WHAT MATHEMATICAL EXPLANATION IS NOT

by

MARK ZELCER

A dissertation submitted to the Graduate Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York
2009

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

`Melvin Fitting`

Date — Chair of Examining Committee

`Iakovos Vasiliou`

Date — Executive Officer

`Rohit Parikh, Advisor`

`Melvin Fitting`

`Samir Chopra`

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

WHAT MATHEMATICAL EXPLANATION IS NOT
by
Mark Zelcer

Advisor: Professor Rohit Parikh

What Mathematical Explanation is Not argues against the possibility of there being a plausible, non-psychological, account of mathematical explanation.

Providing explanations has generally been considered to be an important goal of science. Therefore philosophers of science have considered it vital to give an analysis of what explanations are and what form they take.

Our study begins by first distinguishing between various notions of "-explanation". Generally, "explanation" is used when we are offering some pedagogical elaboration, clarification, exposition, or analysis of a part of the physical universe. But since the ancient Greek period philosophers and scientists have been using the term "explanation" in a more technical sense to refer to an account of some phenomena, or the origin of the phenomena, in the physical universe. Famous accounts of explanation include Aristotle's "four causes" account, Hempel's Deductive-Nomological Model, Michael Friedman and Philip Kitcher's unification models, and Bas van Fraassen's why-question account.

Recently, various philosophers have claimed that there is an analogous concept in mathematics - mathematics offers explanations in the same way that science does. I argue that there cannot be explanations in mathematics akin to those in science. I argue that accounts of explanation modeled after existing accounts of scientific explanation - the D-N model, the causal model, unification models, and erotetic accounts - all fail when applied to mathematics.

Mark Steiner's model, which has no counterpart in the philosophy of science is also shown to be problematic. I also examine Paolo Mancosu's views on mathematical explanation and show that his stance on mathematical explanation is also flawed.

Finally I conclude with some positive arguments showing that there are desiderata for any theory of explanation (as we currently construe it) many of which will not accommodate mathematics. So if the necessary components for explanations cannot apply to mathematics, we will never have an adequate theory of explanation in mathematics.

Acknowledgements

Writing this dissertation, I have accumulated more debts of gratitude than I could possibly repay. While I know that these few sentences cannot convey much, certainly not all I wish to convey, I wish to record an acknowledgment of the help I received, as well as an offer of my genuine and heartfelt gratitude.

First, Rohit Parikh has always been a quintesential *mensch.* I benefited immensely from his extraordinary mathematical abilities, hard work, generosity, and not least for introducing me to the finer points of game theory and social software. I will always cherish his friendship and will never forget all those Indian lunches. I was fortunate to work with him and he will always be a model I will endeavor to emulate in his breadth of learning, depth of understanding, and good-hearted interactions with others.

Noson Yanofsky's help was indispensable and I could not have made it through some of the stages of this dissertation without his patience, mathematical insight, and tasty shnitzel. Thanks too to Hadassah for the tests that kept me on my toes. Arnold Koslow has been helpful to me since I was an undergraduate and I hope that this dissertation is at least somewhat

convincing. I owe a debt of *hakarat hatov* to Tziporah Kasachkoff too who made my stay in graduate school (and once in Jerusalem) far less painful than it would have been otherwise and to David Rosenthal, who has always been extremely generous with his time and his advice. I always appreciate it (even though I suspect he thinks I never listen). Mel Fitting, Joseph Dauben, and Samir Chopra were also extremely helpful as members of my dissertation committee. Thanks also to Michael Levin for comments on an early version of a few parts of this dissertation. Rosemary Iannuzzi and the late Shirley Tolbert at the philosophy program have been invaluable on those not-so-rare occasions when I had no idea what I was doing. Doug Lackey and Nickolas Pappas were great bosses in Baruch College and at CCNY respectively. Both fostered great environments in which to teach and do philosophy. I would also like to thank my friends here at the Graduate Center especially Hung-Tzu Yhan for the thousands of cups of coffee. I especially owe much to Leib and Sarah Litman who have been good friends throughout.

Mark Steiner in Jerusalem was very helpful to me when I started, and his influence is reflected on these pages. I must however state here that nothing that follows is endorsed by him. I would also like to especially note Hanna Caine-Braunschwig who was particularly kind when I spent time at the Van Leer Jerusalem Institute in 1999-2000, when I first started thinking about this project.

Thanks also to the Mid-Western Philosophy of Mathematics Workshop

(MWPMW 8) at Notre Dame for their invitation to present an earlier version of chapter 2.3.1. and the University of Western Ontario where a draft of part of chapter 3.2 was presented.

Acharon acharon chaviv; my family has been quite supportive and in an important sense they are responsible for my being able to complete this. So thanks to my parents, and siblings Esty, Shloimy, Malky, and their families for their encouragement. Finally, to Dahlia, who put up with this dissertation long enough through our courtship, I am grateful for your patience.

Table of Contents

List of Figures

אין שמחה כהתרת הספיקות[1]

---

[1] Traditional Hebrew proverb.

Chapter 1

Introduction

> *When certain inconvenient matters are*
> *brought to a mathematician, he can always say,*
> *"These are psychological [or physical, or historical]*
> *matters. I do not have to deal with them."*[1]

1.1 Thesis

Mark Steiner argues that "mathematical explanation exists" ([Ste78a]: 135). The thesis of this dissertation is that this is false; there is nothing that answers to the description "mathematical explanation", at least not in any of the various familiar senses in which philosophers and most scientists have come to use the term "explanation" (or at least ought to use it when they wish to be precise). Different ways of using "explanation" are discussed in this chapter.

It is significant that there is no mathematical explanation. Why? Because many philosophers (e.g., [Gia05]) take it for granted that we can truthfully

---

[1] Stephen Pepper, *World Hypotheses*: 1.

talk about mathematical explanation, and they even point out "examples" of mathematical explanation. In a discussion of paradoxes Graham Priest ([Pri02]: 288) does so, and in passing endorses the existence of explanation in mathematics. Stathis Psillos ([Psi02]: 2) writes "...there are non-causal explanations (most typically, mathematical explanations) ..."[2] Mary Leng, after a very brief critical review of some of the literature on mathematical explanation, allows the presupposition that some mathematical reasoning is explanatory ([Len05]: 170-171).

Furthermore, some philosophers who take for granted that there are explanations in mathematics, exploit these "explanations" for a variety of purposes. Penelope Maddy ([Mad90]: 82, 86-87), for example, uses the existence of mathematical explanation as partial support for an argument in favor of realism in mathematics. Samuel John Butchart ([But01]) uses the existence of mathematical explanation to shed light on the nature of evidence in mathematics. Roughly, he argues that given the similarity between mathematics and science in that they both make use of evidence (and explanation), naturalism is the correct philosophical stance toward mathematics.

There are at least two reasons why understanding mathematical explanation is useful to philosophers: (1) It helps us better understand explanation. Knowing where there are explanations and knowing where there are not helps

---

[2] Explanation in mathematics is often taken for granted by philosophers in other fields. See for example the discussion of Democritus and Anaxagoras in Catherine Osbourne's [Osb04]: 77.

us understand what explanation is and how to go about a conceptual analysis of explanation in general. W. H. Newton-Smith ([NS00]: 132) claims that the lack of a decided account of explanation in the philosophy of science is an embarrassment to the field. Here I endeavor to show that it is not an embarrassment to the philosophy of mathematics as well. (2) Understanding mathematical explanation helps us understand mathematics. We are better able to delimit the benefits that philosophers of mathematics get from granting that there are explanations in mathematics. If there are explanations in mathematics then perhaps we can better understand mathematics' relationship with other fields that have explanations. But if, as we argue here, there are not, then we have less warrant for a naturalized mathematics because we have a significant disanalogy between mathematical and scientific methodologies.

Philip Kitcher and William Aspray ([KA88]: 17, and cited approvingly by David Corfield in [Cor03]: 18) claim that the issue of mathematical explanation involves questions which are asked only by maverick mathematicians and philosophers of mathematics. Corfield implies that the question of explanation in mathematics is a question about the very nature of mathematical practice, though he does not elaborate on what explanation is or in what sense explanations are countenanced by practicing (maverick) mathematicians.

To date, unlike in the case of the physical sciences, there are few theories of what mathematical explanation is and few examples of mathematical explanation. Even among proponents of explanation in mathematics there is nothing resembling a consensus on whether there exist mathematical explanations, as opposed to merely explanatory proofs. Thus far intuitions diverge even on this point. This dissertation will show that not only is there no adequate notion of mathematical explanation at present, but any account of explanation worth having cannot accommodate mathematics. We show this by first assessing all the current cases (and some hypothetical ones) in favor of mathematical explanation and demonstrating that they are faulty. Then we show that mathematics leaves no room for explanations.

Thus the pattern of the first part of our main argument will be an exhaustion of cases. Our general strategy will be to look at the main theories of explanation and deliver a series of negative results - i.e., for each theory of explanation we show that it is not simultaneously viable and applicable to mathematics. The assumption behind Chapter Two is that when we consider the large variety of theories of explanation, if there is such a thing as mathematical explanation, there is a reasonable expectation that it will fall under one of the existing accounts of scientific explanation. For each account of explanation we assume that it is adequate as it stands as a model of scientific explanation. We attempt to determine whether the model is still adequate when we assume it can cover cases in mathematics as well. We will show that the account will simply not hold in the mathematical cases for each account of scientific explanation.

After examining the main plausible accounts of scientific explanation in Chapter Two, we spend Chapter Three reviewing the proposed account of explanation in mathematics that does not have counterparts in theories of scientific explanation. We also examine a number of suggestions about the nature of mathematical explanations. We find that this model and these suggestions are both inadequate.

