, these introduced terms and the content in virtue of which they are capable doing this kind of guiding should be taken to play a role that is in fact intrinsically mathematical.

Example: Lüroth’s Theorem

As a specific example of the difference geometric language can make to a proof consider two different ways of proving Lüroth’s theorem,\textsuperscript{44} one making use of the theory of field extensions and another using the language and results of the theory of rational curves.\textsuperscript{45}

\textbf{Theorem 1.1.} (Lüroth’s Theorem – Algebraic Form) If \( \mathbb{C}(t) \) is a simple transcendental extension of \( \mathbb{C} \) and \( \mathbb{C} \subsetneq L \subseteq \mathbb{C}(t) \), then \( L \cong \mathbb{C}(t) \).\textsuperscript{46}

\textit{Proof.} If \( \alpha \in L-\mathbb{C} \), then \( \alpha = p(t)/q(t) \) for some polynomials \( p, q \in \mathbb{C}[x] \). Therefore, \( aq(t) - p(t) = 0 \), and \( t \) is algebraic over \( L \). Let \( m(x) \) be the minimal polynomial for \( t \) over \( L \).

\( L[x] \subseteq \mathbb{C}(t)[x] \), so by factoring out any shared terms and clearing denominators it’s possible to write \( m(x) \) as \( \beta f \) for \( \beta \in \mathbb{C}(t) \) and \( f \in \mathbb{C}[t][x] \) with \( f \) primitive in \( \mathbb{C}[t][x] \). Therefore,

\[ f = \beta^{-1}m = a_0(t) + a_1(t)x + \ldots + a_n(t)x^n, \]

where the \( a_i(t) \) are pairwise coprime. Since \( m(x) \) is monic, \( \beta^{-1} = a_n(t) \) and

\[ m = \frac{a_0(t)}{a_n(t)} + \frac{a_1(t)}{a_n(t)}x + \ldots + x^n. \]

(Note that this shows that \( n = \deg_x f = \deg_x m = [\mathbb{C}(t) : L] \).) Since \( t \) is transcendental over \( \mathbb{C} \), there is some coefficient \( a_i(t)/a_n(t) \) of \( m(x) \) in \( L - \mathbb{C} \). Reduce this coefficient to get \( \delta = b(t)/c(t) \in L - \mathbb{C} \) with \( b \) and \( c \) relatively prime in \( \mathbb{C}[x] \).

\textsuperscript{43} Another well-known example of this sort is discussed in Michael Spivak’s \textit{Calculus of Manifolds} (1965: xiii-ix). The proof of the modern form of Stokes’s Theorem, which unifies several classical forms of the theorem, is a triviality once the right definitions are in place. This is often praised as being a major conceptual and mathematical advance.

\textsuperscript{44} (Lüroth 1875)

\textsuperscript{45} This example and the geometric proof that follows is drawn from (Kovács n.d.).

\textsuperscript{46} See (Garling 1986: 145-146). The theorem holds for fields in general.
Let \( r = \max\{\deg_x b, \deg_x c\} \). Then \( \delta c(t) - b(t) = 0 \) implies \([C(t) : C(\delta)] = r\).

\([C(t) : C(\delta)] = [C(t) : L][L : C(\delta)]\), so \( r \geq n \). If \( r \leq n \), then \([L : C(\delta)] = 1\) and \( L = C(\delta)\). Therefore, this inequality is all that remains to be shown.

Define \( \gamma = b(t)c(x) - b(x)c(t) \). \( \gamma \) is nonzero since \( b \) and \( c \) are relatively prime. Further,

\[
c(t)^{-1}\gamma = (b(t)/c(t))c(x) - b(x) = \delta c(x) - b(x),
\]

so \( t \) is a root of \( c(t)^{-1}\gamma \). Therefore, \( m(t)c(t)^{-1}\gamma \) in \( L[x] \); that is, \( c(t)^{-1}\gamma = gm(x) \) for some \( g \in L[x] \). Working in \( C(t)[x] \), we have \( \gamma = c(t)gm = c(t)g\beta f \). Therefore, \( f|\gamma \)

in \( C(t)[x] \). Using Gauss’s lemma, this gives \( f|\gamma \) in \( C[t][x] \). So, for some \( h \in C[t][x] \),

\[
\gamma = fh.
\]

\( \deg_x \gamma = \deg_x \leq r \) by the definition of \( \gamma \) and \( r \). \( \deg_t f \geq r \) since \( f \) has both \( b(t) \) and \( c(t) \) as factors of its coefficients. Since \( \gamma = fh \), \( \deg_x h \) must be 0. That is, \( h \in C[x] \).

Any elements of \( C[t] \) dividing all the coefficients of \( h \) must be in \( C \) (i.e., they must be units), so \( h \) is primitive in \( C[t][x] \). Gauss’s lemma implies \( \gamma = fh \) is primitive in \( C[t][x] \). Since \( \gamma \) is skew-symmetric in \( x \) and \( t \), it’s also the case that \( \gamma \) is primitive in \( C[x][t] \).

Together, \( h|\gamma \) and \( h \in C[x][t] \) imply that \( h \) is a unit of \( C[x] \). That is, \( h \in C \).

Therefore,

\[
n = \deg_x f = \deg_x \gamma = \deg_t \gamma = \deg_t f \geq r.
\]

This proof is clearly heavy on computation and algebraic manipulations. When recast in geometric language and made eligible for the application of the tools of (algebraic) geometry, the theorem has a conceptually more satisfying meaning and proof.

**Theorem 1.2. (Lüroth’s Theorem – Geometric Form)** If \( C(t) \) is a simple transcendental extension of \( C \) and \( C \subsetneq L \subsetneq C(t) \), then \( L \cong C(t) \).

**Proof.** The function field of the complex projective line \( \mathbb{P}^1 \) is isomorphic to \( C(t) \); that is, \( K(\mathbb{P}^1) \cong C(t) \). It’s possible to find a smooth compact curve \( X_L \) whose function field is isomorphic to \( L \); that is, \( L \cong K(X_L) \). There is, therefore, an injection from \( K(X_L) \) into \( K(\mathbb{P}^1) \). This injection gives rise to a dominant rational map from \( \mathbb{P}^1 \) to \( X_L \). The Riemann-Hurwitz formula implies that the genus of \( X_L \) must be 0. (Essentially, this is due to the fact that a sphere can’t be mapped onto any \( n \)-torus without tearing it.) That is, \( X_L \) is birationally equivalent with \( \mathbb{P}^1 \). Therefore, these curves give rise to isomorphic function fields; that is, \( K(X_L) \cong K(\mathbb{P}^1) \). Putting these equivalences together yields the following.

\[
L \cong K(X_L) \cong K(\mathbb{P}^1) \cong C(t)
\]
Thus far, I have argued that the introduction of abbreviations in general provides for the advantage of more convenient formulations of proofs and theorems, and that contentful concepts play essential guiding roles within mathematics as a whole, as well as within the subject’s individual subfields. I have not yet, however, made the case that geometric abbreviations and geometric contents are preferable to other abbreviations and content. The next section will make this case by looking at some relevant work from cognitive science and a recent strand of research in linguistics.

A Wonderful Aid to the Mind

Francis Bacon declaims in the *Novum Organum* that “there arises from a bad and unapt formation of words a wonderful obstruction to the mind.”\(^{47}\) The inverse of Bacon’s aphorism often holds true for mathematical language. In particular, an apt formation of words in the language of geometry is often able to function as a wonderful aid to the mathematical mind.

That the human mind is somehow especially tuned to geometric thinking has, by now, become a familiar thought. Conventional wisdom has it that posing a problem geometrically allows for pictures to be drawn, which then make it possible for the powerful resources of mathematical intuition to be brought to bear on the search for a solution. Alain Connes and Sir Michael Atiyah together exemplify this standard view in the following passages. Connes suggests that “what is difficult and essential in mathematics is the creation of enough mental images to allow the brain to function.”\(^{48}\) Once these images are available, Atiyah explains the effectiveness of allowing our intuition to act on them.

\[\text{Spatial intuition} \ldots\text{is an enormously powerful tool, and that is why geometry is actually such a powerful part of mathematics—not only for things that are obviously}\]

\(^{47}\) (Bacon 1620/1902: Bk 1, Aph 43)
\(^{48}\) Cited in (Berger 2010: viii).
geometrical, but even for things that are not. We try to put them into geometrical form because that enables us to use our intuition. Our intuition is our most powerful tool.\textsuperscript{49}

According to this just-so story, a problem’s geometric formulation makes pictures possible, and the fact that our brains have evolved to be capable of processing a great deal of spatial information very efficiently allows us to use these pictures to great effect. This efficacy is even supposed to be able to extend beyond the domain of geometry proper. As common as this kind of view is, I shall argue that it is incomplete in an important way, and, further, that, due to its narrow focus on geometric images, it under-appreciates the role of geometric language.

The lack of any real detail is one fairly unsurprising way in which the view expressed by Connes and Atiyah is incomplete. Aside from the brevity of their statement of this type of view, one cause of this lack of detail is the view’s appeal to features of the perennially difficult notion of mathematical intuition. Because of the variety of ways ‘mathematical intuition’ has been understood, it is unclear how we are meant to interpret this appeal here. In philosophical writing, for instance, ‘mathematical intuition’ is mostly used in ways deriving from Kant’s sense (\textit{Anschauung}), referring to the kind of immediate representation perception can afford us. In the cognitive science literature, on the other hand, ‘mathematical intuitions’ often refers to things like results of the direct perception of the cardinalities of small collections (\textit{i.e.}, subitizing) or the approximate size discrimination between larger-sized collections.\textsuperscript{50} Sometimes in this literature, ‘intuition’ is also used in the sense of what Efraim Fischbein calls an “anticipatory intuition”: a feeling, prior to full justification, that one has solved a problem or stumbled upon something important.\textsuperscript{51} Obviously, it would be rash to force any of these accounts of what mathematical intuition is like onto these explanations of the usefulness of geometric language. These accounts are far too brief to attempt

\textsuperscript{49} (Atiyah 2002: 6)
\textsuperscript{50} See, among many others, (Dehaene 1997).
\textsuperscript{51} (Fischbein 1982: 10)
to read any kind of theory of mathematical intuition into them.\textsuperscript{52} However, if an account of this sort is to really explain the wide-ranging efficacy of geometric methods, some actual details are called for.

A second, more significant, incompleteness this sort of view exhibits is a failure to offer any explanation of the success of geometric language in settings where the requisite pictures cannot be drawn. While it is true that pictures are available in many of the cases for which geometric language has been most fruitful, in general using the term ‘geometry’ does not imply that pictures can be expected. Rather, in contemporary mathematical practice, describing something as ‘geometry’ often signals, say, working over an algebraically closed field, or using sheaf-theoretic methods, or being able to define some sensible notion of dimension or distance, \textit{etc}. In none of these settings is an ability to draw pictures or to visualize the objects of interest necessarily available. In fact, in some of the cases in which geometric language has been successfully employed, the language is even judged to have “no geometric content.”\textsuperscript{53} Geometric language has been found to be important in many cases such as these, and in these kinds of case it is very much the \textit{language} of geometry that must somehow be doing the work, since pictures either are not to be found or are at best more properly referred to as (scare quotes) “pictures.”\textsuperscript{54}

As an example of this sort of case, consider the use of Minkowski’s theorem\textsuperscript{55} to provide a quick proof of Lagrange’s Four Squares theorem.\textsuperscript{56} Here we have a result formulated in geometric language being of use in number theory, but there are no pictures correlated

\textsuperscript{52} None of these various ways of thinking about mathematical intuition does a very good job of mapping onto the normal use of ‘intuition’ in mathematics anyway. See, \textit{e.g.}, the brief discussion of this matter in Section 1.

\textsuperscript{53} (Schneider 2011: v)

\textsuperscript{54} (Lorenzini 1996: 1) offers an example of this use of “pictures.” He writes that a geometric perspective on the objects of commutative algebra may be said to provide “added insight” by allowing someone who takes it “to draw “pictures” of ring extensions/maps of curves and of various concepts attached to extensions, such as ramification.”

\textsuperscript{55} Minkowski’s theorem states that if \( L \) is a lattice in \( \mathbb{R}^n \) with fundamental region \( F \) and \( X \) is a convex set symmetric about its center with \( \text{vol}(X) > 2^n \text{vol}(F) \), then \( X \) contains a non-zero lattice point of \( L \). See, \textit{e.g.}, (Jones and Jones 1998: 211) for an elementary proof.

\textsuperscript{56} \textit{I.e.}, every positive integer can be written as the sum of four squares.
with this language because Minkowski’s theorem in this instance relates the volume of a 4-dimensional ball to a lattice in 4-dimensional space.⁵⁷ Perhaps appeal to the power of spatial intuition can be made in order to explain the expediency of geometric language in 3-dimensional cases, but in the more important cases in which such visual tools are not available, where the pictures available are only “pictures,” spatial intuition seems to be left without anything spatial to work with.⁵⁸

Even supposing that there were a way to draw real pictures in high-dimensional cases such as these, the evolutionary explanation of the power of intuition to process these images should not be expected to carry over. There is no sense in which there has been pressure for our minds to evolve to be capable of quickly taking in and processing vast amounts of high-dimensional spatial data. In fact, only fairly recently has there even been any sort of comfort with using spatial or geometric language in settings where its employment would imply spaces of dimension greater than three. For example, in a popular algebra textbook from the early twentieth-century we read the following.

In using these terms [viz. geometric terms], we do not propose even to raise the question whether in any geometric sense there is such a thing as space of more than three dimensions. We merely use these terms in a wholly conventional algebraic sense.⁵⁹

Poincaré also notes the novelty of this linguistic innovation in *Science and Method*:

Must we say that, departing from the limited domain where our senses seem to wish to confine us, we must no longer count upon pure analysis and that all geometry of more than three dimensions is vain and useless? In the generation which preceded us the greatest masters would have replied “yes.” We have nowadays become so familiar

---

⁵⁷ This example is also an interesting one because it speaks against the idea that the possibility of a geometric proof should be taken as showing the theorem to be geometrical at heart in some way. Harold Davenport, for example, considers the proof *via* Minkowski’s theorem to be “far from” an idea proof of the four squares theorem. This wouldn’t likely be the case if the theorem was really geometrical.

⁵⁸ It should be noted that there are some common ways of visualizing spaces of more than three dimensions indirectly; *e.g.*, thinking of the points in a 3-dimensional space as being able to vary their shade from, say, white to black adds an additional degree of freedom into a 3-dimensional drawing. Regardless of one’s view of such indirect methods, the point of the example still stands since this kind of visualization does not come into play here.

⁵⁹ (Bôcher 1907/1964: 9)
with this notion of more than three dimensional space that we may speak of it even in the university without arousing astonishment.\textsuperscript{60}

Again, if neither pictures nor a powerful tool like spatial intuition, which has been conditioned for the task, is doing the work, something else must account for the fruitfulness of geometric language in settings like these.

One might object here that spatial intuition can be just as powerful in high-dimensional cases as it is in the familiar 3-dimensional case since, after all, it is a processor of spatial information and there is still spatial information being processed even in these more exotic settings. Although this is a natural enough thought, it is mistaken for at least the following reasons. The first is that, despite its naturalness, it seems to be empirically false. While the mathematically untrained can rely on their spatial intuition to determine the result of, say, the rotation of a cube in 3-space, the correct prediction of the effect of rotating a cube in 11-space or of what happens to the volume of the unit ball in \(n\)-space as \(n\) increases should not be expected. If spatial intuition only needs spatial information to function properly, there should not be such a sharp contrast between these kinds of cases. Secondly, one of the most familiar experiences in mathematics is discovering that a theorem that holds in low dimensions fails in higher dimensions, and the failure of a result to transfer from a finite number of dimensions to an infinite number is even more common. Although spatial intuition often serves us well in \(\leq 3\)-dimensional settings, the ubiquity of this kind of occurrence suggests that it is not able to gain the same kind of traction beyond this. There is no doubt that this low-dimensional intuition is indirectly a powerful tool, due to its ability to suggest things to prove in higher-dimensional settings, but in cases like these it is important to see that it is not powerful because its abilities with the processing of spatial information extend into these further dimensions.

\textsuperscript{60} (Poincaré 1914/2001: 384)
If the efficacy of geometric language outside of geometry proper is not explainable solely in terms of its ability to make pictures available, perhaps we ought to suppose, once again, that the language itself provides some kind of advantage. Section 1 noted the psychological advantage of abbreviated statements of definitions, theorems, and proofs. If geometric language is capable of allowing us to use these abbreviated forms in an especially effective way, by so doing, it will afford us important benefits even without appealing to pictures or to spatial intuition. It appears as if geometric language is capable of doing just this.

A great deal of research has been conducted (largely by Elizabeth Spelke and her collaborators) into ways in which language seems to allow concepts originally formed in isolated modules to enter into domain-general thought. This work shows that, though we share the same so-called “core knowledge” systems with other primates (including, e.g., a module responsible for processing spatial information, one that grounds subitizing abilities, an approximate number size discriminator, etc. as discussed above), the ability to compare and combine the outputs of these various modules appears to be distinctively human and offers an answer, according to Spelke, to the question “What makes us smart?” Linguists exploring this sort of research have attempted to link this ability of domain-general thought to another distinctively human ability: language.

One fundamental supposition of this linguistic work is that, in human language, the concepts originating in the various core knowledge systems must be lexicalized prior to entering into basic syntactic structures. That is, somehow, the language faculty generates new lexical items correlated to these concepts that then form the material for the larger linguistic objects it generates. The idea, then, is that once lexicalized a concept can enter into syntactic structures with any other lexicalized concept. In other words, once lexical-

---

61 Some of this work is summarized in (Spelke 2003). See also (Dillon et al. 2013) and (Spelke et al. 2010: 875).  
62 See also (Izard et al. 2011), (Spelke and Kinzler 2007), (Dehaene 1997), and (Dehaene 2009).  
63 See especially (Boeckx 2010)
ized, from the perspective of the language faculty, “radically different types of concepts are 'just words’” and are, in effect, de-modularized.\textsuperscript{64} This allows concepts formerly separated by “impenetrable boundaries” to mix freely and creates the possibility of domain-general thought.\textsuperscript{65}

This picture of the relationship between the concepts of the core knowledge systems and the items that lexicalize them suggests a way in which the language of geometry can be particularly advantageous even in the absence of images. If we suppose that a geometric term is counted as geometrical at least in part because it lexicalizes a concept which was originally based in one of the core-knowledge systems,\textsuperscript{66} it should additionally be expected that this lexical item will continue to be connected to this module, even when it is employed in non-domain-specific thought.\textsuperscript{67} This connection may then be taken to account for the availability of geometric “feel” or intuitions that arise in connection with this new use of the language, but, more importantly, it may also ground an increased processing ability when this language is employed. That is, choosing a geometrical term to introduce in a mathematical definition may enable the co-opting of the processing power of the spatial information processing module due to the fact that this term remains tied to this module even when it is mixed with lexical items outside of the module’s domain.

There is a significant body of work that appears to support this initial speculation. For instance, an fMRI study conducted by Anjan Chatterjee showed that, although subjects presented with sentences involving spatial metaphors, such as, “The man fell under her spell,” exhibited greater activation in portions of the brain associated with verbal processing than subjects presented with literal motion sentences like, “The man fell under her slide,” there were not differences in the respective activation of brain regions located in the more spatial

\textsuperscript{64} (Ott 2009: 265)
\textsuperscript{65} (Boeckx 2011)
\textsuperscript{66} (Spelke 2011)
\textsuperscript{67} In the language of \textit{(Pietroski 2007)}, it will continue to “indicate” concepts from this module.
right hemisphere. Similar results are reported in (Wallentin et al. 2005b), where the investigators found that the region of the brain that responds to motion is still activated by the use of motion verbs appearing in static contexts; e.g., “The path enters the garden.” These cases are similar to the geometric-language-without-pictures cases in that, for example, despite the language of motion’s being used, there is no real motion involved. Considering the data gathered in another similar study, Wallentin et al. suggest that

> These findings support a model of language, where the understanding of spatial semantic content emerges from the recruitment of brain regions involved in non-linguistic spatial processing.

If this model of language is accurate, it can, perhaps, underpin the explanation of how geometric language can continue to make use of the resources of our powerful spatial processing abilities even when pictures are unavailable.

Of course, one must always be skeptical of this kind of data. Studies of brain activation in the setting of abstract concepts have, in particular, often seemed to yield conflicting results. For example, one set of data “suggest[s] the right hemisphere is much better at processing concrete than abstract words,” while another may be “interpreted as support for a right hemisphere neural pathway in the processing of abstract word representations.” However, (Pexman et al. 2007) offers an appealing explanation for this fact: The “semantic retrieval of abstract concepts involves a network of association areas.” That is, abstract concepts yield activation in seemingly conflicting regions because there is just not any uni-

---

68 (Chatterjee 2008: 33)
69 (Wallentin et al. 2005b: 649)
70 See (Wallentin et al. 2005a: 221). These authors take their findings to support non-linguistic processing only in the handling of a concrete motion sentence like “The man went through the house” (the Danish, “Manden går gennem huset,” was presented in the original). They were unable to establish spatial activation with sentences like, “The man went through the sorrow” (Manden går gennem sorgen). This may be due, however, to the fact that ‘through’ in this context loses its spatial meaning and takes on an aspectual one. Cf. (Wallentin et al. 2005a: 231).
71 (Binder et al. 2005: 1)
72 (Kiehl et al. 1999: 225)
73 See (Pexman et al. 2007: 1407, emphasis added). These findings, therefore, support the view of abstract concepts put forward by Lawrence Barsalou. See, for example, (Barsalou et al. 2003).
formity in where they are processed. It seems as if the brain tries to deal with abstract concepts, which are evidently especially hard to work with, by spreading out the processing load to wherever it can. This again lends support to the view presented here. By making use of geometric language when encoding our abstract mathematical concepts, we may facilitate this co-opting of processing power and encourage the tying of an abstract concept to a particularly powerful processing center.

A final, slightly more tangential, piece of evidence in favor of this view of the functioning of lexicalized geometric concepts is how well comports with Poincaré’s famous account of mathematical discovery, which “still stands as the most viable and reasonable description of the process of mathematical invention.”\(^\text{74}\) Poincaré’s description of his realization that hyperbolic geometry could “provide us with a convenient language”\(^\text{75}\) for understanding the Fuchsian functions is rightfully well-known, but is couched in highly metaphorical language. He suggests that intensely studying a subject for a period time generates a number of mathematical ideas, which being shaped like “Epicurus’s hooked atoms,” are initially lodged by their hooks in compartmentalized regions of the mind. It can sometimes come about, however, that something may jolt and “liberate” these ideas from their homes, setting them in motion to “circulate freely” and recombine. Poincaré suggests that it is in these novel recombinations that “we may reasonably expect the desired solution” to be found.\(^\text{76}\) Despite the strangeness of this description, Poincaré’s account has resonated with many mathematicians—\(e.g\.), Hadamard thinks this description is “striking and remarkably fruitful”\(^\text{77}\)—as well as with researchers into mathematical discovery more generally. There is a conspicuous similarity between Poincaré’s story and the account of freely mixing lexical items offered above, and this fact should be taken as at least a minor point in favor of the

\(^{74}\) (Liljedahl 2004: 250)

\(^{75}\) Cited in (Gray and Walter 1997). (Note that the role of language is again emphasized.)

\(^{76}\) See (Poincaré 1914/2001: 398) for the full account.

\(^{77}\) (Hadamard 1949: 46)
correctness of the latter. That is, if working with a concept over time leads to its being lexicalized as well as to a new ability to freely mix with other lexicalized items, some of Poincaré’s language can be cashed out to yield a less metaphorical description.

With this view of the connection between lexicalized geometric terms and the concepts of the spatial reasoning module in place, one part of the argument for why geometric abbreviations are better than others, which began in Section 1, is complete. I argued there that abbreviating definitions are necessary in mathematics, minimally, for the psychological reason that the human mind cannot operate on strings consisting of trillions of symbols, but it was left open at the time why an abbreviation with geometric content might be a better choice than any other. We can now see one reason why geometrically-formulated abbreviations are preferable. The relation between a geometric term and its spatial-concept counterpart may allow for the harnessing of the power of the spatial processing module when operating with such a term. Thus, geometric abbreviations are expected to be capable of more efficient use than other sorts of mathematical abbreviation.

**Geometric Scaffolding**

One of the most fruitful techniques employed when studying some mathematical object $X$ (or some collection of mathematical objects) has been to attempt to “put a topology on it.” This has especially been true in cases where the original $X$ is a set or algebraic object not related in any obvious way to a topological structure. Endowing $X$ with a topology introduces notions of nearness and farness into the study of $X$, and, with these notions in place, the ability to use other pieces of the language of geometry is not far off. I shall argue in this section that one of the reasons this technique has enjoyed such success is that, in addition

---

78 In set-theoretic terms this involves specifying a collection of “open” subsets of $X$ meeting a few closure conditions.
to providing more structure to work with, the connections among these bits of introduced
geometric language spark development by helping determine what the interesting questions
to ask of \( X \) are. In mathematics, as in any subject, finding the right questions to ask is
often one of the most difficult tasks. The aid geometric language provides in this respect,
therefore, is one of the advantages that makes this language most valuable.

As a simple, but important, case of providing an algebraic object of study with a
topology, consider the following.\(^{79}\) Polynomials of degree less than or equal to some \( n \)
can be treated as a vector space, and this vector space can then be endowed with an inner
product, which allows for the importation of a notion of distance. One example of such an
inner product is specified below.

\[
P_n = \{ a_n x^n + \cdots + a_1 x + a_0 : a_0, \ldots, a_n \in \mathbb{R} \}
\]

\[
\langle p, q \rangle = \int_0^1 p(x)q(x)
\]

Inner Product Space of Degree \( \leq n \) Polynomials

Once our set of polynomials is placed in this environment, using the inner product, we
can ask questions about these polynomials that could not even be formulated before. For
example, one can ask about the length of a polynomial in this space, or about whether or
not two polynomials are orthogonal, and so on. Like the previous cases discussed above,
we seem to have found a way of importing geometric language into the study of purely
algebraic objects. However, whereas with these previous examples the basic questions were
already known and it seemed possible to simply reformulate them in the new geometric
language, in this kind of case, there were no such questions as, “When are two polynomials
orthogonal? or “What is the length of polynomial \( p \)” to ask before we found a way to

\(^{79}\) Other important examples include, e.g., the Zariski topology on Spec \( \mathbb{Z} \) or topological groups and the Haar
measure on those that are locally compact.
import the geometric language. This new language, therefore, can inject a whole series of new questions into an area of study. It is also capable of introducing an entirely new perspective on these unfamiliar objects or on those that are familiar in other terms.

The way geometric language can have this effect may be seen as being analogous to the more familiar process known as “syntactic bootstrapping.” A standard toy example of how this kind of bootstrapping process works is provided by examining pieces of Lewis Carroll’s Jabberwocky. “[T]he slithy toves / Did gyre and gimble in the wabe” is full of nonsense words, but because these words are framed by so-called closed class functional words, we can conclude, say, that ‘slithy’ is an adjective, that ‘gyre’ is a verb, and that “Slithy cats gimbled on the porch” is grammatical. The syntax of this sentence, therefore, provides a key to its semantics. Similarly, there may initially be no intuitive notion of length or orthogonality predicable of a polynomial, but we know certain facts about things with length—because the concept of length is part of a cluster of related geometric concepts—that can function analogously to the functional words in the Jabberwocky case in order to determine how ‘length’ is to be used in this new environment. It is at least partly in this fashion, that the framing of an area of study in geometric language can guide the way in which the objects to be studied are thought about. This language can also guide the area’s future development, which (among other things) will generally continue to expand on the connections between the new geometric language and the old. These connections and expansions, again, are part of what stitches the various colorful parts of mathematics together into the whole that becomes ever more unified.

There is a tradition that spans some of the major schools in linguistics that takes this analogy with Jabberwocky presented above very seriously. For instance, there is much

---

80 Steven Krantz notes a related phenomenon when discussing the language of sheaves, “Without the language, the theorems cannot be formulated; with the language, new insights are obtained” (1992: 293).

81 See, for example, (Boeckx 2015: 88-89), (Carey 2004), and (Carey 2009).
work in the field of cognitive semantics that is based on the idea, presented, *e.g.*, in (Talmy 2000: Ch. 1), that language is essentially composed of two sorts of concepts: the lexical (or meaning bearing) concepts such as those picked out by ‘cat’ and ‘dog’ and the structural concepts that do the organizing and framing or these lexical concepts. These structural concepts include those corresponding to the functional terms in the *Jabberwocky* example, but, more importantly for present purposes, also include “closed-class spatial forms” such as ‘near’, ‘far’, ‘adjacent’, and spatial prepositions more generally.\(^{82}\) As with other closed-class forms, despite surface differences, there is evidence that these spatial forms can actually be found in all human languages.\(^{83}\) In line with the proposal above, this work finds that spatial terms form “a “pre-packaged” bundling together of certain elements in a particular arrangement” that can be selected when linguistically encoding a particular situation and which can also be “deformed” and generalized when handling structures not previously encountered.\(^{84}\) According to Ray Jackendoff, the significance of this packaging “cannot be overemphasized” since it means that, for example, when handling concepts without familiar perceptual content or perceptual counterparts,

we do not have to start *de novo*. Rather, we can constrain the possible hypotheses about such concepts by adapting, insofar as possible, the independently motivated algebra of spatial concepts to our new purposes. The psychological claim behind this methodology is that the mind does not manufacture abstract concepts out of thin air, either. It adapts machinery that is already available, both in the development of the individual organism and in the evolutionary development of the species.\(^{85}\)

Thus, the role claimed above for geometric language in structuring new areas of research and injecting new questions to ask of old mathematical objects seems to have been confirmed for this language’s use in more concrete and mundane settings. This language’s ability to play this role in one domain suggests capabilities to do the same work in the other.

\(^{82}\) Cf. (Talmy 2005: 200).
\(^{83}\) See, for example, the “cartographic” work in (Cinque 2010).
\(^{84}\) (Talmy 2005: 200)
\(^{85}\) (Jackendoff 1983: 188-189)
Even if it is granted, however, that the connections between the concepts of our geometric language generate a range of questions to ask of anything couched in this language, an argument, again, must be made that the set of connections among our geometric concepts is to be preferred over any other language with a connected set of concepts. After all, there is a web of connections among the concepts of, say, the language we use to set out the demands of etiquette as well. Why not employ this language instead when developing new mathematics?