In Chapter Four we show, in a series of positive arguments, that because of the nature of explanations and the nature of mathematics, there are insurmountable conceptual difficulties involved in claiming that there are explanations in mathematics. We examine various expectations we have of a theory of explanation and show that mathematics cannot satisfy them.

The remainder of this chapter covers important preliminaries. First, §1.2 gives some general background to the philosophical question of scientific explanation, and situates the problem of mathematical explanation in the philosophy of science and mathematics. Second, in §1.3, we limit the scope of our investigation to precisely those cases where mathematical explanations are offered within mathematics and not in other domains. In §1.4 we address the question of whether we should assume that there must be a unified account of explanation for mathematics and science. §1.5 spells out an important distinction that will be relevant throughout our study: the distinction between notions of explanation that are psychological, clarificatory, or subjective, that plays to the stylistic preferences of the particular mathematician, and a notion of explanation that is objective and accounts for the “phenomena”. Significant parts of our argument rely on this distinction holding as we are only interested in the latter.

1.2 Background

The general background for the question of mathematical explanation lies in the philosophy of science[3] and its history. Scientists have always seen it as one of their principle duties to explain the phenomena. Aristotle’s discussion of the four causes is the first known attempt to precisely articulate what a scientific explanation consists of. To fully explain something, Aristotle claimed, we must analyze what it is made of, what shape it is, what process made it, and what it is made for. Modern scientists too seek to offer explanations, not exactly the way Aristotle did, but in many respects not all that differently. To explain something is to somehow account for its existence or presence. Scientists, for example, account for the phenomena of earthquakes, that is, they explain their occurrence, by appealing to plate tectonics, the theory that describes the nature of the Earth’s plates that move against each other in particular ways.[4]

---

[3] See [Sal89]: §0.1 for the history of scientific explanation through 1989.

[4] Russell Kahl’s [Kah63] is an interesting collection of 27 explanations. These explanations are culled from a range of disciplines to provide a clear set of examples for the students of philosophy of science to use; selections are taken from many areas from Ovid’s explanation of the creation of the cosmos to Weber’s explanation of the social causes of the decay of ancient civilization. There are examples from physics, biology, psychology, sociol-

More recently the question of scientific explanation was reinvigorated by the Logical Positivists. Logical Positivism[5] thrived mainly during the period between the two world wars, and began to seriously wane around the early 1960s. During the positivist era, there was a growing discomfort with the extant philosophical outlook which had been dominated by the Fichte, Schelling, post-Kantian, and neo-Hegelian traditions. Scientific and philosophical explanations were laden with transcendental metaphysics, theology, teleology, and entelechies. Positivists believed that the primary task of philosophy was to describe the scientific enterprise, and discard all the unnecessary metaphysics that was (and sometimes still is) identified with philosophy in general. Once that was done, philosophers would have contributed all they can to advance human knowledge and a period of real knowledge acquisition, i.e., the accumulation of scientific information, could then begin without intractable metaphysics haunting them.

Once philosophers started to see scientific knowledge as the only real knowledge, and Frege's new logic - which promised to become the new scientific language - became more widely known, there were few remaining philosophical tasks; foremost was to describe precisely what science is, what scientists do, and what the underlying logic of science is. The project went as

---

ogy, and history, though there is no mention of mathematical explanation, or explanation in mathematics. similarly in David-Hillel Ruben's discussion of non-causal explanations, examples from quantum chemistry, formal linguistics, physics, and genetics are invoked, but no references are made to mathematical explanations.

[5] [Aye36] is a good exposition of the methodology and main doctrines of Logical Positivism, written by one of its most ardent proponents.

follows: enumerate the various scientific tasks, and then ascertain that there is a clear "conceptual analysis" of each one. What are the tasks of science? Among the things that scientists, qua scientists, do are make predictions, discover laws of nature, formulate theories, assert probabilities, propose hypotheses, set taxonomies, and offer explanations.

But what is a law of nature? What is a theory? Are there better and worse methods of structuring taxonomies? Can testing a hypothesis prove a theory? And what is an explanation? We constantly encounter questions about the nature of scientific theories. Evolutionary biology, for example, is dismissed by its detractors as "merely a theory". Scientists themselves on the other hand are very impressed at having something as well worked out as a theory of the origins of life on Earth. But neither evolutionary biologists nor their detractors would have an easy time telling even an intelligent audience exactly what they mean when they use the word "theory". Is a theory an educated guess? Is it something we are only somewhat confident of? Is it a way of putting together the data? What sort of things are included in theories? What is the point of a theory? How does a theory differ from a recitation of the known facts?

Similar questions can be asked of the term "explanation". Offering a scientific explanation of a phenomenon is a major accomplishment for a scientist or for science. But just what counts as a scientific explanation is a matter of contention. While believing the same set of facts, reasonable people disagree on whether evolution "explains life", given differing conceptions of "explanation".

Immediately proceeding the positivist era, some nineteenth century thinkers including Gustav Kirchnov[6], Ernst Mach[7], and Auguste Comte (following Hume[8]) considered explanations as part of the old metaphysics-laden approach to scientific explanation.[9] For them it would have been a significant achievement to show how science could be formulated without referring to explanations at all. Philosophers like Emile Meyerson argued otherwise, and claimed that explanations were an indispensable part of the practice of science. In the end the positivists saw it as one of their major accomplishments that they came up with a non-metaphysical analysis of scientific explanation that was true to their ontological views and consistent with scientific practice. Carl Hempel ([HO48]) is credited with formulating the Deductive-Nomological (D-N) model of explanation. This became the canonical positivist account of metaphysics-free explanation.[10] For Hempel, explanations meet specific criteria, and can be recognized solely by their logical form. Hempel's model of explanation is clear about what counts as a scientific ex-

---

[6] See [Car66]: 12-17.

[7] To the extent that Mach had a view of explanation, it was purely psychological, not metaphysical. See [Mac93]: 7. We address this distinction below.

[8] See Ch VII of [Hum77] "The idea of a necessary connexion".

[9] Pierre Duhem was also an opponent of explanation in this period, though his motivations were different from those of Mach's or Comte's. See the second appendix to [Duh54].

[10] Rudolf Carnap ([Car66]), Richard Braithwaite ([Bra53]), Ernest Nagel ([Nag61]), and Karl Popper ([Pop59]) all promoted related views.

planation, and is clear about what fails to be an explanation - or when an explanation fails.

It was becoming clear by the 1970s that the positivist model of explanation would not pan out. There were many technical problems: many things that were clearly taken to be scientific explanations did not fit the model, and there were many things that fit the model that were not seen as scientific explanations. As people began to study the D-N model more closely, patches were added to it and alternative models were proposed.[11] A large body of literature grew around the question of what counts as a scientific explanation. Today there are at least four ways of analyzing "explanation" that are taken seriously by many philosophers. We discuss each of these in turn.[12]

The project of this dissertation starts when we ask the question: is everything we say about scientific explanation applicable to mathematics as well? Are there analogs to scientific explanation in mathematics? Do we give explanations in mathematics in the same way we give explanations in science? At first glance this sounds absurd. Science offers explanations, not mathematics! However, there are a number of reasons to think that it does make sense to speak of mathematical explanation. The first reason is that historically, often

---

[11] See Chapter 2 of [Sal89] for a good discussion of the D-N model and its problems.

[12] Actually, there are many more theories of explanation. Most however are so far from being applicable to our concerns that there is no plausible way to include them in our discussion of mathematical explanations. Those models include Humphreys' aleatory model, Fetzer's probabilistic model, Railton's D-N-P model of probabilistic explanation, and Hempel's I-S (Inductive-Statistical) model.

when someone spelled out what an explanation is in science they provided an example from mathematics without any indication that there might be a difference between the two. Also, there are many instances in the history of mathematics where mathematicians exhibited preferences for one type of proof or method over another. Some philosophers argue that these claims are expressions of preference for explanatory mathematics over non-explanatory mathematics. Next, we tend to use the word "explanation" to describe a variety of things that we do. So perhaps explanations are also in play in the course of elucidating mathematics. Finally, there are many philosophers who for various reasons do not see a significant difference between mathematical and scientific methodologies or mathematical and scientific ontologies. This group includes some mathematical realists and some mathematical naturalists. They argue that if their respective methodologies and ontologies are the same, then by analogy, they should share the fact that explanations exist in their respective domains.

The present work investigates whether mathematics has explanations. When mathematicians discuss mathematics are they looking for explanations or are they just looking for proofs? Do [some] proofs explain? Mathematics deals with proofs, constructions, and deductive inferences. However, can we say that in the same way that science offers us explanations of objects, events, and phenomena in the world, mathematics also offers us explanations, or does mathematics only give us proofs, theorems, and the like? Does mathematics provide us with something identical or analogous to that particular kind of account that in science would be called "explanation"? In other words, do we have that ability to account for some mathematical "phenomena" (or the mathematical analogue to phenomena), or do we just get end results and proofs in mathematics? That is the question we address in this dissertation.

Contrary to most of the contemporary philosophers who have dealt with this question, our conclusion is that mathematics only gives us final answers and proofs. When you have a mathematical proof, that is all you have. There are none of those special things called "explanations" which account for mathematical phenomena. Anything called "explanation" in mathematics is an explanation of a very different sort than the kind philosophers of science have been explicating since the presocratics. We will see what sorts of explanations those are in §1.5.