Although it should not simply be assumed obvious that attempts to formulate mathematical questions in this alternative language would be fruitless, there are at least the following reasons to suspect that such a project would be less successful than one involving the language of geometry. First, as discussed above in Section 2, the use of geometric language allows us to make use of our brain’s powerful spatial processing resources. An important role that an axiomatization or formalization often plays—and one that is not often emphasized in philosophical writing—\footnote{According to Frank Quinn (m.s.: 71), “[e]ssentially no philosophers understand” this role.} is providing a “framework to which we can attach our intuitions.”\footnote{See (Strichartz 2000: 113). Cf. (Corfield 2003: 152), where it is noted that axiomatization can “act as an invaluable tool in the forging of new mathematical theories and the extension of old ones.”} Deciding to use geometric language when formulating new ways of approaching a mathematical object or group of such objects (\textit{i.e.}, when establishing a new formal setting for this object or these objects), facilitates the attaching of intuitions because the language is familiar, but also makes the intuitions attached more effective due to the connection with this processing power. The immediacy of our geometric intuitions, further, allows this kind of language to be usable without any prior preparation. This being the case, development in these terms can be guided by the geometric language immediately.

A second reason to expect geometric language to serve better in this role is that the language of geometry has already been developed extensively over centuries of mathemat-
ical work. This fact is an advantage offered by geometric language for the simple reason that these geometric concepts have already been given the more precise formulation that enables them to be treated mathematically, whereas the language of etiquette has not. But, more importantly, in addition to being to a large extent a concept-forming enterprise, a great deal of what mathematics does is create and extend connections between its concepts. Theorems in mathematics often take the form of “All of the following are equivalent: (1) … , … , (n) …” or “If x is Y, it is also Z” and so on. It is natural to view results like these as forming connections between the concepts involved.88 If the advantageousness of a choice of language to use when developing a piece of mathematics is a function of the extent of the connections between the concepts of that language, and if the theorems of mathematics multiply connections among a particular set of concepts, a set of concepts that has already been thoroughly studied mathematically will be more advantageous as a developmental tool than one that has not.

If geometric language functions as a kind of blueprint for mathematicians developing certain aspects of a piece of mathematics, it should not be surprising that this language provides serious advantages or that it continues to be of great use after the mathematics has been developed. However, one might think that this view is too deflationary. In particular, it is commonly thought that the reason geometric language is so useful in non-geometric settings such as these is that these settings are not really non-geometric after all. That is, the use of geometric language may be seen simply as revealing the hidden geometry of the situation. This hidden geometry may then be taken to be what really accounts for the language’s effectiveness where it is successful. There are reasons to question such a view, however.

Although the idea that there is some hidden geometry that the geometric language is discovering might be a position that one could taken in relation to why geometric language

88 This is the view taken by Wittgenstein. See, for example, (Wittgenstein 1956/1983: III.§31).
is so useful in cases like these, there seems to be little motivation for actually taking it. First, from the perspective of trying to take the facts of mathematical practice adequately into account, this view appears to deem too much as discovery, when mathematicians would often be more inclined to describe some of these developments in other terms; i.e., as creations.\(^89\)

Second, this view of the situation fails to distinguish between that which is natural from that which is somehow already there waiting to be found. Consider for instance, David Corfield discussing the importation of geometric language into the study of sets of functions.

We may believe […] that it is simply natural for functional analysts to take a function to be a ‘point’ belonging to a space, allowing geometric notions to provide them with the idea of nearness between functions. And we may well imagine that had Riemann, Ascoli, and Arzelà not started this way of thinking, then somebody else would have done so.\(^90\)

Although it is easy enough to allow this kind of consideration to suggest that the geometric structure was hidden and awaiting discovery, rather than something implemented along with the geometric language, it is wise to keep in mind the fact that naturalness does not imply inevitability or the preexistence of some structure grounding that naturalness. The availability of simpler alternative explanations for the usefulness of geometric language ought to make this view, with its stronger assumptions of hidden structure, seem less attractive since it is not forced upon us in light of these alternatives.

A more fundamental worry about this perspective, however, is that the view seems to be without real content. Many of the concepts that we now deem to be geometric started out as being employed in 2- and 3-dimensional spaces. But as mathematics has developed, more and more structures have proved to be usefully investigated using similar methods and concepts. In fact, at present, calling something geometric seems to just mean that it is a structure that can be investigated with tools that are somewhat similar to those used to

\(^{89}\) Cf. (Gowers 2011) and the commentary in (Rosen 2011).

\(^{90}\) (Corfield 2003: 82)
investigate 2- or 3-dimensional spaces. So, to say that the usefulness of geometric language in an area shows that it was geometric all along and, in particular, that it is this inherent geometric structure that makes the language useful, appears to get the order of explanation backwards. Instead, we simply decide to call the structure that was there to begin with “geometrical” now because geometric language and methods have proved to be successful. This being the case, this account of the advantages of geometric language seems to collapse into a triviality.

This, then, completes the account of the advantages provided, specifically, by the use of geometric language outside of geometry proper. In Section 1, we saw the need for abbreviated formulations of complex theorems and proofs, as well as the need for contentful concepts to ground correctness judgments and delineate the subject matter of an area of mathematical study. But why geometric abbreviations and concepts should have been chosen over any other type was left as an open question. Section 2 explained a first reason to prefer geometric language: By making use of this language, we can co-opt our special abilities with spatial reasoning, even in settings where spatial imagery is not available as it is often is not when geometric language is used in extra-geometric settings. The present section added another reason to expect geometric language to provide advantages over other sorts of abbreviating language. Since it has been developed by mathematical study for such a long time, the already numerous connections between the concepts of our natural geometric language can serve as a more robust framework for the development of new mathematics and for the creation of new fruitful perspectives on old mathematics. This type of conceptual framework both offers us something to attach our intuitions to and serves as a guide for determining which questions to ask of unfamiliar mathematical objects and domains. All of these advantages combine in such a way that it should come as no surprise that geometric language is such a powerful tool, even outside of its natural home in geometry.

93
Conclusions and Implications

Since the time of Fermat and Descartes, the usefulness of attaching some algebraic structure to a geometric object of study has been recognized and well-understood. For example, we can see that by introducing algebraic structure a construction problem can be reconstrued as the often simpler, calculation-based problem of solving a particular equation. The success of geometric methods in algebra and number theory, on the other hand, continues to arouse feelings of mystery and to motivate strong platonist positions among mathematicians and philosophers of mathematics. If we agree with John Bell that “the ultimate purpose of the philosophy of mathematics is to demystify mathematics while at the same time celebrating it,” this state of affairs should be seen as wholly unsatisfactory. In the preceding, I have argued that one way to begin demystifying the fruitfulness of geometry outside of geometry proper is by conducting an investigation into the role and possible advantages of using geometric language in these settings. As a result of this inquiry, we have several reasons to expect geometric language to offer important, and sometimes essential, benefits wherever it is used in mathematics, and this fact alone goes some of the way towards explaining its initially surprising effectiveness in such a wide range of environments.

So, some progress has been made towards answering the difficult overarching question: Why do we seem to find geometry everywhere we look in mathematics? In part, it is because, everywhere, we look for ways to use geometric language in mathematics. When we call for it, we find it there. The exploration of the topics undertaken here, as well as the methodology used, do, however, also have implications for several other central questions in the philosophy of mathematics, and it is worth briefly canvassing some of them before closing.

91 (Bell 2008: 28)
First, questions about the role of geometric language in non-geometric settings are often (and rightly) construed as being closely tied with the subject of mathematical understanding. An example of this kind of tying can be found when Grothendieck, in a letter to Jean-Pierre Serre, writes, “one can only understand properly if geometric language is available.”\textsuperscript{92} Mathematical understanding, however, is a subject that has traditionally fallen outside the bounds of research in the philosophy of mathematics, and which, despite some recent interest, remains, as Jeremy Avigad has described it, “a sprawling wilderness.”\textsuperscript{93} This is, as Avigad notes, to a large degree because “the philosophy of mathematics still lacks an appropriate language and analytic methodology for discussing mathematical concepts.”\textsuperscript{94} This may be a result of the fact that the tools philosophers generally use to investigate language-specific aspects of mathematics, deriving primarily from proof theory and model theory, are tailored to focus on different types of questions and to deal with more formal objects than the everyday language in which mathematics is practiced. By pursuing questions about mathematical language using the best tools from linguistics and cognitive science, as we did here, perhaps further new insights can be achieved.\textsuperscript{95}

Second, investigating the effectiveness of geometric methods outside of geometry is, in part, investigating a particular instance of one of the major “varieties of mathematical explanation” delineated by Johannes Hafner and Paolo Mancosu. This sort of explanation involves cases in which a mathematical fact “is understood from a certain point of view but one looks for alternative explanations.”\textsuperscript{96} The relevant kind of example here is one in which a mathematician is interested, say, in finding a geometrical explanation of some already-proven arithmetic fact. Clearly, a better understanding of the importance of geomet-

\textsuperscript{92} (Grothendieck 1961/2004: 117, emphasis in the original)
\textsuperscript{93} (Avigad 2008: 313)
\textsuperscript{94} (Avigad 2003: 277)
\textsuperscript{95} See (Hofweber 2009) for a similar suggestion.
\textsuperscript{96} (Hafner and Mancosu 2005: 220)
ric language in non-geometric settings can help aid in understanding the nature of this type of mathematical explanation. More generally, any instance of a mathematical explanation of a mathematical fact from an alternative perspective will involve the kind of “impurity” of which geometric language outside of geometry proper is an instance. This being the case, a broader understanding of the advantages particular mathematical languages can provide in various settings is prerequisite to a complete picture of this prominent type of mathematical explanation.

Finally, the recent flurry of interest in purity of methods as an ideal of proof draws inspiration from important figures like Aristotle, Bolzano, and Hilbert, but is not balanced by an inquiry into the views of other central figures, such as Klein, Poincaré, and (at times) Dieudonné, who stressed the importance of impure methods. It is essential to keep in mind the many dimensions along which mathematicians value specific proofs when studying any particular ideal in order not to weigh that ideal too heavily. An investigation into the advantages of geometric language can help us to appreciate the special kind of value that might attach to proofs employing impure methods like using geometry outside of geometry proper. Furthermore, the study of the role of impure methods such as these may also provide some insight into the value of purity by the *via negativa*.

All of these implications suggest that the continued study of the specific role played, and advantages provided, by various types of mathematical languages should amply repay the efforts expended.

---

97 Mathematical impurity is defined by contrast to the so-called ideal of purity that has recently received significant philosophical attention, for example, in (Detlefsen 2008), (Hallett 2008), (Detlefsen and Arana 2011), and (Baldwin 2013).

98 Passages from, e.g., (Dieudonné 1969) and (Dieudonné 1985) cited in (Detlefsen and Arana 2011: 5-6) suggest that Dieudonné and other Bourbakistes understood purity of methods as a value. However, as I’ll show below, Dieudonné also emphasizes the central importance of impure methods; e.g., in (Dieudonné 1975), (Dieudonné 1981), and (Dieudonné 1992).

99 See, e.g., (Tao 2007) for a discussion of some of this variety.
Cantor’s paradise,\textsuperscript{2} and the world of mathematics more generally, has seemed to many to be much like Dante’s Paradise, where “law eternal” ordains that “no trace of chance can find a place.”\textsuperscript{3} However, mathematical coincidences or accidents, phenomena recently under discussion in the philosophy of mathematics,\textsuperscript{4} threaten to disturb this idyllic setting by injecting a bit of the unexplained and unordained into paradise. This chapter presents a view of mathematical coincidence which is informed by what we find when we ask (following the realist precept recommended in Chapter 1) not what a mathematical coincidence \textit{is}, but what work calling something a mathematical coincidence \textit{does}. The view to be presented might be called—only \textit{half}-jokingly and in keeping with the divine theme—Pascalian. In one of his so-called minor works, \textit{The Art of Persuasion}, Pascal suggests that God wants divine truths to enter our lives by proceeding from the heart to the mind, not the other way

\begin{flushright}
\textit{[L]es saints […] disent en parlant des choses divines qu’il faut les aimer pour les connaître}.\textsuperscript{1}
\end{flushright}

\begin{flushright}
\textemdash\textit{Pascal, De l’Art de persuader}
\end{flushright}

\textsuperscript{1} “[T]he saints […] say, in speaking of things divine, that we must love them in order to know them.”
\textsuperscript{2} (Hilbert 1926/1983: 191)
\textsuperscript{3} (Alighieri 2007: Canto XXXII, 52-56)
\textsuperscript{4} See, e.g., (Baker 2009), (Lange 2010), and (Lange 2017: Ch. 8).
around. The view I’ll be presenting here is Pascalian in the sense that it suggests that our talk of mathematical coincidence is most appropriately seen as being involved in making us love mathematics in a way leads us to mathematical knowledge and understanding. Even in this abstract domain, we need a loving guide to lead us along the path to knowledge.

The plan for the chapter is as follows. I shall begin in Section 1 by presenting several quick examples of what have been called mathematical coincidences. I shall then go on in Section 2 to discuss some of the initial troubles that make the idea of coincidence in the realm of mathematics seem so puzzling. Next, Section 3 will present and criticize the default theory of mathematical coincidence in the literature. Section 4 then explains and motivates the alternative Pascalian view, which I defend from what I take to be the most obvious and potentially threatening objection to the view in Section 5. I close in Section 6 by reflecting on the role the realist methodology defended in Chapter 1 has played in coming to the improved understanding of mathematical coincidence and its role in practice achieved here.