Like many of the significant enduring questions in philosophy, we can see the first inquiries into mathematical explanation in Plato's writings. In the *Phaedo* we find Socrates interrogating Cebes regarding the way we account for number:

> I will not even allow myself to say that where one is added to one either the one to which is added or the one that is added becomes two, or that the one added or the one to which it is added become two because of the addition of the one to the other. I wonder that, when each of them is separate from the other, each of them is one, nor are they then two, but that, when they come near to one another, this is the cause of their becoming two, the

> coming together and being placed closer to one another. Nor can I any longer be persuaded that when one thing is divided, this division is the cause of its becoming two, for just now the cause of becoming two was the opposite. At that time it was their coming close together and one was added to the other, but now it is because one is taken and separated from the other. (96e-97b)

Socrates appeals to the Forms to account for what makes two two, and three three, etc. Further on Socrates continues:

> . . . would you not avoid saying that when one is added to one it is the addition and when it is divided it is the division that is the cause of two? And you would loudly exclaim that you do not know how else each thing can come except by sharing in the particular reality in which it shares, and in these cases you do not know of any other cause of becoming two except by sharing in Twoness, and that the things that are to be two must share in this, as that which is to be one must share in Oneness, and you would dismiss these additions and divisions and other such subtleties, and leave them to those wiser than yourself to answer. (101b-101d)

When looking for what accounts for the properties of some numbers, i.e., the odd ones, we find Socrates (103e-105d) trying to demonstrate that it is Oneness that makes a number odd, since although two odd numbers, when added together, must be even, nothing, even Oneness can produce its opposite, i.e., something even.

Plato thus alludes to questions about mathematical explanation in mathematical examples about the nature of the Forms, as the forms are often taken to be explanations.[13] Plato's general attitude was undoubtedly spawned by the Pythagorean doctrine of the One, or Unity, being the cause of everything. One was said to be the cause of both odd and even numbers (which in turn was the cause of everything else).[14] This attitude persisted through the middle ages. Abraham ibn Ezra, a twelfth century philosopher and an author of works on arithmetic, in his religious work *The Secret of the Torah*, in a somewhat Husserlian moment, follows the Pythagorean doctrine quite explicitly in claiming that "One is the cause of all numbers. However, it is not itself a number ..." ([Ezr95]: 159).[15]

So the question of mathematical explanation is at least as old as Plato or Pythagoras. In modern times Nicholas Rescher has asserted that mathematics has explanations:

> ... "what sorts of things can be explained - what is the potential range of explanatory problems?" The answer is "Any and all facts whatsoever." ... The properties and states of virtually anything, any and all occurrences and events, the behavior and doings of people, indeed every aspect of "what goes on in the world," can be regarded as appropriate objects of explanation. And here "the world" can be taken in the widest sense, including not only the physical universe, but also the "world of mathematics" and so on. ([Res70]: 3-4)

---

[13] See [Rub90] for more on Plato's notion of explanation.

[14] See Aristotle's *Metaphysics* 986a 14-20, and the fragment of Theo of Smyrna in [Whe60] page 224.

[15] This very Pythagorean attitude is still sometimes seen in contemporary scientific literature (see e.g., [Hen05]).

We discuss the history of mathematical explanation throughout this work. Some philosophers, particularly Paolo Mancosu, see the issue of mathematical explanation as having a long continuous history pervasive throughout the history of mathematics. I shall argue that concern over mathematical explanation seems to only appear in the history of mathematics sporadically, if at all.

But before we can go on to explore the questions of explanation, we still require a few distinctions.

1.3 Explanations involving mathematics

There are three ways explanation can be involved in mathematics.[16] (1) First, mathematics or some part of it may provide explanations for some non-mathematical facts or phenomena. Mathematics has, for example, been invoked in the explanation of various physical facts. Pythagoras is the first known thinker to invoke mathematics, specifically Number, to explain the physical universe. Plato would later adapt Pythagoras' style of explanation and use the Forms instead of Number to explain the natural world.

In a 1638 letter to Mersenne, Descartes writes that he is

> resolved to quit only abstract geometry, that is, the study of questions that serve only to exercise the mind; and I do so to have at much more leisure to cultivate another sort of mathematics that takes on as questions *the explanation of the phenomena of*

---

[16] See Mancosu's [Man01]: §I.2, Hafner and Mancosu's [HM05], and [Man08b].

> *nature*. For if he will please consider what [I] have written on salt, on snow, on the rainbow, etc., he will understand that my entire physics is nothing other than mathematics. (Quoted in [Mah90]: 462, *emphasis added*.)

Since Galileo mathematics has become more and more significant in any discussion of science. There are now numerous examples. To pick one at random: the Fourier series can be invoked to account for the similarity of the wave equation and the diffusion equation.[17] In a 1979 interview, Richard Feynman declared that we are well past the time where any real science can be done without the tools of mathematics.[18]

Indespensibility arguments which stress science's dependence on quantification over mathematical entities alongside physical and theoretical entities, begin with the assumption that mathematical entities are indespensible for accounting for scientific phenomena.[19] The most common usage of the phrase "mathematical explanation" is, as Leo Apostel ([Apo61]: 14) claims, to give a mathematical model for physical phenomena. Eugene Wigner's question about the applicability of mathematics to natural science is an important part of this subject.

---

[17] For more contemporary discussions of mathematics explaining science see e.g., Grahm Nerlich's [Ner79] where he claims that geometry explains features of the shape of space. John Barrow ([Bar04]) explores Gödelian incompleteness as it accounts for various physical results. Mark Steiner's [Ste89], [Ste95], and [Ste98], which builds on Eugene Wigner's [Wig60], can be seen in a similar light. (Ronald E. Mickens' [Mic90] is a good collection of articles on this topic.)

[18] See [Fey99]: 193-194.

[19] See Mark Colyvan's [Col01] for recent work on indespensibility arguments.

That is the first way that we can speak of mathematical explanations and I believe the most common.[20]

(2) Second, is mathematics, or some part of it explained by facts, objects, or structures that are outside the domain of mathematics? The idea that it is would undoubtedly find sympathy with many who take the continuity of mathematics and science seriously. This would also strike a Millian[21] or other mathematical empiricist as plausible.[22] Heidegger, who also raises a question of mathematical explanation sees mathematics as something that is known together with the things in the world ([Hei77]: 273-278). More recently some have claimed that it has become fashionable in some mathematical circles to invoke empirical data and empirical methodological techniques in mathematical proofs (see [Hor93]). (Much of the early history of mathematics is also intimately connected with physical techniques and results.) Thus we often have non-mathematical information involved in explaining the mathematical.[23] Chris Swoyer ([Swo99]: §5) argues that many aspects of number theory are often explained by appealing to metaphysical entities, i.e., properties. So

---

[20] For a good discussion of mathematical explanation in the physical sciences see Mary Leng's [Len05]. See also Lyon and Colyvan's [LC08] for a recent discussion.

[21] See [Mil41]: Bk 3, ch. 24 §5, esp. p. 400-1.

[22] We discuss Mill's work in a related vein in the next chapter. Mancosu specifically invokes Mill (in [Man01]: 103-4) as advocating what Mancosu calls the "h-inductivist tradition" which he claims is a tradition of mathematical explanation.

[23] Interestingly enough, we sometimes find mathematicians offering a physical model (e.g., a circuit diagram) of a mathematical result to illustrate a point about mathematics. These models can claim to explain the mathematical structure they model. However, they generally use very mathematically idealized physics, such as assuming that some of the physical objects have zero volume, or are frictionless, etc. So they do not literally conform to the mathematics, they merely give us an intuitive but imprecise analogy.

we have a second way of invoking explanation in mathematics.

(3) Third and finally, we can inquire whether there are mathematical facts or structures that explain other mathematical facts or structures.

Our present discussion will focus only on the last of these three where both the explanans and explanandum are mathematical truths, structures, proofs, facts, etc. If there are philosophical concerns that emerge from (1) and (2), and I suspect that there are, they will need to be dealt with very differently from the way we will be dealing with (3). One reason for this is the nature of mathematical objects and mathematical facts, and because in this case both the explanans and explanandum are clearly on the same ontological and methodological footing - they are both in the domain of mathematics, and thus will not need an argument to defend the (counter-intuitive) notion that they are the same or that there is an (explanatory) relationship between them. The first two of our cases would first seem to need to show how objects in one domain interact with and account for those in mathematics. That is far afield from our present concerns. So while some philosophers see the first and third kinds of explanation as intimately related, and sometimes run them together, we do not.

1.4 Methodological dualism?

Another question which should be addressed prior to our inquiry is the question of whether or not to insist on a methodologically monistic view of science and mathematics. Should we consider accounts of explanation that are exclusively designed for mathematics (such as Steiner's) on the grounds that it might give us a good account of what mathematical explanation is, regardless of whether or not it can inform us on the question of scientific explanation.[24] Or do we see it as a defect of any account of explanation if it cannot handle both science and mathematics? Shall we dismiss an account of explanation out of hand if it cannot accommodate both mathematical and scientific cases?

This is a relevant question here since we will be invoking explanation - a concept that has been traditionally located inside the domain of scientific methodology - and we are looking at explanation from the perspective of mathematics.

Jamie Tappenden sees it as a virtue in a theory of explanation if it can accommodate both science and mathematics ([Tap05]: 158). Mancosu ([Man01]: 102) claims that since "theories of scientific explanation often attempt to characterize the 'scientific' aspect of an explanation independently of the subject area in which it might be offered", "they should thus be able to capture mathematical explanations". Their failure to do so shows a serious limitation of their theories. But this begs the question of the continuity and similarities of mathematics and the natural and social sciences. Mancosu as much as admits this when he claims that if we cannot find a common theory

---

[24] See Sandborg's [San97]: 2-3 for more on this question. Mancosu (in [Man01] and with Hafner in [HM05]) also addresses this. Questions about methodological dualism can be found in [KS89]. We shall also briefly return to this point in connection with Sandborg's [San98].

of explanation for both it “will reveal a very interesting difference between the two domains”. But clearly mathematics and science exhibit many structural and methodological (if not ontological) differences; e.g., observations, theories, contingincy, and laws are just some of the features of science that it does not share with mathematics. Why should the inclusion of “explanation” in this list be particularly interesting? Mancosu does not elaborate. Nonetheless he insists that “it is thus clear that mathematical explanations can be used to test theories of scientific explanation”.