A Budget of Coincidences

At least since (Davis 1981), a number of mathematicians and philosophers have noticed that within mathematical practice facts or pairs of facts are often classified as being “coincidental.” The following examples are just a small sample of this phenomenon.⁵

Example 1. The thirteenth digit of both π and e is 9.⁶

\[
\begin{align*}
\pi &= 3. 1 4 1 5 9 2 6 5 3 5 8 \mid 9 \mid 7 \ \ldots \\
e &= 2. 7 1 8 2 8 1 8 2 8 4 5 \mid 9 \mid 0 \ \ldots 
\end{align*}
\]

⁵ So-called “monstrous moonshine” (Conway and Norton 1979) is another more famous and much more difficult source of examples of surprising coincidences relating facts about finite groups and modular functions. See, e.g., (Gannon 2006).
⁶ This example is from (Davis 1981: 312).
Example 2. 1!, 2!, …, 100! are all “Niven” (or “harshad”) numbers. That is, each is divisible by the sum of its decimal digits.\(^7\)

Example 3. The expression $\sqrt{1141y^2 + 1}$ isn’t an integer for values of $y$ between 1 and 1,000,000,000,000,000.\(^8\)

Example 4. Each of the following is a prime number: 31, 331, 3331, 33331, 333331, 333331, 3333331.\(^9\)

Example 5. Pick any “normal” three-digit number (i.e., aside from 111, 222, …); for example, 713. Take the number that results from arranging the digits in descending order and subtract from that the number that results from arranging the digits in ascending order: 731−137 = 594. Repeat this process again: 954−459 = 495. And again: 954−459 = 495. Choosing a few other starting numbers also yields a repeating 495 after several iterations of this operation—a so-called Kaprekar operation.\(^{10}\)

One striking fact about these examples is the different sorts of questions they prompt and the different levels of desire for explanation they elicit. I assume that most wouldn’t be inclined to ask whether or not it’s a coincidence that $\pi$ and $e$ have the same thirteenth decimal digit or be desirous in the least of an explanation of that fact. The primality of numbers of the form $33\ldots31$ in general, however, is hard not to immediately ask after. Certainly, one would not expect all numbers of this form to be prime, but the question nevertheless naturally arises. Finally, and this will become important in Section 3, at least two of these examples which have been called mathematical coincidences have explana-\(^{10}\)

---

\(^7\) This example is from (Guy 1990: 10).
\(^8\) This example is from (Davis 1981: 313)
\(^9\) This example is from (Guy 1988: 699).
\(^{10}\) This operation is named after Dattathreya Kaprekar, who studied it and discovered a number of facts about it. See, e.g., (Kaprekar 1955). Kaprekar also gave the name ‘harshad’ to harshad numbers.
tions (to be provided presently): e.g., the coincidence that any normal three-digit number we repeatedly apply the Kaprekar operation to reaches 495 eventually is easily explained by elementary considerations. Before moving on, then, I will briefly revisit each of these examples, asking and answering a few natural questions regarding each.

**Example 1 (bis).** This seems to be merely a coincidence.

**Example 2 (bis).** It’s natural to ask whether \( n! \) is a Niven number for all \( n \). It turns out that there is no counterexample until 432!. The sum of the 953 digits of this number is \( 3^2 \cdot 433 \). Since

\[
432! = 432 \cdot 431 \cdot \ldots \cdot 2 \cdot 1,
\]

and 433 is prime, the sum of the decimal digits can’t divide 432!.

**Example 3 (bis).** The first natural question to ask about the expression \( \sqrt{1141} y^2 + 1 \), given that there are no small values of \( y \) for which it is an integer, is whether it is ever an integer. The answer is that it is and this fact is explained by the theory of the Pell equation: if \( d \) is a positive integer which isn’t a perfect square, the Pell equation \( x^2 - dy^2 = 1 \) has nontrivial integral solutions. It’s not until \( y = 30, 693, 385, 322, 765, 657, 197, 397, 208 \) that we find our first integer value, however. So, it’s also natural to wonder why we don’t find any integral values until such large values of \( y \). This is again explained by the Pell equation. If the period of the continued fraction expansion of \( \sqrt{d} \) is very long, the first solution to \( x^2 - dy^2 = 1 \) will come at a very large \( y \). In this case the period of the continued fraction expansion of \( \sqrt{1141} \) has length 58:

\[
[33; 1, 3, 1, 1, 12, 1, 21, 1, 1, 2, 5, 4, 3, 7, 5, 16, 1, 2, 3, 1, 1, 1, 2, 1, 2, 1, 4, 1, 8, 1, 4, 1, 2, 1, 2, 1, 1, 3, 2, 1, 16, 5, 7, 3, 4, 5, 2, 1, 1, 21, 1, 12, 1, 1, 3, 1, 66]
\]

So, we should expect the first integral value of \( \sqrt{1141} y^2 + 1 \) to be at a very large \( y \).

---

Example 4 (bis). Again, it’s natural to wonder whether all numbers of the form $3 \ldots 31^n$ are prime, but one would also naturally expect not. This expectation is the correct one; e.g., $333333331 = 17 \cdot 19607843$. We might also want to ask if it’s possible to know in advance whether $3 \ldots 31^n$ is prime for any given $n$ though and continue to investigate numbers of this form.

Example 5 (bis). It can be shown fairly quickly that applying the Kaprekar operation involved in this example to a three-digit number always terminates in 495. Each iteration of the Kaprekar operation has the following form, where $0 \leq c \leq b \leq a \leq 9$.

\[
\begin{array}{ccc}
  a & b & c \\
- & c & b & a \\
A & B & C
\end{array}
\]

And the relations between $a, b, c$ and $A, B, C$ can be seen to be the following.

\[
\begin{align*}
  C &= 10 + c - a \\
  B &= 10 + b - 1 - b = 9 \\
  A &= a - 1 - c
\end{align*}
\]

Since we’re looking for a fixed point of this operation, we need $A, B, C$ to be some permutation of $a, b, c$. There are six possible permutations, but only $(A, B, C) = (c, a, b)$ solves the equations. And for this permutation, $(a, b, c) = (9, 5, 4)$. That is, applying the Kaprekar operation in the form of $954 - 459$ results in the fixed point of 495.

A similar procedure can be used to show that 6174 results from the repeated application of the Kaprekar operation to normal four-digit numbers. This leads to a natural further question: For which length numbers does the Kaprekar operation terminate in a single value? And is there a reason that certain length numbers behave this way under the operation or is it merely coincidental that, e.g., three- and four-digit numbers have a fixed point but five-digit numbers do not?

---

12 For four-digit numbers, the operation ends with 6174. Neither two-digit nor five-digit numbers have a repeating value (or “kernel”). See (Nishiyama 2012: 370).


101
Initial Puzzles

As observed in the last section, remarks about a fact’s being a coincidence or an accident are both common and somehow natural in the setting of mathematics. There’s nevertheless something very strange about the idea of a mathematical coincidence. Coincidences seem to involve the simultaneous occurrence of accidental, unlikely, or unconnected events.\(^\text{14}\) But neither events nor anything else occurs in the world of mathematics; mathematical facts aren’t events of any kind; and there’s no place for time here either. Further, we tend to think of the world of mathematics as being a place where nothing is true by accident or chance: everything is necessary in this sphere. So, despite the fact that people talk about coincidence in mathematics fairly regularly within mathematical practice, it’s not clear that the means we have for thinking about coincidence in ordinary language are up to the task of being translated over to this new setting. Some rethinking and care is at least necessary.

One familiar way to proceed in the face of this sort of puzzlement is to start with a list of examples like the one given above and then attempt to find necessary and sufficient conditions for membership on the list and for being a mathematical coincidence more generally. These conditions, if we can find them, will likely show us what to put in place of the occurrences, events, and times integral to our ordinary understanding of coincidence and, further, might allow us to categorize pairs of mathematical facts as being coincidences or non-coincidences without fail. This has been the way the phenomenon of mathematical coincidence has been approached in the literature so far.\(^\text{15}\) Examining a version of this kind of view and its shortcomings will continue to clarify the phenomenon of mathematical coincidence and help prepare the way for the Pascalian view presented in Section 4.


\(^{15}\) See especially (Lange 2010).
The Dictionary Analysis

The move from *The Oxford English Dictionary*’s definition of ‘coincidence’ as “[a] notable concurrence of events or circumstances having no apparent causal connection,”\(^{16}\) to what I’ll call here the *Dictionary Analysis* proceeds by discarding ‘notable’ as being inessential,\(^{17}\) substituting a pair of mathematical facts in for concurrent events, and suggesting that the analogue of “no apparent causal connection” is “lacking a common explanation.”\(^{18}\) After these adjustments we arrive at (in Lange’s formulation) the following.

For mathematical truths \(A\) and \(B\), it is a coincidence that \(A\) and \(B\) are true iff \(A\) and \(B\) have no single, unified explanation.

This definition makes use of the concept of a single, unified mathematical explanation, and so can’t be fully understood without some basic idea of what such a thing is supposed to be.\(^{19}\) The discussions of mathematical explanation in Chapter 2.6(a) and in Chapter 5, however, are sufficient here since the criticisms I will make of this view do not depend in any crucial way on the fine details of the analysis.

Before raising some worries about this account, it must be admitted that it seem like it’s getting at something correct and interesting. What else makes it merely a coincidence that the thirteenth digit of \(\pi\) and \(e\) is 9 if not the fact that there’s no unified explanation of that fact? Despite this initial appeal, however, the following four reasons suggest that we shouldn’t be satisfied with the Dictionary Analysis.

The first concern stems from the fact that this understanding of mathematical coincidence seems not to mesh well with some of the things mathematicians themselves have said


\(^{17}\) (Lange 2010: 316)

\(^{18}\) See (Lange 2010: 316-322) and (Baker 2009: 141).

\(^{19}\) The qualification essentially aims to prevent one from claiming that by putting together an explanatory proof of \(A\) and an explanatory proof for \(B\) one has given an explanation for \(A\) and \(B\).
about coincidence in mathematics. For example, several of the examples in Philip Davis’s original article on mathematical coincidence (including Example 3 above) do not fit the Dictionary Analysis. In his account of Example 3, Davis notes that the first time \( \sqrt{1141 y^2 + 1} \) is an integer is when \( y \) is greater than \( 10^{25} \); he then wonders whether this is a coincidence. But *prima facie* this example does not consist of two mathematical facts being joined together in want of a unifying explanation. The only obvious choices for two facts involved in this sort of example don’t seem to be very promising ones either. E.g., maybe it’s the fact that \( y \) is very large and the fact that \( \sqrt{1141 y^2 + 1} \) is an integer that need a common explanation; or the facts that \( y \) is an integer and that \( \sqrt{1141 y^2 + 1} \) is one too. Yet, it’s hard to imagine how any explanation of the fact that \( \sqrt{1141 y^2 + 1} \) is an integer could explain why \( y \) is very large or why \( y \) is an integer. In fact, it seems as if there just is no explanation for why a number is very large or why an integer is an integer.

Since accidents don’t necessarily involve a conjunction of events or facts, it may be thought that Davis’s question would be better put in terms of whether or not it’s an *accident* that the first integral value of \( \sqrt{1141 y^2 + 1} \) comes at such a large \( y \). This may be the correct way to go, but the Dictionary Analysis shouldn’t be expected to easily provide a good account of mathematical accidents. An accident can have a fully satisfying explanation and be an accident nevertheless: e.g., the *Challenger* disaster was a horrible accident, but it is explained by the failure in the below-freezing temperature at launch time of O-rings meant to seal segments of the booster rockets. Given this fact about explanation and accidents, the lack of a certain kind of explanation is unlikely to adequately capture the notion of a mathematical fact’s being accidental.

Further, not only do some purported examples of mathematical coincidence not seem to fit with the Dictionary Analysis, both Davis and the mathematician E. H. Moore suggest what is essentially the exact opposite of the Dictionary Analysis in general. They both claim
that, “The existence of [a] coincidence implies the existence of an explanation”; that is, once you see something interesting or unexpected going on, you’re going to be able to find an explanation for it in some way. The Dictionary Analysis says that coincidence implies that there is no explanation though, so clearly both can’t be correct. The way Davis and Moore speak of coincidences as having explanations may seem unusual, but it is paralleled to some extent in ordinary discourse as well. If coincidences by definition have no explanation, “There’s no explanation for this coincidence,” would be tautologous, and something of the form, “X explains the coincidence,” would be contradictory. I take it that neither statement is tautologous or contradictory, however. Certainly, one might suggest that it would be more careful to say, “X explains the apparent coincidence,” in the latter case, but even if this were so, the fact that “X explains the coincidence” isn’t nonsense, in combination with the claims of Davis and Moore from before, ought to provide some evidence against the Dictionary Analysis.

The second worry about the Dictionary Analysis is that it has the odd consequence that nearly everything in mathematics turns out to be a coincidence. Take practically any two mathematical facts you like and there just isn’t going to be a unified explanation that connects them. Of course, it ultimately would not be so bad if there turned out to be many more coincidences than we might have expected, but it’s not even clear that it makes sense to call two randomly selected facts either a coincidence or a non-coincidence. Consider non-mathematical cases first: Is it a coincidence that I like Scotch and Beijing was very hot in April of 1993? Is it a coincidence that my birthday is in January and my wife’s is in

\[20\] See (Davis 1981: 320) and (Moore 1909), cited in (Krieger 2003: 214). (Guy 1988: 698), e.g., disagrees though. He suggests that early coincidences are actually “the enemy of mathematical discovery” since they tend to send us on wild goose chases for proofs of theorems that are simply false.

\[21\] Note that there remain questions about how seriously to take and weigh the opinions of mathematicians on these sorts of issues. Recall that it’s not part of the realist methodology being illustrated here to accept something as false simply because a mathematician said something different. It’s still worth taking note of this data though.
September? It’s not at all apparent what to say in these cases since the notion of coincidence seems to be so unsuited to this kind of usage. I would suggest that to the extent that we can make sense of these sorts of questions, we do so by trying to imagine a situation where the questions would make sense—e.g., the questions about birthdays might make sense if my parents birthdays were also in January and April, say—but making sense in a different context doesn’t imply that sense is made with this kind of question whenever it’s asked.\textsuperscript{22}

Lange is aware that his view has this odd consequence, and attempts to explain away its strangeness by an appeal to salience: we tend only to ask whether two events are coincidental if there’s some reason to suspect that a common explanation could be found.\textsuperscript{23} So, it may be the case that it really is a coincidence that the twelfth digit of $\pi$ is 8 and the fifth digit of $e$ is 2, but we would not ordinarily say so because there’s nothing that would lead us to expect a unifying explanation in the first place. This attempt to deal with this consequence of his view, however, makes the mistake noted above of taking a question to make sense just because we can find a different context in which it makes sense: if there was something salient that made us ask about the twelfth digit of $\pi$ and the fifth digit of $e$, we could decide whether it was coincidental or not that one is 8 and the other is 2, so it’s safe to conclude that it makes sense to say that it’s a coincidence now. But this conclusion cannot safely be drawn. That Lange is aware that his response here is not completely satisfactory is made clear by the fact that he makes the following altogether different claim later in the same footnote currently under discussion, “Without some salient feature, it may not make sense even to ask whether or not [a pair of facts] is a coincidence.” This claim jeopardizes the whole Dictionary Analysis though: if it doesn’t make sense to ask whether or not two randomly chosen facts are a coincidence, then the fact that the Dictionary Analysis says that they almost always are coincidental is clearly problematic.