Some philosophers like J.J.C. Smart ([Sma90]: 2) assert that all explanations take the same pattern. Other parsimonious philosophers like Philip Kitcher and Jamie Tappenden, claim that we should have only one theory. The grounds for this claim is the alleged continuity of mathematics with science. If mathematics and science are continuous, there is no justificatrion for, or need of more than one methdology.

Insofar as we are concerned, this continuity is a poor reason to advocate a unified methodology. From the perspective of our concern, it begs the question in favor of there being explanation in mathematics. Also it does not strictly follow that because the sciences are continuous with mathematics, that they ought to have a unified philosophical underpinning. There can be a unified body of knowledge within which there are diverse methodologies of explanation. We will take it as undecided whether or not mathematics is continuous with science, and certainly whether it ought to be treated identically from a methodological standpoint. There does seem to be a relevant difference between mathematics and science that suggests that the methodology ought not be unified. What will apply to mathematics need not apply to science. We treat this more extensively in Chapter Four.

If we do not need a unified account of explanation for mathematics and science we then have license to explore both theories of scientific explanation and see if they can fit into the domain of mathematics and also theories of explanation designed specifically for mathematics. We thereby do not need to worry that we would prefer an account that works in the sciences as well.

1.5 Different types of explanation

In the first paragraph of this chapter we qualified the statement that there is no such thing as mathematical explanation. Here is why: It has been noted[25] that in ordinary English (and many other languages) we use “explanation” in a variety of ways. For example, we explain how to pitch a tent, explain why we lied, we explain *Macbeth*, explain how to fly an airplane, or explain what “melancholy” means. In these cases “explain” means to describe, justify, paraphrase, demonstrate, and define, respectively. We may encounter empathetic explanations, ideological explanations, (see e.g., [Net70]), philosophical explanations, historical explanations, etc. Therefore, the claim that there is *no* explanation in mathematics will not be the whole story.

---

[25] See also Mancosu’s [Man01]: 100.

There is an obvious first objection to the assertion that there is no mathematical explanation. The objection can take many forms. But it will generally invoke a common usage of “explanation” and claim that of course people explain mathematics. Many textbooks claim to explain their subject matter, or explain it better than another textbook. Many students and mathematicians have heard explanations of particular pieces of mathematics.

But this objection is predicated on a confusion, as it runs together various senses of “explanation”.[26] “Explain” is ambiguous and using the wrong sense of “explanation” will lead to philosophical confusion. This confusion frustrated Hempel. In [Hem63] (126-127), as a response to an objection by Michael Scriven, Hempel wrote that there can be no deductive model of explanation that covers explaining “to the mechanic in a Yugoslav garage what has gone wrong with a car.” Hempel then goes on to point out that this confusion is similar to objecting to some definition of “proof” (in the metamathematical proof-theory sense) because it fails to accommodate how “proof” is used in “90 proof gin”.

Standard philosophical models of explanation were never intended to cover all uses of the word “explanation”. They were designed to capture answers to questions that ask *why* something happened or, more precisely, *what accounts* for some phenomena, pattern, event, or regularity.

Some philosophers have looked to mathematical literature and noticed

---

[26] This is the kind of confusion that Francis Bacon might call an “Idol of the Marketplace”.

the occasional use of the term “explain”. They then cite this as evidence that mathematicians are interested in explanations, taking the word “explanation” as it is used in general philosophical discussions of scientific explanation. However, the philosophers’ use of “explanation” in these cases does not match the mathematicians’. In addition, philosophers may look at mathematical literature and notice that mathematicians are (or appear to be) pursing a goal that extends beyond merely proving theorems. Philosophers then take this as tantamount to the claim that mathematicians are involved in a project of offering explanations. Further on we examine other goals mathematicians may be pursuing. They are not seeking explanations.

Thus as a preliminary to our project, our first task is to show that in mathematical contexts, “explanation” is used, but in ways other than the philosophically interesting ones.

To stave off charges of our project being question-begging *ab initio*, we do not merely assert that in mathematical contexts “explain” means something other than what philosophers of science intend. We show that whatever “explanation” is in science, it fails to be that in mathematics. We argue that the only thing “explanation” can be in mathematics is something that reflects various subjective preferences and experiences of individual mathematicians.

Another version of the above objection to the claim that there are no mathematical explanations appeals to the difference between proof and explanation. That objection claims that in mathematics only the proofs need to be objective. The explanations in mathematics are subjective. Thus there are ways to explain mathematics, and the ways must be subjective.

I am sympathetic with the sentiment behind this objection. That objection highlights the underlying theme of this work. We claim that to the extent that explanations in mathematics are taken to be expositions that appeal to some people in some circumstances, then they exist in mathematics. However, the literature in the philosophy of science makes no place for such explanations in science. The literature in the philosophy of mathematics implicitly assumes that explanations in mathematics serve a similar function in mathematics as scientific explanations serve in science. If it did not then the question of explanation in mathematics is trivial.

Few philosophers claim that in scientific discourse a “result”, “observation”, or “experiment” is objective and an explanation is subjective. One assumption behind much of the literature in scientific explanation is that whatever explanations turn out to be, they are objective, or at least can be objectively determined. Moreover, if it was the case that mathematical explanations are subjective, quests for models or cases of mathematical explanation would be pointless. Any exposition or popularization of mathematics would be sufficient evidence that there are explanations in mathematics, and that exposition would illustrate what explanations are. But there is general agreement that at the very least, if there are explanations in mathematics, they are more than popularizations.

The claim should be easier to accept when we realize that scientists will often describe their trade as essentially an explanation-seeking enterprise. Mathematicians make no similar claims. Therefore, to the extent that there are explanations in mathematics - and there are certainly *some sorts* of explanations offered in mathematics - they will be of a far different sort than those in science.

In the remainder of this section we rule out four (somewhat related) uses of "explanation" as being irrelevant for our task of finding explanations in mathematics. In future chapters we show that when philosophers claim that there is explanation in mathematics they are using "explanation" in one of the following ways. And given that these uses are very different from the use we seek, there is no explanation in mathematics in a sense that will be methodologically useful to mathematics. So I am arguing throughout this work that the philosophers who do believe that there are mathematical explanations, really allow "explanation" to take on one of the following meanings. However, these meanings bear little resemblance to the meaning of "explanation" as it is traditionally seen in the literature of the philosophy of science.

(1) "Explanation" is sometimes used in what might be called its "pedagogical" or "clarificatory" sense, as opposed to the sense of explanation that we use when we want to talk about accounting for phenomenon.[27] Isaac

---

[27] J. D. Trout uses the term "pedagogical" in reference to explanation, specifically to Kitcher's unification account. See his [Tro02]: 221. Humphreys ([Hum93]) ties pedagogy to

Asimov, the great science and science-fiction writer earned a reputation as "the great explainer" of science. This is because he helped disseminate, popularize, and clarify scientific concepts. It is also despite the fact that he did little scientific research in his life, either theoretical or applied, and was generally not involved in accounting for previously unaccounted-for scientific phenomena. Asimov was mostly involved in purely pedagogical projects. In the pedagogical sense of "explanation" one makes a topic clear in a *psychologically* useful way[28] to the individual it is being explained to. It is in that "popularizing" sense that Asimov earned his moniker. In a case where we seek to elucidate science in a psychologically useful way, we speak of explainers and explainees,[29] not explanans or explananda. Asimov was an explainer of science to explainees. He did not necessarily provide sets of explanans and explananda in any rigorous way. To the extent that Asimov might have presented sets of explanans and explananda to popular audiences, he earned his title not in virtue of that, but in virtue of his ability to present and describe those explanans and explananda in a way that was comprehensible to a non-scientific audience.

This distinction is often made explicit. Ruben Abel lists seven uses of the

---

understanding and then understanding to explanation. We agree with the former pairing, though not the latter for reasons we explain shortly. Mancosu's [Man01] also addresses this (see the references on page 101, and later on) when he criticizes Hersh for conflating "explanation" with "pedagogy". David Hillel-Ruben ([Rub90]) however takes pedagogy to be the difference between the epistemic and the ontological sense of explanation.

[28] See [Hem63]: 127.

[29] We deliberately avoid Bromberger's tutor/tutee terminology here to differentiate between the agents in the different types of explanations.

term "explanation", many similar to those mentioned above. He clarifies:

> There are of course many kinds of explanation; they are the answers to many kinds of questions. …
>
> Why is there no greatest prime number? Why is the sum of the interior angles of Euclidean triangle 180°? Here the explanation would be the *analysis* of what is logically entailed by certain postulates of logic and mathematics.
>
> …All of these explanations contain a core of what Mill calls "considerations for the intellect to give its assent." The explanation releases the tension that provoked the question. It evokes the "aha!" response: oh, so Brutus thought Cesar wanted to be emperor! So the Trojan horse was full of soldiers! so Miss Prism was the nurse! so there was sabotage on the plane! If the puzzlement is not in fact eliminated, the explanation is not accepted. …This psychological aspect of explanation varies with the person, and with his degree of bewilderment and sophistication; however, the explanations which science provides, by means of logical inference from a law, are independent of persons. ([Abe76]: 91, emphasis in original.)

Abel seems to use "analysis" to mean "proof". This is contrasted with explanations that merely remove bewilderment. Hersh too ([Her93]: 398) is clear that the "educational value of proof is the value of complete explanation. The teacher decided whether *any* explanation is called for."[30] Wesley Salmon also expresses the notion that pedagogical explanations are not related to philosophical explanation:

---

[30] See also [Man01]: 107.

> In everyday life we enounter many types of explanation, which appear not to raise any philosophical difficulties .... Prior to take-off a flight attendant explains how to use the safety equipment on an aeroplane, in a museum a guide explains the significance of a famous painting. A mathematics teacher explains a geometric proof to a bewildered student. ([Sal92]: 132)

We can think of it as follows: A work of popular scientific writing attempts to explain something to the student or lay reader. The work does this by taking existing explanans and the explanandum that scientists have provided for them and then rephrasing it in a way, or adding some detail that makes it easier for some explainee to understand, grasp, or comprehend. Thus, a mathematics explainer might use easier mathematics[31], mathematics that is learned at an earlier age in a standard school curricula, or mathematics that uses shorter chains of reasoning.[32]

The scientific project, in contrast, has always been to take existing explanandum and provide appropriate explanans. The explanans are appropriate in this case if they conform to certain formal scientific and philosophical considerations, like concision, parsimony, symmetry, consistency with theory, etc. Paraphrases and explainees play no part in this latter type of explanation.

As Alex Rosenberg notes:

---

[31] See [Tap05]: 179.

[32] "Mathematical explanation" is used often in the context of various kinds of mathematical pedagogy: see e.g., [JBM99], [EAFD97], and [LSL02].

> these [logical empiricists] sought a notion of scientific explanation that would constitute an objective relationship between *explanandum* and *explanans*, a relation like the relation of mathematical proof, which obtains regardless of whether anyone recognizes that it does . . . the logical empiricists rejected the notion of scientific explanation as an attempt to allay curiosity or answer a question that might be put by an inquirer. It is relatively easy to "explain" complex physical processes to children by telling them stories that allay their curiosity. The subjective psychological relevance of the *explanans* to the *explananda* in such cases may be very great, but they do not constitute scientific explanation. . . . they needed a relationship which made explanatory relevance a matter of objective relations between statements and not the subjective beliefs about relevance of less-than-omniscient cognitive agents. ([Ros00]: 26-27)

Here we consider whether there can be analogous cases in mathematics. It is uncontroversial that we find analogues to the first, pedagogical-type of explanation. There are mathematical results, proofs, and theories. There are also many popular works and pedagogues which explain these mathematical results, proofs and theories by rephrasing it for the student, lay reader, or other explainee. But what we shall see is that the latter kind of explanation that is analogous to scientific explanation does not obtain in mathematics.

That is the first sense of explanation that is of purely psychological interest, not scientific/philosophical. We shall see that many philosophers conflate these pedagogical explanations with proper explanations.

(2) The second point we make about "explanation" is that some expla-

nations vary from person to person. Those too shall not interest us. We are looking for explanations that do not succeed or fail depending on the person receiving the explanation. Explanations that can be said to satisfy some individuals and not others are unsuitable for our purposes.

This type of explanation more closely resembles something that produces understanding. Understanding is traditionally seen as an important part of scientific explanation. It is usually invoked in any account of explanation.[33] Peter Achinstein believes that there is a "fundamental relationship between explanation and understanding" ([Ach83]: 16). David Lewis requires that "the recipient understand and believes what he is told" for an explanation to work ([Lew86]: 218). Michael Friedman (in [Fri74]: 77) claims that the central problem of scientific explanation is the relation of phenomena that "gives understanding of the explained phenomena". Kitcher too (in [Kit81a]: 168) claims that a theory of explanation should "show us *how* a scientific explanation advances our understanding". Finally, James Woodward (in [Woo93]: 249) claims that a theory of explanation should "identify the structural features of such explanation which function so as to produce understanding in an ordinary user".[34] Similarly Newton-Smith claims that "A good explanation increases our understanding."([NS00])

However, there are a number of reasons I reject equating explanation with understanding. First, in trying to analyze "explanation" few philoso-

---

[33] See Jaegwon Kim's [Kim94].

[34] See([Tro02]).

phers consider psychological repercussions of understanding in their accounts of explanation. There are many varieties of explanation that might "clear things up" or make something more comprehensible to an individual.[35] Each of these types of explanations would be situation and individual-dependent. Because of that, it is hard to have a non-psychological account of understanding. If there are no non-psychological accounts of understanding then to the extent that explanation is tied to understanding there can be no non-psychological accounts of explanation.

A not atypical article in a scientific journal may make reference to an explanation that can be provided given the author's results. Naturally, such an article will be comprehensible to, and only to, those who already understand the field(s) that the results are in. The mere presence of an explanation will not make anything understood. The article will provide an account of say, how some newly observed phenomenon accounts for previously unexplained but already known phenomena. The reader has an explanation where he didn't have one before, but is not guaranteed (any) new understanding. Nor does the reader necessarily come away with a sense that the subject is less complex, nor does he have a greater appreciation for the field. It is possible

---

[35] Abel ([Abe76]: 93), under the heading of "Misconceptions About How Science Explains" states that

> Scientific explanation is not the same as "understanding" in the sense in which it is said, for example, that D. H. Lawrence understood women, or T. E. Lawrence understood the Arab mentality, or that an experienced nurse understands children. Such understanding is more like knowledge by acquaintance or like knowing how . . . than like science.

that while a scientific article does provide an explanation, the reader may actually *understand less* than when she started.

Explanation and understanding are clearly not the same thing,[36] and a theory of explanation will not be identical to a theory of understanding. Any theory of explanation must not fail to distinguish between the two. Perhaps "understanding" relates to "mentally grasping" something. Perhaps it relates to something else. Both successful and unsuccessful explanations can lead to or increase understanding, and one can often understand something or understand something better without an explanation present.

Those who research physics or mathematics education often distinguish between students understanding material and students merely memorizing material. When researchers measure understanding they attempt to determine a number of things: does the student know how to apply formulas or concepts in new contexts?, does the student see how various concepts connect with each other, or grasp that two separate fields or topics are really one coherent topic? Getting students to understand in this sense is not part of the mathematician's or scientist's job, *qua* mathematician or scientist respectively.

When we look back to the positivists and other early theorists of scien-

---

[36] Explanation and understanding are often run together (e.g., [Tap05]). Sometimes this is deliberate, sometimes it is not. However, Kenneth Manders sees it as a goal to understand mathematical understanding independent of explanation and Jermey Avigad is conscious not to use 'explanation' interchangeably with 'understanding' in mathematical contexts. (See [Avi06].)

tific explanation we find that their aim was generally to make explanations objective. For example, Hempel's stated conception of explanation did "not require relativization with respect to questioning individuals any more than does the concept of mathematical proof." ([Hem65a]: 426)[37] The positivists, like us, assume that providing explanations is part of what physical scientists do when they do science. The positivists did not take scientists, *qua* scientists, to be providing a psychologically subjective pedagogical function. It is rarely (though sometimes) a useful contribution to science to rephrase a known result (though it may often be useful to some scientists). If it was generally useful, it would show that science was not completely objective as it is dependent on the particulars of some scientists.[38] Here too, we set aside any non-objective factors as we look for mathematical explanations.

Understanding is a psychological notion (formalizable or not[39]), or per-

---

[37] [Hem65a] §4 gives a number of other arguments for the same point having to do with the logical character of explanations. Paul Humphreys ([Hum89]: 127) expresses a similar sentiment when he explicitly ignores, in his discussion of explanation, the type of explanation that is "a relativization of explanations to an individual" as it tends to be concerned with "the epistemic state of an inquirer".

[38] See Hempel's [Hem65a]. Though note that he does at one point claim that increased understanding is a by-product of explanation (337).

[39] In outlining his "humanist" conception of mathematics (as opposed to the one he terms "absolutist") Ruben Hersh ([Her97]: 59-61) describes the differences between mathematical research and mathematical teaching. He writes:

> The absolutist teacher wants to tell only what he intends to prove... The main purpose of proof isn't explanation. The purpose is certification: admission into the catalogue of absolute truths.
>
> The view I favor is humanism...
>
> Proof is complete explanation. Give it when complete explanation is appropriate, rather than incomplete explanation or no explanation.
>
> ...some proofs don't explain much. They're called "tricky," "pulling a rabbit

haps something that explanation *leads to*.[40] But if some theory of explanation fails to guarantee a corresponding increase in understanding, all the worse for one's theory of understanding. On our account, a theory of explanation needs to meet certain objective [formal] constraints. Understanding is generally a phenomenologically familiar feeling ([Tro07]: 564). Philosophers may argue about what the formal constraints on explanation are, but the constraints must not give rise to a theory whose outcome allows for the success of an explanation to be dependent on who it is explained to, or how, why or in what context the explanation is offered or to a theory whose telltale sign of truth is as unreliable as as sense of understanding.

Given the above considerations, it is hard to imagine what is meant by, or what is gained by claiming that e.g., Aristotle's, Hempel's, Salmon's, Kitcher's, or Friedman's style of explanations increase understanding. In presenting them, each claims that greater understanding seems to fall out of their account naturally. That is, their respective accounts produce explanations that increase understanding. In each of the cases, "increased under-

---

> out of a hat." Give that kind of proof when you want your students to see a rabbit pulled out of a hat. But in general, give proofs that explain ...
>
> This policy uses the notion of "understanding," which isn't precise or likely to be made precise. Do we understand what it means "to understand"? No. Can we teach to foster understanding? Yes. We recognize understanding though we can't say precisely what it is ...
>
> *Proof can convince, and it can explain. In research, convincing is primary. In high-school or undergraduate class, explaining is primary.* (Hersh's *emphasis.*)

[40] This seems to have been Aristotle's view in Chapter 11 of the *Prior Analytics*.

standing" is left undefined. Ultimately we are left with a circular definition. Explanation (of some specific sort) is supposed to increase understanding, and understanding is what is gained by (some specific sort) of explanation. Few attempts to break out of this circle by analyzing understanding have even been made.