\textsuperscript{22} Cf. (Wittgenstein 1969: §10).
\textsuperscript{23} (Lange 2010: 328n17)
Thirdly, and what I expect to be most controversially, if a mathematical coincidence is supposed to be a certain kind of mathematical fact, there’s reason to doubt whether $A$’s being a mathematical fact and $B$’s being a mathematical fact together imply that $A \land B$ is a mathematical fact as well. Since it’s an event that’s a coincidence or not (e.g., two friends meeting unexpectedly at the store), if we want to maintain some kind of connection to the ordinary concept of a coincidence, it’s important not to try to explain simply the facts $A$ and $B$ together, but to explain the fact $A \land B$. Philosophers have from time to time wondered whether $A$’s being a physical fact and $B$’s being a physical fact automatically make $A \land B$ into another physical fact; that is, one can ask whether or not there are conjunctive facts or events really in the world. My concern here is not based on a metaphysical worry about conjunctive mathematical facts however. Rather, I take it that if someone were writing up a textbook or a research paper that included something like the following, “Theorem 1: if $G$ has order $p^a q^b$, where $p$ and $q$ are prime, then $G$ is a solvable group, and $x^n + y^n = z^n$ isn’t (non-trivially) solvable in integers for $n > 2$,” a reviewer would surely say, “That’s not a theorem: that’s two theorems.” Or if it was a student who made an odd choice like this in presenting her work, a professor would surely say, “That’s not how we do things.” Certainly, it’s to be expected that if one were to ask a mathematician whether two theorems could simply be joined together to get a new theorem, she would probably say, “Yes.” But if we take the actual practices of mathematicians as seriously as we should, it becomes clear that a mathematician would not actually act on this answer. If that’s an accurate assessment of the norms of mathematical practice, one might be inclined to say that there are no mathematical coincidences. There are coincidences of course, but as soon as $A \land B$ gets the status of being a mathematical fact, it is by that very process not a coincidence anymore. What makes $A \land B$

---

25 This theorem consists of the Burnside’s theorem and Fermat’s last theorem.
a mathematical fact and not just two facts $A$ and $B$ is precisely that it’s been given a single (and at least somewhat) unified proof.\textsuperscript{26}

Finally, the Dictionary Analysis fails to explain the appropriateness or inappropriateness of certain attitudes we might think one ought to adopt upon deciding that some fact is a coincidence or not. This worry is somewhat analogous to a concern that has led many people to embrace something like expressivism in ethics.\textsuperscript{27} It has seemed to many to be obviously correct that, if you sincerely think that you ought to give money to charity each month, then you ought to be motivated to do so in at least \textit{some} way.\textsuperscript{28} Expressivism has seemed appealing in light of this apparent connection between moral beliefs and motivation since it suggests that the way our ethical beliefs are related to the world is different from the way in which our other beliefs are. If one holds that moral beliefs do something like express a person’s view of how the world ought to be rather than having something to do with the world as it is, a ready-made story is available about why a moral belief would automatically imply some kind of motivation.\textsuperscript{29} On a view for which we aim to have our beliefs about morality accurately represent the world in the same way we want our beliefs about, say, crustaceans to do, this connection between motivation and belief is not so easily accounted for. In the case of mathematical coincidence, if a mathematician were to think that the fact $A$ and $B$ is merely a coincidence and then spend the next twenty years of her research career trying to figure out why $A$ and $B$ holds, or if she were to think “Aha! This is no coincidence!” and then show absolutely no interest in finding out what connects $A$ and $B$, something would seem to be amiss. We ought to ask our thinking about coincidences in

\textsuperscript{26} In fact, this way of thinking about coincidence in mathematics may go some of the way to explaining the thinking of Davis and Moore referred to above.

\textsuperscript{27} See (Schroeder 2010: Ch. 1.4).

\textsuperscript{28} See (Smith 1994: 71-76) for an influential argument for this claim from someone not moved to expressivism by this “motivation problem.” (Shafer-Landau 2000) and (Railton 1986), on the other hand, object to this sort of motivational internalism.

\textsuperscript{29} Again, see (Schroeder 2010: Ch. 1.4) here.
mathematics to do some of the work of accounting for the appropriateness of these attitudes in one way or another, but the Dictionary Analysis does not appear to provide us with the necessary tools. This fact should count as a mark against the view.

A Pascalian View

The objections to the Dictionary Analysis raised in the last section should clear some room for an alternative proposal regarding mathematical coincidence and its role in mathematical practice. The connection with expressivism made in this previous section, suggests one path along which we might find a more fruitful account. A philosopher inspired by the realist approach to the philosophy of mathematics described in Chapter 1 will try to follow this path and do better by eschewing attempts at definition and focusing instead on aiming to understand the place of coincidence talk within the ordinary practice of mathematics. It may be hoped, further, that this kind of understanding will provide a more satisfying picture of mathematical coincidence than even a correct definition could yield.

As mentioned in the introduction to this chapter, Pascal had the idea that, when it came to things divine, one must love them in order to know them. I would like to suggest that there’s good reason to think that something quite like this is the case with things mathematical as well. Other sciences can rely more on human needs and necessities to push forward investigation and research: new diseases come along; the climate changes; people demand faster computers; etc. Some of the science involved in these pursuits will push mathematics forward as well obviously: in order to build a better bridge, you need to understand, say, some geometry, differential equations, and so on. Even the purest number theory can sometimes be driven forward by real world applications such as the modern need for

30 And if something like Marx’s view about the development of human needs through the exercise of human powers is correct, we can be confident that this source of motivation will only ever become stronger. Cf. (Marx 1844/1988: 111-112) for example.
encryption online. But there’s plenty of mathematics that certainly does not now and may
never have real world applications or be used to solve human problems. If we think that
the general practice of mathematics is a worthwhile and a significant human endeavor that
needn’t be justified in terms of applications, we need to continue finding ways to generate
particular interest in it and even love for it. One way of creating this kind of care, which
we find in ordinary mathematical discourse already, is putting forward a given mathemat-
ical fact or pair of facts as being no coincidence.\footnote{Although he wouldn’t agree with the
details of the Pascalian view presented here, Lange agrees that coincidence talk can sometimes make it easier to recognize interesting issues. See \citep{Lange2010}.} To make a claim like this is a way of
saying that there might be something worthy of interest here—something that’s of interest
independently of anything else that might make a piece of mathematics worth caring about.
Similarly, saying that something is \textit{just} (or \textit{merely}) a coincidence is a way of suggesting that
it’s not something worth spending time on. (Hopefully, the examples presented in Section
1 elicited just these sorts of judgments.) These sorts of evaluative judgments are in a real
sense essential to the practice of mathematics. We don’t have unlimited time or resources,
so we need to focus attention on one area or problem rather than another now and then;
researchers need to get other people interested in their problems to sustain their work, to get
grants and postdocs; and so on. Talking in terms of mathematical coincidence is one way
of contributing to the doing of all this all this work.

To see this kind of interest-generating role for mathematical coincidence in action in
a simple case, consider the following pair of “worm-eaten arithmetic” problems.\footnote{\citep{Nishiyama2012}}

\[
\begin{array}{cccccc}
\Box & \Box & \Box & \Box & \Box \\
\times & \Box & \Box & \Box & \Box & \Box \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{array}
\]
Each of these problems is solvable by the same method, but due to the coincidence that the product arrived at in the first is 123456789, there’s a natural desire to find the answer, while 123456784 is somehow less exciting. I don’t want to claim that the difference in interest is very great here at all or that the problem is itself of any great significance. However, I do claim that such a difference in interest is there nonetheless, and that fact is worth trying to understand.

There are a few other roles that talk of coincidence in mathematics often plays and which are worth mentioning here as well. Although these roles are clearly important, they are the sort of roles that philosophers of mathematics have, in general, been slow to pay attention to unfortunately. The first is that fact relevant to mathematical pedagogy that presenting a new piece of mathematics as a case for which there’s a question about whether or not a coincidence is involved can motivate students to learn the fact better as they do the work required to find out the answer on their own, and there are many studies in the cognitive science literature and the literature on mathematics education that bear out the effectiveness of this approach. One important part of mathematical practice is the training of new generations of mathematicians, and noticing and making use of the effect that this style of presentation can have on students seems to be a wise thing to do. Generally seeing mathematical coincidence talk as essentially aimed at creating or dissolving a certain kind of care can play an important role in taking this kind of fact about pedagogy into proper consideration.

\[
\begin{array}{ccccccccc}
\times & & & & & & & & & \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 4 \\
\end{array}
\]

33 123456789 = \(3^2 \cdot 3607 \cdot 3803 = 10821 \cdot 11409\) and 123456784 = \(2^4 \cdot 11^2 \cdot 43 \cdot 1483 = 10406 \cdot 11864\).

34 See, for example, (Buchbinder and Zaslavsky 2011) and (Diaconis and Mosteller 1989: 859). (Dannenberg 2008: 90) suggests that the desire for explanation generated by coincidence also plays a significant role in literature.
The idea that talk of mathematical coincidence can play an effective role in mathematical education is supported by the fact that in mathematical textbooks the most common use of ‘coincidence’ is in the context of saying that something is “no coincidence.” This is a way of saying that the facts presented so far may seem to be merely coincidental, but that one ought to stay tuned for a more satisfying story. The use of coincidence talk plays a different, but still central, role in research-level mathematics. In contrast to textbook writing, a search for the word ‘coincidence’ on MathOverflow, “a question and answer site for professional mathematicians,”³⁵ reveals that the term occurs most often by far in the context of questions of the form, “Is it a coincidence that X?” Asking whether or not something is a coincidence on a forum like MathOverflow is one way of gauging whether other mathematicians find the particular fact worth investigating or not and generating the kind of attention needed for progress to be made. This kind of sparked interest can play the role of prompting other researchers to get involved in the assembly and correlation of answers to other non-why-questions from which answers to coincidence-related why-questions tend to emerge.³⁶

Finally, and I hope not too fancifully, it’s easy when trying engage in the cold, objective observation of a science to forget that people study mathematics and pursue mathematical research for the unique kinds of pleasures that can be obtained in such pursuits. There’s beauty to be appreciated in certain proofs and there’s enjoyment to be had in observing certain kinds of cleverness in mathematics that are specific to the practice and worth pursuing for their own sake. One further kind of enjoyment that one can experience when engaging in this kind of work is wondering whether something is a coincidence or not and then having

³⁶ Cf. Sylvain Bromberger’s advice to someone seeking an answer to a why-question: “My guess is that the rational thing for him to do is to forget about the why-question and to turn to other questions instead, remembering that answers to why-questions usually emerge from work on questions with more reliable credentials” (Bromberger 1992: 169).
the question settled.37 This kind of resolution can come about through new research or can be put before you by someone writing an article or textbook. This kind of pleasure is very similar to one that many find in reading works of fiction. Many people have had the experience of reading Dickens, e.g., *Martin Chuzzlewit*, and having the thought, “Ah, another very convenient coincidence, Charles!” But of course one needn’t respond to the book that way. One might also wonder whether it really *is* a coincidence that such-and-such person just happens to show up at just the right moment, and this wonder can add to the enjoyment of the novel. When it’s revealed at the end of the book that the main characters have basically been following each other around the whole time, the coincidences make sense and can again generate a kind of satisfaction and enjoyment. Of course, some people will hate this kind of story and experience, but not taking note of this phenomenology would be to miss an important part of what reading and writing novels is all about. Similarly, we would be missing some of the true multicolored mixture of mathematics and its practice if we did not pay attention to this role for mathematical coincidence within the practice.

Clearly, I have not offered a *theory* of mathematical coincidence here. And given a commitment to the ordinary realism being illustrated in this dissertation, this is should come as no surprise. What I have done, however, is point to some of the important roles that the apparently strange phenomenon of mathematical coincidence plays within the practice of mathematics. My hope is that through the presentation of these roles and facts the phenomenon needn’t feel so strange anymore—that is, that some of our philosophical puzzlement has been dissolved—and that some insight into mathematics and its practice has been achieved at the same time. Let me emphasize one key characteristic of mathematical coincidence as I have described it here though before moving on to discuss the most

37 Note that this kind of question can be settled without coming to the conclusion that it’s *true or false* that $X$ is a coincidence. One way of settling the question would be to come to the opinion that the fact isn’t interesting and so just a coincidence or that it is interesting and so is no mere coincidence.
obvious objection to the view, which arises in relation to this characteristic. Calling something a mathematical coincidence or not, according to the Pascalian view presented here, primarily engages one in evaluating that fact as being uninteresting or worthy of attention, respectively. This being the case, its being true that some mathematical fact, say, is a coincidence shouldn’t be understood in terms of the world of mathematics verifying this claim.\footnote{Cf. (Floyd 2012: 232) for a similar claim about surprises in mathematics.} Rather, if we’re to talk about truth at all here, to say that it’s true that \(X\) is just a coincidence is to do no more than agree that \(X\) isn’t worthy of special attention.\footnote{\(X\) may be worthy of attention for other reasons, of course. I’m only claiming here that its interest doesn’t derive from there being a coincidence that sparks attention.} To say that it’s true that \(X\) is no coincidence is similarly merely to put it forward as something worthy of our mathematical attention.

**Does Anything Go?**

The suggestion that whether a mathematical fact or pair of facts is coincidental is not determined by what the world of mathematics is like is bound to meet with vigorous resistance. Perhaps it makes sense of how coincidence talk is used in practice; perhaps it nicely squares with judgments that coincidence comes in degrees in mathematics\footnote{For the desirability of such a feature, see (Baker 2009: 148).}, and perhaps it even helps us understand otherwise strange claims made by mathematicians (e.g., Davis claims that the mathematician to some extent brings mathematical coincidences into existence—if coincidence talk is in the business of focusing attention here or there, this seems like it might be a reasonable enough claim; if not, perhaps not\footnote{See (Davis 1981: 320).}); but surely there are facts of the matter here. Marc Lange is getting at something like this thought when he claims that

\[
\text{“[e]ven from a mathematically omniscient perspective, there are some mathematical coinci-}
\]

\footnote{\text{\[Floyd 2012: 232\] for a similar claim about surprises in mathematics.}}
Isn’t someone who takes the Pascalian view described above committed to a kind of extreme relativism here where anything goes when it comes to talking about coincidence in mathematics?

In responding to this worry and the surrounding cluster of issues it raises, it should first be noted that, as a matter of fact, mathematical practice does tend to converge on whether or not something is appropriately called a coincidence. So, evidently, not just anything does go here. Part of what mathematical practice does—like all other sufficiently developed practices—is establish norms for the usage of the terms of art, like ‘mathematical coincidence,’ within the practice. ‘Coincidence’ is a term we’re all familiar with prior to any exposure to its use within mathematics, but it’s important to see clearly both the similarities and differences between our ordinary usage and the uses the word finds within mathematics. Given the fact that to be part of the practice of mathematics is to be governed by the professional and informal norms of that practice, to be at least largely in line with other mathematicians’ judgments about coincidence will in part be constitutive of being engaged in that very practice. This being the case, there will be better or worse opinions about whether particular facts are coincidences or not, and these opinions will be judged by the standards of other practitioners. This fact alone should go a long way towards reducing the sting of this relativist worry. Nevertheless, the further question that ought to be asked when deciding how seriously to take the objection that anything like the Pascalian view of Section 4 implies that there’s no fact of the matter about whether something is a mathematical coincidence is, “What best explains these facts about mathematical practice?”

---

42 See (Lange 2010: 316). According to an account of explanation for which explanations are intimately related to the answering of why-questions, this claim of Lange’s would likely be judged to mistaken. E.g., (van Fraassen 1980: 130) suggests that an omniscient Being wouldn’t be in the business of explanation at all, so the distinction between coincidence and non-coincidence in Lange’s terms would disappear.