Additionally, the fact that each of the above authors believes that explanations that fit the pattern of their models increase understanding suggests that there must be very different views of what understanding is and that there are a number of ways to gain it. A possible reason for that is that "understanding" itself is a vague concept. As Philip Gasper puts it

> A satisfactory explanation of an event or phenomenon should provide us with understanding of what has been explained (the "explanandum"). But understanding is a notoriously vague and subjective state. Different inquirers may disagree about what is sufficient for understanding and whether or not understanding has actually been achieved. If the search for explanation has a central role in scientific reasoning, then it is important to ensure that our concept of explanation is free from this kind of vagueness. We want to be able to tell accurately when the goal of scientific explanation has been reached and when a new theory has sufficient explanatory resources to be accepted or preferred over its competitors. ([BGT91]: 288)

Moreover, J. D. Trout, gives a few reasons against using understanding as a criteria for judging explanations. He claims that understanding is just a type of psychological confidence brought about by psychological factors -

specifically the phenomena of hindsight and over-confidence, and they are likely to mislead ([Tro02]). As Trout concludes: "…there is little practical wisdom in relying on the sense of understanding as a cue of a good explanation, or even a potential one." ([Tro07]: 574) Thus the unreliability of our "understanding" is even more reason to divorce it from explanation.

Explanations are not only not subjective, but they are objective. If something serves as an explanation for someone it is also an explanation for someone else. There is near-universal agreement among biologists that the theory of evolution explains the various forms of life on Earth. We speak of this as *the* explanation, not *an* explanation. Those who do not believe in evolution as the explanation for life on Earth still speak of something else, say creation, as *the* explanation for life on Earth. Neither would seriously contend that there are two legitimate explanations - one for those who believe x, and one for people who believe y. The way we increase understanding about evolution may vary. An individual may come to understand evolution only after seeing an overabundance of examples of fossils, or reading popular books on the subject, while some may understand it after seeing evidence of similar DNA patterns in different species, and reading technical biology textbooks.[41] As

---

[41] A Story in Wired Magazine ("Be There and Be Square" Volume 14.06, 2006, p 114) has a mathematician, Will Murray, claiming not to understand Ron Graham's mathematics of juggling. With reference to the seminal "siteswap" paper, he is quoted as saying "I learned a lot of math from that paper, but in the end I didn't understand why the theorem was true. Others have come up with theorems that, to me, are more connected to juggling." But what does Murray mean? What would aid his understanding of the relevant theorems? It is unlikely that fitting whatever he is missing into any account of mathematical explanation will help him put the pieces together. [Gia05] apparently subsumes both of these uses

Dale Jacquette points out: “we explain the occurrence of disease differently to a three-year-old than to a thirty-three-year-old.” ([Jac06]: 80). But the sense of explanation we seek must be objective.

The take-away lesson here is that explanation, as something objective, is not intrinsically connected with explanations that help people understand. I aim to show that they are often conflated. Also, I stress that while I do not take the former kind of explanation to express something philosophically significant about mathematics, I do not say anything about the importance of the latter form of explanation.

(3) Our third distinction about explanation is that the context in which something is discovered (as opposed to the context in which it is justified) often provides what is referred to as an “explanation” of the theorem that is later rigorously proved. This view of explanation is rejected by Mancosu as well ([Man01]: 100). We occasionally find mathematicians adding information to theorems and proofs, especially in textbooks that allows the student to see how a proof can be “built up” or how it can be seen in the context of other related theorems. Sometimes this information shows how a theorem could have easily been discovered given the context it is put in. Sometimes the context is clear, other times the context of discovery can depend on the discoverer or on specific idiosyncratic intuitions. This precludes its relevance to us here in a discussion of mathematical explanation.

---

under “explanation”.

We should also note that the question of explanation should be clearly distinguished from the question of discovery. Many philosophers have discussed whether or not there could be a logic of scientific discovery. Some (e.g., Corfield) have also asked if there can be a logic of mathematical discovery. And while this is an interesting philosophical question, it does not speak to the question of the existence of mathematical explanations.

(4) Finally, the fourth use of "explanation" we do not find suitable is that which is motivated by cognitive science. The question which motivates these accounts is of the form: what explains or accounts for x's *belief* that α; or, why does a community of scientists (or mathematicians) adopt a particular set of beliefs. For example, we might inquire into what psychological and/or social forces motivated the mathematical community to accept Appel and Haken's computer proof of the Four Color Theorem or unsurveyable proofs in general. Or, perhaps we want to account for why fifteenth century Europeans did not have the concept of convergence to a limit, and hence not have calculus, differential geometry, or analysis despite the fact that mathematicians had been well aware of Eudoxus' and Archimedes' method of exhaustion to find the area of a circle for already nineteen centuries. We might then invoke Aristotle's dismissal of infinity as only potential infinity - i.e., infinity as merely an abstract fiction - as an explanation. Many questions of this sort have very interesting and important philosophical ramifications. Tappenden, for example, addresses the question of "*why* graphic statics held sway over

its analytical rivals during the time that it *did* hold sway." ([Tap05]: 178)

These sorts of problems are taken to have intimate connections with philosophical questions connected to mathematical practice.[42] However, they are mainly interested in explaining beliefs, not mathematical facts and are therefore not what we seek in our quest for mathematical explanation.

1.6 Conclusion

The more inclusive we become about what counts as an explanation, the further we get from any similarity with the standard philosophical project of explicating scientific explanation - and hence mathematical explanation.

Scientific explanation is important as a foil for us because in science, though there is little agreement as to what counts as an explanation, there is no significant debate on whether there are scientific explanations. Intuitions of philosophers and scientists currently affirm that it is reasonable to speak about explanations in science. But in the case of mathematics it is still at best an open question. Too few mathematicians or philosophers have argued over or expressed views on whether or not mathematics has explanations.

We are looking for a model of explanation in mathematics that is different from analyses of the kinds of explanations we just described. We seek a model of explanation that is (1) not pedagogical, (2) not dependent on "un-

---

[42] See also [Res70] for discussion of a pedagogical or 'pragmatic' notions of explanation. [Woo03] also discusses this as a preliminary to his discussion of scientific explanation. See [Tha99] ch. 1 and [Chu89] chs. 6 and 10 for what I refer to as cognitive science inspired accounts of explanation.

derstanding" or the particular psychological dispositions of the individual receiving the explanation, (3) is objective, (4) does not involve the context of discovery and (5) is not a model that accounts for a belief.

Given these constraints I argue in what follows that we will not find any model that will work for mathematics. Moreover, I contend that the philosophers who have seriously argued in favor of mathematical explanation are in fact arguing that there are examples in the mathematical literature that conform to the types of explanation mentioned in the last section; that is they violate at least one of (1)-(5) above. Let us now see how this is done with existing models of explanation.

Chapter 2

Models of scientific explanation

> *I'm not being deliberately gnomic, . . . , I assure you. There is no time for explanations now, and for some things, it seems, there may actually* be *no explanations.*[1]

In this chapter we look for theories of mathematical explanation in theories of scientific explanation. Since so much work has been done on theories of scientific explanation[2], we examine them to see if they can do double-duty as theories of explanation in mathematics. We ask of each theory of explanation whether it holds up as a theory of explanation for science and also mathematics. We wish to know if we have a theory of explanation that is preserved when the theory is taken to include mathematics.

---

[1] William Gibson, *All Tomorrow's Parties*: 126.

[2] The literature on scientific explanation is quite extensive. [Sal89] has the most comprehensive bibliography of scientific explanation as of 1989 and it is thus quite out of date, though it remains the definitive history of the topic. [Pit88] is a useful collection. [Woo03] is a good recent overview. I will explain each view as it becomes relevant. I use in various ways the expositions of [Rub90], [SEG$^+$92], [Hun97], [Kle97] and [Ros00]. This chapter presupposes some familiarity with the main theories of explanation whose main points are well known and repeated often in various literatures (e.g., [Lev01] §3.2). We repeat the details of specific theories here only as they directly relate to our discussion. Many introductory books in the philosophy of science (e.g., [Oka02] or [SEG$^+$92]) provide a basic enough outline of the main questions about explanation.

Wesley Salmon ([Sal84], [Sal89]: 94, 118ff) outlines three major types of theories of explanation: ontic, epistemic, and modal. The ontic conception is the one he advocates for throughout much of his own work. It asserts that our knowledge of the process of causation does the explanatory work. Many (e.g., Sosa, Mellor) advocate a modal conception and hold that explanations involve a knowledge of why some facts are necessary, why they must take place. Other modal conceptions (sometimes identified with Gödel) claim that explanations involve knowledge of why some things must be true. The epistemic conception comes in three varieties. The Deductive-Nomological approach is most closely identified with Hempel. Salmon calls this model an "inferential" conception which ties explanation to nomic expectability. Another epistemic conception is the "information theoretic" view which relates to the subsumption of much information under a more compressed rubric, and is most closely identified with Philip Kitcher. Finally, the erotetic or why-question approach is the last epistemic conception of explanation and is most closely identified with Bas van Fraassen.

In this chapter I first look at the causal account of explanation and its applicability (or lack thereof) to mathematical explanation. Then I do the same for the three epistemic conceptions. We postpone discussion of the modal conception to Chapter 4.

In §2.1 and §2.2 I oppose the causal notion of explanation and the D-N model of explanation respectively. I address these models as part of an attempt to stave off potential criticism that charges that I omitted two important models of explanation. The literature records no claim that the causal or D-N models of explanation should be applied to mathematics, though it does contain claims that other models should be considered. It is probably unnecessary to stress that the intuition that these models of explanation will not apply to mathematics are much stronger than the ability of an argument to convince us of that.