43 See, e.g., (MacIntyre 1981: Ch. 14) on this role of practice.

44 See (Field 2001) for a similar view in relation to his “evaluativist” account of apriority.
Gilbert Harman and David Wiggins have each suggested that realism (in the metaphysical sense) is the view one should opt for when the best explanation for convergence like the one just described is the existence of a relevant fact. In Harman’s examples, the convergence of a group of scientists on the thought, “There goes a proton,” is best explained by the existence of the proton, but convergence on the thought, “Setting cats on fire is wrong,” is not best explained by the fact that this action is wrong. Rather, this convergence can be better explained by other psychological and sociological facts about human beings in certain places and times. These facts can be used to give an argument for scientific realism and against moral realism. The relevant question for present purposes is whether convergence within mathematical practice on a judgment that, say, $X$ is coincidence is best explained by the fact that $X$ really is a coincidence, whatever we ultimately take that claim to mean. My claim is that the real existence of something coincidental is not what leads to the convergence here. Instead, this convergence is explained, as suggested above, by the shared practice governing the interests and appropriateness of certain judgments of those making the relevant judgments. In fact, even the Dictionary Analysis would seem to be committed to this view of the explanation of convergent views about mathematical coincidence: convergence on the view that $X$ is no coincidence tends to happen prior to any explanation being found, and since there’s generally no way to prove that $A$ and $B$ can’t be given a single, unified explanation, convergence on the view that $A$ and $B$ is merely coincidental can’t be reached on the basis of the fact that no explanation of this kind will be found.

Another familiar test for whether or not a domain of discourse should be thought of as being (metaphysically) realist is whether it exhibits what Crispin Wright calls “cognitive command.” A domain of discourse exhibits cognitive command, roughly, if it’s a priori that when there’s a dispute, one side or the other is mistaken. According to the Pascalian

---

45 See, e.g., (Harman 1977: Ch. 1.3) and (Wiggins 1987: 147, 149-151).
46 (Wright 1992: 92-93)
view, talk about coincidence fails to be a domain exhibiting cognitive command, but this fact should be considered a point in its favor. One of the objections to the Dictionary Analysis raised in Section 3 was that it had the consequence that any two mathematical facts turned out to be either a coincidence or not, but that this did not fit with our ordinary ways of thinking about coincidence at all. In fact, for many pairs of mathematical facts, it does not seem as if it really make sense to ask whether they’re coincidental or not. In these cases, one could insist that they’re either coincidental or not, but there’s nothing like an a priori certainty that one side or the other must be mistaken in this imagined insistence.

These responses to the worry that anything goes with mathematical coincidence talk as far as the Pascalian view is concerned are relatively brief, and the back-and-forth could obviously be extended (perhaps indefinitely). However, I take it that the general line of response presented so far is enough to both show the Pascalian view not to have unacceptable consequences and to indicate how this further conversation could be continued for those parties interested in such a continuation.

Conclusion

One’s view of the world and coincidences in it appears to be determined to a large degree by one’s intuitive ideas about how the world works. If I think that good things generally come to those who wait, I might not find it coincidental that my patience is so often rewarded. If I have no idea about how the world and occurrences in it are governed, nearly anything could be deemed a coincidence: the sun’s rising two days in a row might be a shock; the milk’s going bad after being left out on the radiator might be baffling; etc. The world of mathematics is one that we’re often in the position of not knowing our way about very well and of not having an intuitive understanding that allows us to see certain things as expected.
This is especially true when we are trying to break into unknown parts of the subject while new research is underway. It’s no surprise, then, that we seem to find coincidences so often in this domain. As we come to have hunches or insights or intuitions about how a particular corner of world of mathematics works and is governed, we aim to remove our feelings of uncertainty and confusion by turning the coincidences we find there from surprises into trivialities.47

By focusing on how talk of mathematical coincidence functions within normal mathematical practice, I hope to have shed some light on how this talk and these judgments push us forward in the quest for understanding in the domain mathematics. I also hope to have shown once again how the realist approach on display in this dissertation is a more effective one than the ordinary, and largely un-theorized, “practice-first” approach most often employed within contemporary philosophy of mathematics. Marc Lange, for example, clearly advocates the practice-first approach to the subject, but as soon as he finds himself in possession of the hammer of an account of mathematical explanation, other phenomenon in the field suddenly seem to start looking more and more like nails. The realist methodology being illustrated here aims to resist this sort of tendency as much as possible—it aims almost above all else to “teach you differences.”48 It’s only in the light of these differences and without theory-driven homogenization that we’ll be able to come to the complex and variegated understanding that this complex and multicolored subject demands.

47 Cf. (Rota 1997a: 93): “The quest for ultimate triviality is characteristic of the mathematical enterprise.”

48 This line of Kent’s in King Lear—“I’ll teach you differences” (1.4.82-83)—was considered by Wittgenstein as an epigraph for the Philosophical Investigations. See (Monk 1990: 536-537).
Chapter 5
Mathematical Induction and Explanation

In fact he was beginning to like very much arbors and ardors and Adas.
—Nabokov, Ada, or Ardor

Philosophers of mathematics have commonly taken mathematicians to judge proofs by mathematical induction to be paradigm cases of proofs demonstrating *that* a proposition is true without revealing the reasons *why* it is.\(^1\) Attempts to justify these largely intuition-based judgments, however, have been far less common. The focus in what follows will be on one attempt to supply this kind of justification through argument, along with some of the responses this attempt has generated.\(^2\) The debate is of interest here not so much because of the arguments and replies on offer, but rather because of what it shows about current thinking about mathematical explanation more generally. This being the case, it’s possible to avoid some of finer details of the debate and to instead use it to provide the context for a closer look at some of the assumptions made by those involved in the larger discussion about explanation in mathematics. There seem to me to be serious problems with the typical

\(^1\) See, *e.g.*, (Steiner 1978a: 136-137), (Hanna 1990: 9), (Giaquinto 2005: 78), and (Hafner and Mancosu 2005: 237). For the contrary opinion, however, see for example (Kitcher 1975: 265) and (Brown 1997: 177).

\(^2\) Marc Lange’s (2009) sparked the current debate. Replies to his paper include (Baker 2010), (Baldwin 2016), (Hoeltje *et al.* 2013), and (Cariani ms).
manner in which the subject is approached created in part by the charm the word ‘explanation’ can have on us. The project of this paper is primarily to point out some of these problems and offer suggestions about how to better pursue the relevant questions.

More specifically, the plan for the chapter is as follows. Since the general background to the concept of a mathematical explanation has already been provided in Section 6(a) of Chapter 2, after a brief remark about this background in Section 1, Section 2 will immediately begin discussing Marc Lange’s argument for the claim that proofs by induction are not explanatory. This section will also begin the consideration of a few of the replies his argument has spurred. Section 3 then catalogues some of the background assumptions made by Lange and his respondents, and argues that they are unwarranted. I shall further argue that these problematic assumptions can ultimately be seen to be based on the premise that mathematical explanation and the kind of explanation found in the natural sciences will share many common features. Section 4 examines reasons that might justify the acceptance of a premise such as this. After making a case that these reasons don’t do the necessary justificatory work, Section 5 shifts gears to offer an alternative way of viewing the notion of mathematical explanation. I close, in Section 6, by considering the consequences this alternative view has for the debate over inductive proofs and explanation in mathematics.

A Comment on ‘Explanation’

The display of the two different proofs of the identity

\[ 1 + 2 + \ldots + n = \frac{n(n + 1)}{2} \]

presented in Chapter 2.6(a) is the most commonly employed illustration of the distinction between a proof that explains and one that does not in the literature. However, for the
purposes of trying to decide whether or not inductive proofs are explanatory, fixing on this central distinction by means—almost alone—of a proof by induction to show what a non-explanatory proof looks like appears to be not particularly evenhanded. Rather than offer further examples, however, which might be more distracting than beneficial, I will proceed for now as if the ‘that vs. why’-distinction is clear enough for present purposes. Although it’s not ideal, this way of proceeding is common practice. The thought seems to be that examples of explanatory proofs that everyone agrees on are rare, and those that can be presented quickly without requiring much background are even rarer, so it’s necessary to make do. It’s simply assumed that the distinction being drawn is clear enough or familiar enough that these kinds of examples suffice.

Perhaps not though. In a trivial sense, of course, it’s not quite clear what distinction is being aimed at yet. The whole point of the investigation into explanatory proofs is to try to find a way to characterize what’s special about them. It cannot be expected, then, that a useful characterization of this distinction can be had right at the outset of the study. As long as the extra weight given to the example used to help latch onto the right concepts is kept in mind, it should be possible to counterbalance it if need be.

There’s surely something right about this line of thought. Yet, there should also be some serious concern about whether the familiarity of this distinction is merely illusory and about whether the supposed familiarity doesn’t cover up deeper issues that still need settling. That is, it may be that the ‘that vs. why’ distinction seems familiar, but only because it’s recognizable from its appearance it other contexts: Experimentation may reveal, say, that falling objects near the Earth accelerate at 9.8 m/s², and then a theory may do the job of explaining why. It’s not obvious, however, what reasons there could be for thinking that this way of making sense of the distinction will transfer to mathematical cases in any
straightforward way. This being the case, it seems as if more thought needs to be given to the very question of whether or not we should be speaking in terms of ‘explanatory’ vs. ‘non-explanatory’ proofs at all. This would certainly be the recommendation of someone taking the realist approach to the subject as outlined in chapter 1.

While it’s true enough that mathematicians can be found making remarks about explanatory proofs in many places, the question as to whether, how, and in what ways a proof deemed explanatory is used differently from one that is not remains open and critical. Once this type of question is answered, there’s still the further question of whether or not this difference in use and function is usefully characterized in terms related to explanation as well. Suppose, for example, a proof called explanatory by mathematicians is used as a paradigm for other proofs in some subfield or as providing a guide for how the foundations of that part of the subject might be rewritten or reconceptualized. We could perhaps imagine Alfred Pringsheim using a proof or proofs in this fashion when he was inspired to take a different approach to complex analysis based on the method of mean values of functions. It’s certainly not clear that putting a proof to use in this manner is usefully understood in terms of its being explanatory even if, say, Pringsheim himself happened to reach for this word in making sense of his alternative method. In other words, what’s primarily of interest is that some proofs look better than others in certain ways that we hope are usefully specifiable, and there’s good reason to attempt to understand more about what these ways are and in what ways these proofs play unique roles in mathematics and its practice. But, again, it’s plausible to believe that remain at a stage of understanding these phenomena at which we ought to constantly keep the question open as to whether we should really be trying to characterize the differences we’re after in terms of one proof being more explanatory.

---

3 This is not to say that anything as simple as this kind of transfer is going on in the actual cases of interest.
4 This is explicitly argued in (Zelcer 2013). See (Weber and Frans 2017), however, for a rebuttal.
5 See (Pringsheim 1925), and see (Mancosu 2001) for detailed discussion.
than another, instead of being, say, more insightful, more intuitive, more conceptual, more elegant, etc. The terminology of explanation inevitably carries with it a number of expectations, analogies, and clusters of closely-related concepts. We ought to still be seriously asking what reasons there are to take all this baggage on board. Much of the paper will be devoted to suggesting that that there aren’t very many good ones.

Lange’s Argument and Replies

With this basic background in hand, I will move on to examining Lange’s argument that inductive proofs are generally not explanatory. Again, the details of the argument will ultimately be of less importance than the assumptions in play, so we will only need to grasp the main ideas involved.

The Argument

Lange’s argument turns on the following feature of proofs by induction: If I want to prove that something is true for all positive integers, I can proceed in many different ways. I might first show that it’s true of 1 and then show that whenever it’s true of a number $k$ it’s also true of $k + 1$—this is the style of proof that usually comes to mind when people think of a proof by mathematical induction. However, I could have just as easily—in some sense—started by showing that the claim is true of, say, 100 and then showing both that whenever it’s true of $k$ it’s true of $k + 1$ and that whenever it’s true of $k$ it’s true of $k - 1$ for $k > 1$. Either way, I will have shown that the result holds for all positive integers in the end.

Since these two proofs only differ in the starting point chosen, Lange suggests that it would be arbitrary to judge one to be explanatory and the other not. But, says Lange, we can’t take them both to be explanatory. In the first version of the proof, if the proof is
explanatory, the claim’s being true of 1 is part of the explanation for why it’s true of all positive integers. And, since its being true of all positive integers can be taken to explain why it’s true of 100 in particular, the claim’s truth for 1 partly explains why it’s true of 100. In the second version of the proof, the roles of 1 and 100 are simply switched. As a result, if this proof is explanatory as well, the claim’s being true of 100 is part of the explanation for its being true of 1. This can’t be the case, however: “X explains Y” is an asymmetric relation. We’re meant to conclude, then, that neither proof is explanatory. And the conclusion generalizes.

Replies

Since Lange’s argument only relies on a few explicit assumptions, these are the obvious places at which to question it. Unsurprisingly, the replies the argument has generated have targeted just these points. I shall begin with the most direct attacks on the argument and move towards those that I take to strike more deeply at its roots.

The argument’s first potential weakness is the claim that both versions of the proof are equally explanatory. Since the proof that begins by showing the claim to hold of 1 and the proof that begins by showing it to hold of 100 both include a step showing that the claim’s truth for k implies its truth for k + 1, the second proof’s need for an extra step—to show that truth for k implies truth for k – 1 when k > 1—might be thought to make it less explanatory. That is, the first proof is minimal or less disjunctive in some sense, and, therefore, could reasonably be judged more explanatory than the second. This move, if successful, would block Lange’s conclusion.

The second point at which one might object to the argument is the relationship Lange needs to hold between a statement involving a universal quantifier and its instances. In

---

6 This is the path of response taken by (Baker 2010) and (Cariani ms).
7 This is the focus of the reply by (Hoeltje et al. 2013).
order to show that taking both proofs to be explanatory leads to a violation of the asymmetry of the explanation relation, Lange has to rely on some principle that says that the truth of a universal claim explains the truth of its instances. This kind of principle is employed, for example, in his reasoning about the first proof when he argues that the truth of a claim for all positive integers explains its truth for 100 in particular. The guiding idea seems to be that, if it were a law of nature that all dogs go to heaven, then this law would explain why each dog goes. Analogously here: the truth of the claim for all numbers explains its truth for particular cases.

Hoeltje et al. (2013) argue that this explanatory relationship between a universal statement and its instances does not hold in this case, however. They provide a number of reasons for thinking, rather, that a universal claim is explained by its instances, not the other way around. If the direction of explanation Lange needs to hold doesn’t, his argument is not able to establish that the asymmetry of explanation is violated, and the argument can, therefore, be resisted. (Baldwin 2016) also usefully criticizes this part of Lange’s proof without taking such a strong stand on the direction of explanation that “really” holds between a universal claim and its instances.

Other direct responses to Lange’s argument have taken the form of presenting counterexamples. Both (Cariani ms) and (Wysocki 2017) give examples of proofs where the “downward” direction of the second inductive proofs can’t work without reestablishing the base case at 1. For example, Wysocki defines a family of functions on the positive integers as follows. For all \( f \) in this family,

1. \( f(1) > 8 \) and
2. for all \( n \), \( f(n) < f(n + 1) \).

It’s possible to prove that all the values of every function in this family are greater than 8 in two different ways corresponding to the two ways taken in Lange’s original argument.
First, one could choose an arbitrary function in this family, show that it takes a value greater than 8 at 1 and then use the second clause of the definition of the family to show that all other values must be greater than 8 as well. Or, second, one could again choose an arbitrary function and assume that it’s value is greater than 8 at some $k$, show that it must take values greater than 8 for larger values of $k$, and show that from $k$ downwards the values are still greater than 8. Yet, this second proof still must make use of the fact that $f(1) > 8$ just as the first did. Therefore, the second proof is not eligible for partially explaining the first proof’s base case as the argument requires.

A more conceptual response to Lange’s approach to the problem of the explanatory value of proofs by induction can be seen in the insight of (Baldwin 2016) that Lange begins by discussing whether or not a particular kind of proof is explanatory but immediately reduces this to thinking in terms the explanatory value of particular steps in the proof. If it’s proofs as a whole that do explanatory work, as we might think theories do in science more generally, this approach to the problem jettisons a global perspective much too quickly.

Finally, from the perspective of the realism being advocated and illustrated in this dissertation, two responses to Lange’s argument should immediately come to mind. First, discussions of the relative explanatory value of proofs in actual mathematical practice (to the extent that it occurs) seems to be most often carried out in concrete terms: “Why does this particular proof do a better job explaining the result than that one?” While it’s clear enough that sometimes the form a proof takes can render it non-explanatory in some sense (e.g., if it’s a mere enumeration of cases when more unified means are available8), it would be very hard to imagine that a form of argument as necessary and ubiquitous to the subject as mathematical induction could on its own render proofs employing it defective in this

8 Cf. (Hardy 1940: 113).
Secondly, although both of the proofs of the simple arithmetic identity under discussion would be considered equally good from the perspective of the provability of this identity, in practice the second proof Lange offers would certainly be called “wrong” in one way or another. For example, if a student submitted this kind of proof as part of her homework, any grader would remark on the strangeness of the approach; if a paper were submitted to a journal with this kind of disjointed induction in it, it would be returned for revision; etc. The second proof Lange presents us with is, therefore, from the standpoint of mathematical practice already disqualified from competing with the first one as a candidate for an explanatory proof. As such, the argument itself should not be taken to raise any particular problems for the explanatory power of mathematical induction.

In fact, from the standpoint of practice, it looks as if all that really matters in an inductive proof is the inductive step that establishes that the result holds for countably many instances. It’s this that makes the method powerful and in fact necessary for proving many results that hold beyond the finite. Even if finitely many cases fail early on, if an induction can be used to show that further objects satisfying whatever property is in question can be found without end, there’s still a potentially interesting theorem remaining. This being the case, even if the explanatory value of both the proofs Lange gives in his article was to come under consideration, it’s easy to imagine the conclusion being drawn that they’re “essentially” the same, so not in competition. Again, induction as a potentially explanatory method of proof wouldn’t be endangered by the argument.\footnote{Clearly, the practice-based responses being put forward in the last two paragraphs are in tension with one another. To the extent that both are plausible representations of how mathematical practice would view...}

\footnote{Cf. (Cariani ms), which likens the question of whether proofs by induction are explanatory to the question of whether proofs by cases are.}

\footnote{Although, see (Hafner and Mancosu 2008) for a different view from the perspective of real algebraic geometry as practiced by Gregory Brumfiel (1979).}
Background Assumptions

We now have Lange’s basic argument and a few of the responses to it that question explicit (and some not-so-explicit) assumptions he makes on the table. In this section, I would like to step back and look further into some of the deeper background assumptions in play among those involved in the dispute. Each of these assumptions is rooted in the thought that we should expect mathematical explanation and explanation in the natural sciences to be quite similar. I hope to show that the principles which this assumption motivate are either mistaken or of uncertain relevance.

Asymmetry

Lange’s argument, as noted, makes few explicit assumptions. One of the major assumptions he does make, however, is that “[r]elations of explanatory priority are asymmetric. Otherwise mathematical explanation would be nothing at all like scientific explanation.”12

As Alan Baker notes however, Lange’s argument, if successful, only shows that, if both inductive proofs are explanatory, we have a case in which $A$ partially explains $B$ and vice versa. And it’s not obvious that this is impossible even in standard cases of scientific explanation.13 (Woodward 1984) goes so far to suggest that the relation “… explains …” is not always asymmetric in ordinary science either.14 Nevertheless, we still should ask

---

12 (Lange 2009: 206)  
13 (Baker 2010: 682)  
14 An example he gives of potentially symmetric explanation is the fact that in General Relativity the distribution of mass-energy throughout a spacetime depends upon the intrinsic geometry of that spacetime and the intrinsic geometry of a spacetime depends on the distribution of mass-energy throughout the spacetime. See (Woodward 1984: 438-439).
whether we ought to accept the claim that, in mathematics, we can never have a case in which \( A \) explains \( B \) and \( B \) explains \( A \).

Given the usefulness and prevalence of dualities in mathematics, it seems as if there are reasons not to rule out symmetric explanation as impossible. The following are just a two quick examples of well-known dualities. Category theory is another area rich in this phenomenon.\(^{15}\)

The most famous area in which duality plays a major role is projective geometry. Here, the “Principle of Projective Duality” says that if you replace ‘point’ by ‘line’, ‘colinear’ by ‘concurrent’\(^{16}\), etc. in a theorem, you’ll end up with another theorem. As an example, Pascal’s Theorem states that if you extend the opposite sides of a hexagon inscribed in a conic section, the intersections of these lines are colinear. (See the figure below.) The dual of this theorem, known as Brianchon’s Theorem, states that if you connect the opposite vertices of a hexagon circumscribed on a conic section, these lines are concurrent.

---

\(^{15}\) See, e.g., (Leinster 2014).

\(^{16}\) Lines are concurrent if they intersect in a single point.
It’s not clear however why in a mathematical case like this one there’d be any reason to rule out *a priori* that Pascal’s Theorem explains why the lines connecting the opposite vertices of a circumscribed hexagon intersect in a single point, and that Brianchon’s Theorem explains why the intersections of the opposite sides of an inscribed hexagon are colinear; i.e., that Pascal’s Theorem explains Brianchon’s Theorem and Brianchon’s Theorem explains Pascal’s. It’s not as if one causes the other after all, and without causal ordering of the kind that would be expected to be at play in explanation in other sciences, the need for explanations to be asymmetrical becomes far less pressing.

Similarly, in algebraic topology, Poincaré Duality, “[t]he crucial starting point for applications of algebraic topology to geometric topology,” states that for closed, orientable *n*-dimensional manifolds, the $k$th homology group $H_k(M, \mathbb{Z})$ is isomorphic to the $(n - k)^{th}$ cohomology group $H^{n-k}(M, \mathbb{Z})$ for any $k$. Again, as opposed to other scientific contexts in which explanations are often conceived of in causal terms, there’s no obvious reason for saying that it’s impossible that $M$’s having the $k$th homology group it does explains why it has the $(n - k)^{th}$ cohomology group it does and *vice versa*. Taking either direction of explanation to be more fundamental would appear to be an arbitrary choice, but either (alleged) explanation seems to be a perfectly good way to answer to the question, “Why does $M$ have this particular (co)homology group?”

I don’t expect that these cases on their own could make a convincing argument that explanation in mathematics *is*, in fact, often symmetric. It’s just that cases like these—and the many more that could have be given—should make us reevaluate taking as a starting point any assumption that makes this an impossibility. In fact, the primary reason for thinking that explanation in mathematics ought to be taken to be asymmetric seems to be that it

---

17 (May 1999: 149)
allows the maintenance of the analogy with ordinary scientific explanation. However, we shouldn’t let this desire shape our investigation if we want to maintain a realistic approach to the subject.

Occam’s Razor and Simplicity

Like Lange, those replying to his argument take principles about explanation in the natural sciences to transfer unproblematically to the mathematical case. Two of those principles, which we find appeal to here, are the value of ontological parsimony and overall simplicity.

Baker (2010), for instance, appeals to Occam’s Razor, which “give[s] credit to theories that are qualitatively parsimonious, in other words theories that postulate fewer kinds of entities or mechanisms,” to justify his claim that the inductive proof starting from 1 is more explanatory than the one that begins at 100. Fabrizio Cariani (ms) similarly makes an appeal to the virtue of simplicity to arbitrate between the two proofs Lange presents. He writes, “[I]f mathematical explanation is to be anything like scientific explanation, simplicity must function as a tie-breaker of sorts.”

Once again, however, we need to ask what we’re to make of these principles when applied to mathematical explanations. The following (long) quote from Jeremy Avigad speaks directly to the uncertain relevance of Occam’s Razor in mathematical contexts.

[W]e learn that, in a sense, we do not need a very rich universe of mathematical objects, nor do we need strong principles of reasoning [...]. Infinitary objects like topological spaces, manifolds, and measures can be coded as suitable sets of numbers, finitary objects can be coded as numbers, and some basic axioms of arithmetic are then enough to justify the desired mathematical inferences. So, in a sense, mathematics does not need analysis, algebra, or geometry; all it needs is a weak theory of arithmetic.

18 (Lange 2017: 311), however, “gestures” at the argument that the priority of axioms over theorems might be a source for some asymmetry in explanation here. Perhaps some of the findings of reverse mathematics would show this not to be a viable place to look for asymmetry though.

19 (Baker 2010: 684)

20 (Cariani ms: 6)

21 [From various proof-theoretic studies and works in reverse mathematics.]
But certainly there is a sense in which this is false: *of course* mathematics needs topological spaces, manifolds, and measures [...] Put simply, the success of the reductionist program poses the philosophical challenge of explaining this use of the word ‘need’: if logical strength is not everything, what else is there?22

I take it that Avigad is correct here, and that principles of ontological parsimony don’t typically motivate in actual mathematical practice. This can be further seen in the way much larger domains, e.g., working over the real numbers instead of the rationals, is valued even when it’s not strictly necessary. This being the case, appeals to Occam’s Razor shouldn’t be expected to do much work here without a lot of preliminary preparation for its application.

How to apply considerations of simplicity to adjudicate between two potentially explanatory proofs is similarly non-obvious. This is because, as Avigad (ms) again points out, mathematical simplicity doesn’t seem to line up in any straightforward way with notions of simplicity in science: “It seems to have more to do with streamlining our thought processes and modes of expression than simplifying models and reducing the number of parameters.”23

Consider the following examples.24 The Fermat-Euler Theorem, which states

\[ a^{\varphi(n)} \equiv 1 \pmod{n}, \] 25

can be proved using methods of modular arithmetic that don’t require much setup, or using methods from basic group theory that require the introduction of the concepts of a group and a coset, as well as a proof of Lagrange’s Theorem. With the relevant background, the group theoretic proof seems simpler, but there are reasons one might prefer the proof with less conceptual requirements. Similarly, many of the proofs in real analysis are (apparently) simpler once concepts from topology are introduced. They are simpler in the sense that they

---

22 (Avigad 2003: 275, emphasis in the original)
23 (Avigad ms: 1)
24 The first is drawn from (Avigad ms)
25 Where \( \varphi \) is Euler’s “totient” function.
are easier to think through or state, however; not necessarily simpler in the sense of requiring fewer, or weaker, assumptions or a smaller store of theoretical entities. Or consider the way in which ostensibly more complicated counting methods according to which, e.g., what appears to be one point is to be counted twice are considered to be the “right” ways to count in place of our ordinary ways. Once more, it’s not at all apparent how we are to apply these standard kinds of measures in mathematical contexts.

To be clear, I do not mean to suggest that sense cannot be made out of principles like Occam’s Razor or appeals to simplicity in mathematical studies. It’s just that work does need to be done to make sense out of them: We oughtn’t simply assume they are already applicable.

Instances and Generalizations

Recall that the second method of replying to Lange’s argument, pursued by Hoeltje et al. (2013), suggested that Lange was wrong to hold that a universally quantified truth explains its instances. Instead, they claim, we should accept

**Rule-By-Case:** A universally quantified truth is explained by its instances.\(^{26}\)

This principle may sound familiar from non-mathematical contexts: What explains why it’s true (suppose) that all pigs are pink, if not the many instances of pink pigs? But, once again, how this kind of principle should be applied to mathematical cases, if at all, requires some thought.

We should expect at least some trouble applying principles like Rule-by-Case, since universally quantified statements in mathematics often deal with infinite domains, where problems familiarly tend to arise. It seems as if something has gone wrong when a principle like Rule-By-Case leads to claims like the following.

\(^{26}\) (Hoeltje et al. 2013: 6-8)
[I]n the infinite case every instance makes only an infinitesimal contribution to the truth of the universal statement. Nevertheless, an infinitesimal contribution is a contribution; and we see no harm in countenancing cases where an infinite number of instances each partially explain the corresponding universal statement.\textsuperscript{27}

The thought seems to be that every time a statement is true of a number, it does a very small amount of work towards explaining why it’s true for all numbers.

We ought to doubt this kind of explanation though. Of course, it’s unclear what to make of the notion of infinitesimal contributions to an explanation. But, also, if a principle like Rule-by-Cases were valid, any statement about the positive integers with only finitely many counterexamples and any true universal claim would have precisely the same number of instances contributing to explaining why it’s true—infinitely many. True, there might be a principle that says that a counterexample introduces an infinite amount of anti-explanatory force or something similar, but, despite that possibility, this way of thinking about the relationship between instances of a universally quantified statement in mathematics and that statement itself simply seems to be flawed from the start. This isn’t to say that we should embrace the alternative principle that a universally quantified statement explains its instances. Rather, there are reasons for supposing that, in cases like these, to say that a statement is true for all positive integers \textit{just is} to say that it’s true for 1 and that, when it’s true for any \textit{k}, it’s also true for \textit{k + 1}.\textsuperscript{28} In other words, it’s nothing but the fact that a mathematical induction can be carried out that explains the validity of the universal claim.

\section*{Justifying the Assumptions}

I have just spent a lot of time trying to make the case that thinking about mathematical explanation in terms familiar from other forms of scientific explanation leads us to assume,

\textsuperscript{27} (Hoeltje \textit{et al.} 2013: 7-8)
mistakenly, that similar principles can be brought into play in our theorizing; principles
that often do more to muddy the waters than anything else. But maybe there’s still a case
to be made for accepting the analogy between scientific and mathematical explanation that
trumps these considerations. That is, perhaps this analogy creates some false hopes, but is
still in essentials apt.

This section considers some reasons that are often put forward in order to justify taking
this analogy seriously.\textsuperscript{29} Despite the fact that these kinds of sentiments are quite common,
I take it that all but the final reason are bad.

Optimism

One reason a connection is often made between mathematical and other kinds of scientific
explanation seems to be a kind of optimism. Philosophers of science have spent a long time
trying to make sense of explanation in science, and it would be nice if this work could be
applied to the case of explanation in mathematics as well. If it turns out that explanation in
mathematics is very different than the general scientific case, “we won’t be able to rely on
very many important historical case studies for scientific explanation.”\textsuperscript{30}

This can be taken to be true enough. But if these are not actually similar kinds of ex-
planation after all or even usefully thought of in terms of explanation, we of course shouldn’t
rely on many important historical case studies of scientific explanation. No one would deny
that it would be nice if all the work philosophers and scientists have put into to coming to an
understanding of scientific explanation could simply be transferred over to the mathematical
case, but this appears to be no more than wishful thinking at this point.

A closely-related reason for attempting to treat scientific and mathematical explanation
similarly may also stem from a kind of scientism. Philosophical naturalism can have the

\textsuperscript{29} See (Jansson 2017) for a recent attempt to draw out this analogy in detail.
\textsuperscript{30} (Betti 2010: 285)
unfortunate side-effect of leading some philosophers to believe that to the extent that a
subject is unlike the natural sciences, there is to that same extent something wrong with
the subject. We can see this kind of thinking expressed by Arianna Betti (2010: 284) as
follows.