In each case mentioned in this chapter, we conclude that the models will not be able to bear the added constraint of mathematics and still give us what we want in a theory of explanation.

2.1 Causal explanations

In this section we address a prominent theory of explanation and its applicability to mathematical explanation. Despite the notoriously unclear connection between causation and explanation (see [Sal89]), causation is a major contender for the theory of explanation, and some philosophers assume that there is a way to interpret causation such that it is applicable to mathematics. If there are causes in mathematics then a strong case could be made for explanations in mathematics too.

First we show why talking about causation in the context of mathematics may not be as odd as it initially sounds, then we argue that reasonable people have spoken about mathematical causation, and there are reasonable theories of causation that initially appear to be broad enough to accommodate mathematics. Therefore it should be reasonable to speak of causal explanation in mathematics. We then respond to this argument by showing that because of the nature of even the most liberal theory of causation which may initially appear applicable to mathematics, mathematics is lacking the most essential features considered integral to any theory of causation. Finally we look at how Michael Detlefsen sees Frege's conception of mathematical explanation as akin to a causal theory of explanation, and we review Detlefsen's argument against Frege.

First let us explore the relationship between causation and explanation.

Many philosophers see a connection between causation and explanation. Aristotle ([Ari75]), who gives the first recorded careful analysis of scientific explanation, sees causation as just another way of talking about explanation (see [Rub90]: ch. 3). Sosa and Tooley, in their introduction to *Causation* ([ST93]), outright assert the connection when they state that "causation, conditionals, explanation, confirmation, disposition and laws form a cluster of closely related topics in metaphysics, philosophy of language, and philosophy of science."[3] Donald Davidson claims that many uses of the word "cause" are best expressed by "causally explains" ([Dav67]: 162). Hempel and Oppenheimer ([HO48]: 250) claim that their Deductive Nomological (D-N) model of explanation is often called a causal model. Michael Scriven, an early opponent of the D-N model ([Scr75]) refers to the "curious love-hate relationship" between causation and explanation. And Stathis Psillos in his study of causation and explanation concludes that "the fact of the matter is that the concepts of causation, laws of nature and explanation form a quite tight web. Hardly any progress can be made in the elucidation of any of those without engaging in the elucidation of at least some of the others." ([Psi02]: 12) We see that no discussion of explanation can responsibly omit at least a cursory discussion of causation.

Let us next explore the relationship between mathematics and causation. Mathematics is usually not taken as the type of discipline that admits causes.

---

[3] For reasons that will become clear later on it is curious that they omit "prediction" from this list.

We do not generally use causal language in mathematics - even metaphorically. However, as we just saw, some very significant accounts of explanation appeal directly to causes: if you can talk about what causes something, you have explained it, and more importantly, vice-versa, to explain something is [just] to give a causal story, or an account of how events fit into the causal structure of the world. On this account there should not be any way to talk about explanations in mathematics. Resnik and Kushner note that "...we explain in terms of causal processes, and these are conspicuously absent from the realm of mathematics." ([RK87]: 151) Desanti too notes that: "there are no phenomena in the field of mathematics which relate to the explanatory steps and the usual forms of causality ..." ([Des73]: 57) Samuel John Butchard also summarily dismisses causal explanations in mathematics ([But01]: 329).

It would seem that the natural assumption is that there are no causes in mathematics. And if causation really is as relevant to explanation as we claimed a few paragraphs back, then there is no explanation in mathematics.[4] This argument is the simplest and most straight-forward argument against mathematical explanation.

However, the history of mathematics and philosophy is rife with thinkers who do not accept the premises or conclusion of the above argument. Plato (*Phaedo* 96e-97b), many neoplatonists, and pythagoreans were certainly not

---

[4] This is frequently pointed out in the literature on mathematical explanation just before an alternate view of explanation is put forth. See e.g., [Tap05]: 158.

shy about speaking of "the causes of number" or number's explanatory power.[5] Neither were many mathematicians even as recently as the modern period. Aristotle too considered causes in mathematics. Hippocrates Apostle reads Aristotle as follows:

> When one who wishes to have scientific knowledge continues pressing a question as to why something either is or changes or is true, he is asking for the first cause or causes; and the final answer must be one or more of the four causes mentioned . . . Why is the sum of the angles of a triangle equal to two right angles? Because of this, and because of that, and that because of the definition of the plane and of the straight line, and definitions in mathematics are the formal causes and starting points of demonstrations. ([Apo52]: 50)

The thirteenth century logician Richard Rufus referred to what he called a "demonstrational cause" which he treated as a species of Aristotelian ef-

---

[5] In Chapter 1.2 we first saw Plato and the twelfth century philosopher Abraham ibn Ezra ([Ezr95]) describe the number "one" as a cause of the rest of the numbers, or the cause of some other property of number. Bachya ben Asher in the thirteenth century makes lengthy analogies between "one" and God, where God is said to have certain causal powers analogous to the causal powers of the number "one" (and not the other way around):

> The number "one" does not begin a computation. Instead it is the cause of the computation. Now, just as "one" is the cause of the computation [of the numbers which succeeded it] but in itself is not a part thereof, so the Creator is the cause of the existence of the world but is not a part of it. Just as all numbers would be nullified by the removal of the power of "one" - they exist as numbers only because of "one" - so would all that exists in heaven and on earth . . . Furthermore the power of "one" would remain even if you removed all other numbers, . . . Just as "one" is the cause of "two," through which is causes "three," etc until "nine," the last of the units, so the one God is the cause of all things . . . Just as the numeral "one" is the cause of calculation and is not itself a part thereof, so the creator. . . [Bac80]: 362, 375-376)

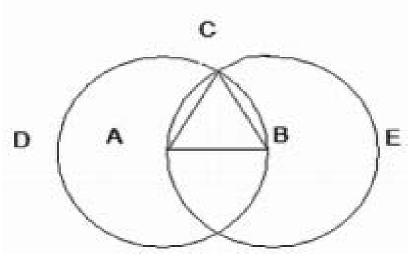

Figure 2.1: Causal Proof

ficient cause. As Rega Wood and Robert Andrews explicate: "Demonstrational causality refers to the dependence of the conclusion on the premises of a demonstration . . . Aristotle and his medieval commentators describe demonstrational causality in the same terms as efficient causality. Aristotle speaks of the conclusion as an "effect" of the premises; his commentators speak of the "sufficiency" of first principles or axioms (*dignitates*) in producing the conclusion" ([WA96]: 325). Rufus specifically cites the mathematical definition of a triangle in this context (331).

In the modern period the question of causes in mathematics was quite predominant. The following was taken as a prototypical example of a "causal proof" (from [Man96]: 18) (See Figure 2.1): Let AB be a segment. Draw two circles with radii of equal length AB and centers in A and B, respectively. Let C be one of the points where the circles intersect. Connect A and C and C and B. Then ABC is an equilateral triangle, since its sides are equal to the radius of the same circle, and thus are equal to each other.

According to Mancosu, the mathematician Biancani relying on an earlier argument of Proclus, argues that this is a causal proof "since it shows that the cause of the equality of the sides of the triangle ABC is that they are radii of equal circles." And the argument ultimately rests on the definition of a circle, which thus acts as the formal cause of the proof. Biancani then deals with material causes by arguing that proposition I.32 [in Euclid's *Elements*] "...proceeds by material causes when it infers the equality of the wholes from the equality of the parts." Arguments for causation in mathematics were also put forth by mathematicians as esteemed as Isaac Barrow.[6]

We note also that at least one contemporary philosopher of mathematics has considered defending a related notion of mathematical causation on similar grounds to ones mentioned:

> ...A more daring response to the claim that causes are not operative in mathematics is to deny it. In the Hellenistic era, when a broader conception of cause prevailed, this would not seem unreasonable. Proclus remarks that 'many persons have thought that geometry does not investigate the cause, that it, does not ask the question "why?", but then argues that they are mistaken (Morrow 1970: 158-9). As Mancosu (1996) explains, the notion of cause was also at play in seventeenth-century mathematics. Today, mathematicians are happy to say that the existence of the octonions is what *causes* the periodicity of the homotopy groups of $O(\infty)$, the inductive limit of $O(n)$ of that the reason that there are exceptional Lie groups is *because* the covering group of $SO(8)$ has an outer automorphism. More generally, when you are study-

---

[6] See Isaac Barrow's [Bar34], especially Lecture VI "Of the Causality of Mathematical Demonstration" where he has a lengthy discussion of what it means to be a causal demonstration. There he also uses the same example as the one we cited above.

> ing a mathematical object which displays a certain property, it is very common to wonder which features of the object are 'responsible for' that property. ([Cor03]: 117, emphasis in original.)

The Renaissance-era motivation behind describing certain mathematical processes as causal is fairly straight-forward given an Aristotelian epistemology. Causation, particularly as Aristotle conceived it (when he speaks of "the four causes") were taken to be the touchstone of a respectable science. Thus defenders of the "dignity of mathematics" had to show that mathematics was epistemologically on par with science, and thus admitted causes just like science did.

We also find in a 1869 address, the mathematician James Joseph Sylvester, a man who significantly participated in the trend toward greater abstraction in nineteenth century algebra, offering the following response to a criticism of T. H. Huxley:

> ...we are told [by Huxley] that 'mathematics is that study which knows nothing of observation, nothing of experiment, nothing of causation.'
>
> I think no statement could have been made more opposite the undoubted facts of the case, that mathematical analysis is constantly invoking the aid of new principles, new ideas, new methods ...([Syl96]: §12)

Further on Sylvester clarifies:

> ...I could tell a story of almost romantic interest about my own latest researches in a field where Geometry, Algebra, and the

> Theory of Numbers melt in a surprising manner into one another ... which would very strikingly illustrate how much observation, divination, induction, experimental trial, and verification, *causation, too* (if that means, as I suppose it must, mounting from phenomena to their reasons or causes of being), have to do with the work of the mathematicians. (§16, *emphasis added*)

We see then that causation persists in the consciousness of mathematicians well into the nineteenth century, and beyond the time when questions of Aristotelianism were prominent.