Surely mathematics is worth its name as a science as much as physics is? If so, explana-
tion in mathematics is a legitimate form of scientific explanation as much as explanation
in physics is.

It’s as if not being in the business of explanation excludes an activity from being science,\textsuperscript{31} and for a philosopher to take anything away from the stature of mathematics by denying that
it’s in this business just as much as physics is would be deemed ludicrous. But again this
kind of reasoning obviously is not going to do much justificatory work even if it is one of
the main reasons philosophers of mathematics have been loath to consider whether or not
distinct forms of explanation really are at work in mathematics and other sciences.

Hyper-Physics

Another bad reason for expecting mathematical explanation to mirror scientific explanation
is created by taking too seriously the metaphors mathematicians sometimes use to describe
their work, e.g., that they are exploring the world of numbers and so on. The thought is,
natural sciences try to explain the physical world; mathematics similarly attempts to explain
the hyper-physical world, using similar explanatory principles. These sorts of metaphors
are a key place for a philosopher taking a realist approach to the subject to be on guard
against the ways prose tends to lead us on.\textsuperscript{32}

Betti (2010: 288), again, appeals to this kind of reasoning. “[A] proper or real sci-
ence” is one (among other things) that “concern[s] a specific set of objects or [is] about a

\textsuperscript{31} Cf. (Brinck et al. 2010: 3-4).
\textsuperscript{32} This is a point continually stressed by Wittgenstein in his attempts to get us to reject the view of mathematics
certain domain of being(s).”\textsuperscript{33} Mathematics, being a real science, is then about a specific set of objects and tries to explain their behavior just as physics, chemistry, or biology does. As such, we should predict similar explanatory practices.

It is, of course, true enough to say that mathematics is about mathematical objects. But to infer that it’s useful to characterize mathematics and physics as nothing more than two instances of \textit{proper} or \textit{real} science from the truism that these subjects are both about a domain of objects is to overlook the fact that these are very different practices—different enough, in fact, that we should not presuppose that the same notion of explanation is in play in both—and objects in very different senses.

Jonathan Shaheen’s recent work on “metaphor-generated polysemy” might offer some insight into the temptation to hyper-physics thinking here.\textsuperscript{34} Consider the simple example he uses to introduce the idea.\textsuperscript{35} It’s reasonable enough to assume that at one time the word \textit{leg} was only used to refer to the body parts of live or dead animals. But as creative language users began to see similarities in shape or purpose, it’s similarly easy enough to imagine that ‘leg’ came to be used for things attached to tables and chairs. At first, this kind of usage seems to be dependent on this kind of creativity and to still be metaphorical in some sense. However, as time again continues on, we understand ‘leg’ to be polysemous with more than one literal sense.\textsuperscript{36} The metaphor has generated the polysemy. The use to which Shaheen puts this concept is in his attempt to understand non-causal explanation of the sort that might be found in mathematical explanation. If we think of causal explanation as playing the role of the anatomical parts referred to by the early uses of ‘leg’ and non-causal explanation as offering a new sense of ‘explanation’ arrived at by and after pro-

\textsuperscript{33} This characterization is drawn from (de Jong and Betti 2010).
\textsuperscript{34} See (Shaheen 2017b).
\textsuperscript{35} (Shaheen 2017b: 557)
\textsuperscript{36} The distinction between metaphorical and literal language here needn’t be drawn too precisely to make sense of this change. See (Yablo 1998) for doubts about the usefulness of such a distinction.
longed metaphorical expansion of the concept of causal explanation, we can see non-causal explanations as being “count[ed] as explanatory insofar as they fit the causal metaphor.” Further, the two senses might seem to be unified through some kind of shared structure that led to the initial metaphorical uses.

This theory of metaphor-generated polysemy can go some of the way to explaining the prevalence of the thought that scientific and mathematical explanation ought to be treated very similarly. And an explanation of this kind is surely necessary. It remains to be asked, however, how useful and extensive this shared structure can be for our thinking about what’s commonly referred to as explanation in mathematics. The doubts about the prospects of this utility that have been raised throughout this chapter should leave us skeptical at least.

Objective Explanation

The thought that only a form of explanation closely modeled on scientific explanation can provide objective explanation is a final motivation for treating mathematical explanation and scientific explanation as being largely one and the same. The route to this conclusion might go as follows.

There’s a very natural sense in which every proof shows, precisely, why a particular proposition is a theorem: if there’s a question about why the sum of the first $n$ natural numbers is $n(n + 1)/2$, a proof answers the question; if someone asks, “Why is $R[X]$ noetherian if $R$ is?” we typically show them a proof of Hilbert’s Basis Theorem; and so on. If someone is presented with these proofs and says, “I see all that, but why is the result true?” there’s some reason to think he has not understood the proof. However, this sense of ‘explain’ has seemed to some to be too subjective or psychologistic. As philosophers, we are not

---

37 (Shaheen 2017a: 860)  
38 (Shaheen 2017b: 554)  
interested in the “mere psychology” of what it takes to explain a theorem, say, to a classroom. Rather, we want to know about the real or deep reasons that make a theorem true.\textsuperscript{40} Since explanation in the sciences is the paradigm case of providing such information, we are quickly led to thinking of mathematical explanation along the lines of scientific explanation. I have argued though, that, once this move is made, misleading analogies abound.

Without getting into the difficult and longstanding debate about what exactly a theory of explanation is supposed to be a theory of,\textsuperscript{41} it’s still possible to say something about how mathematicians seem to think about the desirability of explanations that are objective in this sense in at least some of their work. While it’s correct that the kind of explanation the search for which appears to motivate some research in mathematics is indeed of the objective sort, it’s a mistaken assumption that this is only obtainable on the model of explanation at work in the natural sciences. The next section briefly attempts to make this case.

**Proof and Explanation of Meaning**

The thought that explanation in mathematics can usefully be thought of as involving explanation of meaning, which was first discussed in Section 8(b) of Chapter 2, can further be appealed to in order to give objective explanations. The goodness or badness of an explanation of meaning is something that can be established objectively, and such explanations also are properly thought of as needing to be properly attuned to a particular audience.\textsuperscript{42} These sorts of explanations won’t be explanations of the kind we ordinarily find in the natural sciences these days. However, on a conception of science advocated, e.g., by Aristotle and Aquinas, explanations of meaning were very much to the point.\textsuperscript{43}

\textsuperscript{40} Cf. (Hafner and Mancosu 2005: 218).

\textsuperscript{41} My own view is most in line with those who take a theory of explanation to be essentially a theory of answers to why-questions. See, e.g., (Skow 2016: Ch. 2), (van Fraassen 1980), and (Bromberger 1992).

\textsuperscript{42} See (Thurston 2006) on the importance of attention to audience in offering mathematical explanations.

In order to arrive at a proper understanding of what the good life for man is, Aristotle would suggest that one must consider what ordinary people say it is; consider arguments for and against these various answers; and while doing this come to a better and better understanding of the original object of inquiry. “The good life for man is the activity of eudaimonia” is something that everyone would agree to, but it’s not a statement whose meaning is clear until this further investigation gotten seriously underway.\textsuperscript{44} A similar process seems to be in play when one arrives at an explanatory proof. Perhaps the mathematician starts out with something that seems to be true. This might be on the basis of its holding for numerous cases or on analogy with some other truth. This potentially true claim may then be tested in more cases in an attempt to show it’s wrong. At the same time, reasons for its holding true will be searched for. All the while the meaning of the initial statement will be becoming clearer. It may even be modified slightly in the process. One may stumble on a proof that shows the statement to be true, but doesn’t reveal why it is. So, the process may continue on. It seems as if it’s simply a mistake to think that one can immediately see the full content and meaning of a mathematical statement just because each of the words involved in it is understood. Through the process of trying to prove something to be true, we learn more about what it’s telling us.

Again, it’s not at all uncommon to see this notion of explanation appealed to in mathematical texts. Below are just a few examples, many of which line up nicely with the varieties of mathematical explanation noted in (Hafner and Mancosu 2005: 218-221).\textsuperscript{45}

1. This theorem explains the meaning of the equations of variations: they describe the action of the transformation over the time from \( t_0 \) to \( t \) on the vectors tangent to the phase space [...].\textsuperscript{46}

\textsuperscript{44} See (Aristotle 1999: I.7-8)  
\textsuperscript{45} Chapter 2.(b) contains a few other examples of this notion of explanation being employed.  
\textsuperscript{46} (Arnol’d 1992: 281)
2. Before we proceed to the proof of this classical theorem, let us first record some of its interesting consequences. (These consequences, in fact, help clarify the meaning of the theorem itself.)

This way of conceiving of mathematical explanation also has some philosophical precedent, for example, in the works of Wittgenstein, who often makes remarks like, “If you want to know what is proved, look at the proof,” and, “[T]he proof belongs to the sense of the proved proposition.” Further historical roots for this manner of thinking can be found in Riemann, Poincaré, and Federigo Enriques, if we take—as I do—the idea (also found in Wittgenstein) that “mathematics forms concepts” to be a manifestation of this same idea.

There is, then, evidence that explanation as explanation of meaning is often found in mathematical contexts, as well as a respectable philosophical lineage. This style of explanation also seems able to deliver the objective explanations most commonly sought: If a good proof explains the meaning of the proposition it proves, this is not just an explanation for me. As such, it’s not the case that only the traditional scientific model of explanation can produce the objective explanations we are after.

I will close by quickly considering how this model of explanation might answer the question of whether or not proofs by induction are explanatory.

**Are Inductive Proofs Explanatory?**

First, it seems as if we should be very skeptical of any argument claiming that form alone suffices to rule a proof explanatory or not. Given the wide variety of the kinds of theorems

---

47 (Lam 1991: 157)
48 (Wittgenstein 1933/1978: 369, 375, emphases in the original)
50 (Poincaré 1905/2001: 58-59)
51 (Enriques 1914: 184-185)
provable by mathematical induction,\textsuperscript{53} it should be expected that this decision needs to be made largely on a case by case basis.\textsuperscript{54} This is what I shall suggest.

Recall our first example, the two proofs demonstrating that

\[
1 + 2 + \ldots + n = \frac{n(n + 1)}{2}.
\]

The first proof by mathematical induction can be carried out almost mechanically, and, therefore, does not provide any real insight into what the theorem is actually saying. The more pictorial proof, on the other hand, does tell us something about what the ‘\(n(n + 1)\)’ means and why the division by 2 is necessary. There is, therefore, reason to suggest that this proof is more explanatory than the inductive proof—it does more to explain the meaning of the identity.

\[
\begin{array}{cccccc}
1 & + & 2 & + & \ldots & + & n \\
n & + & n - 1 & + & \ldots & + & 1 \\
\hline
(n + 1) & + & (n + 1) & + & \ldots & + & (n + 1) \\
\end{array}
\]

This is repeated \(n\) times.

In a case where what’s being proved is something like the efficacy of an algorithm, say, the Euclidean algorithm for finding the greatest common divisor (GCD) of a pair of integers on the other hand, an inductive proof seems to show exactly why the algorithm works, as well as reveals the meaning of the number produced as the GCD sought. So, here is a case in which it might be reasonable to say that an inductive proof is, in fact, explanatory.

As a final example, consider a proof of a version of Cauchy’s theorem that any finite group \(G\) has a subgroup of order \(p\) for any prime \(p\) dividing the order of \(G\) that only takes into

\textsuperscript{53} E.g., the Artin-Schreier Theorem, the Cantor-Bernstein Theorem, the Chinese Remainder Theorem, Fermat’s Little Theorem, the Hahn-Banach Theorem, Pick’s Theorem, Ramsey’s Theorem, the Schur Decomposition Theorem, Taylor’s Theorem, and Tychonoff’s Theorem. See (Gunderson 2010).

\textsuperscript{54} This is the verdict ultimately reached by some of Lange’s respondents. See (Hoeltje et al. 2013: 10-11) and (Cariani ms: 8).
account the abelian case. If one is going to prove the abelian version of Cauchy’s theorem
directly,\footnote{That is, instead of deducing it from Cauchy’s theorem proved à la (McKay 1959) or deducing it from the
Sylow theorems or from the fundamental theorem of finite abelian groups as in (Mac Lane and Birkhoff 1999).} the only sane option is to use complete induction.\footnote{See, for example, (Dummit and Foote 2004: 101-102), (Lang 2002: 33), (Herstein 1964: 51-52), (Rotman 1995: 73).} (If one were to try to start at
a finite abelian group of order \(k\), the problem of classifying the abelian groups of order \(k\)
now stands in the way.) Consider the following sketch of the standard proof.

Since the result is trivial if \(|G| = 1\), we can assume that there’s an element \(g \neq 1\) in \(G\).
If \(p\) divides \(|g|\), we can write \(|g| = pn\) for some \(n\). It turns out that \(g^n\) will generate a
subgroup of order \(p\). If \(p\) doesn’t divide \(|g|\), consider \(\langle g \rangle\) the subgroup generated by \(g\).
Since \(G\) is abelian \(\langle g \rangle\) is normal in \(G\). By Lagrange’s theorem, \(|G/\langle g \rangle| = |G|/|\langle g \rangle|\).
Since \(p\) doesn’t divide \(|g|\), it does divide \(|G/\langle g \rangle|\) which is less than \(|G|\). So, by the
induction hypothesis, there is an element \(g'\) of order \(p\) in \(G/\langle g \rangle\). This element can be
used to produce one of order \(p\) in the original group.\footnote{(Dummit and Foote 2004: 102)}

Does this proof illuminate the meaning of the theorem being proved? I think there’s reason
to believe that it does. By considering properties of the various possible nonidentity ele-
ments chosen in the initial part of the proof, one sees what a subgroup of order \(p\) can look
like in a given case as well as how it is constructed. It further helps to clarify what follows
from a group being finite and abelian by allowing one to make sense of the suggestion of
Joseph Rotman that “the more complicated the prime factorization of \(|G|\), the more com-
plicated the group” by starting with the “local” case where only one prime divides \(|G|\).\footnote{See (Rotman 1995: 73).}

Finally, the proof of the theorem allows one to see how information about a quotient group
can be used to determine facts about the original group as well as possible obstructions to
this sort of methodology.\footnote{See, again, (Dummit and Foote 2004: 102).} The proof can, therefore, be seen as providing an explanation of
the meanings of both the theorem itself and some of the key terms involved in its statement.

Again, one might want to argue that this theorem is better explained as an easy con-
sequence of the Sylow Theorems,\footnote{See (Herstein 1964) for this view.} but we shouldn’t expect the explanatory value of a proof
to be an all or nothing affair. A way of thinking about explanation in mathematics that allows us to find explanations of varying degrees of success, therefore, should be preferred in any case.

These are merely a few quick cases illustrating the variability of the explanatory power of proofs by mathematical induction. Others will have to be judged individually, as should be expected. I hope to have pointed out, however, a way of thinking about mathematical explanation that is versatile enough to provide the means to make a reasonable judgment in each case of interest.
Bibliography


Hardy, Godfrey. 1940. A Mathematician’s Apology. Cambridge University Press.


Lange, Marc. forthcoming. Mathematical Explanations that are Not Proofs. *Erkenntnis*.


153


Newton, Isaac. 1769. Universal Arithmetic: Or, a Treatise of Arithmetical Composition and Resolution. W. Johnston.