The historian of mathematics Sanford Segal also uses causal language when speaking of mathematics:

> Mathematics also has a notion of *strict causality*: if A then B ... The notion that it is conceivable that B can be shown to always follow from A is central to mathematics ... Both the necessary process of abstraction and the idea of mathematical causality separate mathematics from more mundane areas. ... Mathematical abstraction and mathematical causality seem to elevate mathematics above the sphere of the larger culture. ([Seg03]: 2, *emphasis added*)

So it is not that unusual to find causal language in mathematical writing. The fact that much of the talk about causation in mathematics is incorporated in defunct debates does not detract from the strength of our claim. It merely shows that there is a plausible interpretation of the history of mathematics that has mathematicians worrying about causes and causal explanations.

But we also want to argue that in addition to mathematics being amenable to causal talk, the nature of causation does not rule out causal explanations in mathematics. To see if there might be a sense in which causation can be taken to include mathematics, let us take a very brief detour into the broader question of causation.

Like "explanation", there is no single agreed-upon analysis of "cause". There are currently three standard approaches to causation taken seriously enough in the literature that we should explore. (1) The Regularity View of Causation (RVC), (2) the mechanism approach, and (3) a pluralist view. (1) The Humean analysis (RVC)[7] treats causation as fundamental to *our thinking about the world.* (2) The mechanistic approach treats causation as a fundamental part *of the world itself.* In mathematics neither (1) nor (2) seem *prima facie* plausible. Let us examine these three notions of causation in tern.

(1) Can Hume's notion of causation as "constant conjunction" be applied to mathematics? Hume's view of causation (as given in [Psi02]: 19) is as follows:

> c causes e iff:
> (a) c is spatiotemporally contiguous to e;
> (b) e succeeds c in time; and
> (c) all events of type C (i.e., all events that are like c) are regularly followed by (are constantly conjoined with) events of type E (i.e., events like e).

---

[7] This is not the place to take sides on the historical questions regarding the actual position Hume held. With apologies to Norman Kemp-Smith and Galen Strawson, we call the RVC the "Humean view".

Many questions emerge when we start replacing c and e with mathematical entities. What is "constantly conjoined" in mathematics? Can we substitute c with premises of some proof, and c with the theorem? Or, perhaps the mathematical objects on each side of an equality are constantly conjoined? No, in both cases they are not; and it is easy to diagnose why we can so easily dismiss this view of causation in mathematics.

First, to even consider of the RVC in mathematics we have to immediately ignore conditions (a) and (b) involving space and time. Whatever mathematical objects are, and however they are known, they are not spatiotemporal. They are abstract in some way. That alone ought to be sufficient to dismiss the RVC as inapplicable to mathematics.

Second, mathematics contains no events to conjoin. Mathematical objects are not events, neither are proofs, equalities, or constructions. Events are indispensable features of the RVC. So the RVC, whatever its merits for scientific causation, is unlikely to be able to fit into a theory of causal explanation.

Third, any premises P can have various consequences C, C', C'' .... So the fact that $P \models C$, $P \models C'$, $P \models C''$, ... tells us that to the extent we "experience" P it may not be conjoined with C, but rather some C'. The fact that C always follows from P is irrelevant.

Fourth, a corollary of the RVC is that there is no necessity in causation. The usual assumption in mathematics is that mathematical statements are necessary. While this reason itself does not rule out the RVC in mathematics, it suggests a significant disanalogy that tells us that we are not dealing with subjects that have similar methodologies.

Finally, mathematical statements are not known because of anything but the fact that they can be proven. Proofs also need not have any observer, as we do not "observe" any mathematical objects.[8] Proofs as we just said, contain no observation statements in support of their conclusions. "Conjunction" then, as it is used in this context is observer dependent. Therefore, here it cannot be taken literally.

(2) The second account of causation, the mechanism view (see e.g., [Mac74]: ch. 8), takes causation to be part of the world itself. The account generally has causation being the transfer of energy or momentum from the cause to the effect or, it posits a mechanism that links cause and effect. Mackie takes the mechanism to consist in the qualitative or structural continuity, or the persistence exhibited by certain processes. What persists can be the total energy of a system, the number of particles, etc.

Can we apply this analysis of causation to mathematics? Again, we cannot. In a proof, the premises P do not transfer anything to the conclusion C, nor does anything "move" from one side of an equation to another. Besides sounding anachronistically scholastic, it is hard to even grasp what this might mean to a mathematician, as there is no "going" from one "side" to the other,

---

[8] Gödel's account of mathematical perception notwithstanding.

nor is there really one side at all. Nor is one side loosing anything in the process. So nothing is preserved. An equality does not transfer anything, nor does an argument. Though equalities, by definition, assert the equivalence of whatever is on each side of an equation, they do not preserve that equality. Nothing in mathematics can even be described in the *a posteori* vocabulary of "transference".

So clearly the two standard accounts of causation do little for us but make the problem look absurd and not worth mentioning. But perhaps if we looked at more inclusive views of causation we can find a more helpful account of causation for mathematics.

(3) A third view of causation which has been gaining prominence of late, traces its lineage to a remark of G. E. M Anscombe ([Ans93]). The theory that eventually emerged is causal pluralism. This view argues that there is no single relation that we call "causation". Rather there is a wide variety of relations that can be said to obtain between causes and their effects. Nancy Cartwright (in [Car06] and elsewhere) can also be said to be arguing that causation (what we call "causation") is not a homogeneous set of phenomenon. Causation might just be whatever we call "causal" in some context, and there is no one thing that is a causal relation. The motivation for the causal minimalism view is the wide variety of relations that we generally take to fall under the "causal" heading. There are no specific physical (or even

metaphysical) constraints to "causation". [9] Peter Godfrey Smith describes such a theory as the "causal minimalism" thesis:

> 1. "C was a cause of E" is true iff the relation between C and E can also be described using some member of set S, or can be described as a chain of relations each of which can be described using some member of S.
> 2. S is a set of causal verbs and other linguistic formulas which represent "special causal concepts" in Anscombe's sense.

Taking S as open we should be able to include all sorts of specific relations as causal, including mathematical ones.[10]

Given such an inclusive view of causation it now seems more reasonable to ask if there are causes (and effects) in mathematics. Consider the following argument: First, claim that some theorem is true, as it has a proof (or that some feature of a theorem is true because of some specific feature of its proof). Then claim that the proof can be taken to be the cause of the theorem because the verb "proves" might be a member of Anscombe's set S. So when we say that α proves p, we are using terminology that would count as causal on this account.

This too will not work. First, "proves" is the wrong kind of verb for S. Here is why: the function of a proof is not in any sense to *make* a theorem true. Its function is to demonstrate that the theorem follows from the

---

[9] A related view is implied by Jody Azzouni ([Azz00]: §6). Implicit in his work is that causation is ontologically neutral, or at least ontologically unimportant.

[10] Peter Godfrey Smith. "Causal Pluralism"; forthcoming.

premises, or to establish the theorem. A proof serves an epistemic function vis à vis the theorem. The theorem followed before the proof, only no one could claim to know it. Thus whatever the constraints on S are, it must be something that makes the theorem itself true, not something that makes us believe that the theorem is true. The same reasoning should hold true for related mathematical verbs like "construct".

The second reason that the pluralist theory of causation cannot work is best expressed in the language of the "platitude camp".[11] Platitudes are what are given (for now) in place of necessary or sufficient conditions for causation. They provide some constraints on any theory of causation. Hence the "Difference Platitude" described as follows: In science there is a clear intuition that whatever causes ultimately turn out to be, things would be different if some of the causes were absent. This is usually phrased counterfactually: if the cause hadn't been, the effect would not not have been either. So, for example, a scientist might claim that if humans hadn't polluted the atmosphere, the ozone layer would be larger. Humans polluting the atmosphere is a cause of the ozone layer shrinking. The cause made some difference.

Given this platitude as a restriction on an analysis of causation, mathematics cannot be causal because it is hard to imagine how to talk about what is absent in mathematics. A proof, let's say, or its absence, only makes

---

[11] Advocates of this approach include Mellor, Menzies, and Armstrong. See [Psi02]: 6.

a difference to some statement about *our knowing* the truth value of the theorem. Nor would a change in the proof matter. A change in the proof would not make a theorem false. It would just be the proof of something else, or a different proof of the same theorem, or it would not be a proof at all.

But a theory of explanation (of the kind we are interested in) is not one in which we are accounting for someone's knowledge of a theorem. Rather we seek to account for the theorem itself.

We cannot reasonably speak about mathematics as a structure or body of knowledge that is different than the way it is, given the actual absence of some part of mathematics. We can talk about some part of mathematics that we do not understand, or did not know at some point in history. And we can talk about what mathematics would be like if we were somehow restricted, say we could not use the axiom of choice, or we could only use constructive proofs. But we could not, for example, sensibly talk about our current mathematical theory and wonder what it would be like if there were no prime numbers.

So whatever the ultimate analysis of causation ends up being, mathematics will not be analyzeable in those terms, as mathematics cannot accommodate one of the fundamental intuitions about causes.

Finally, another way which we can consider causes operating in mathematics is to take something in mathematics to be serving the same functional role in mathematics that causes serve in science. If we can show that something exists that "serves a causal role" then we can argue that there is an analogy between science and mathematics, and hence something functionally equivalent to explanation in mathematics. Michael Detlefsen ([Det88]) examines this approach to mathematical explanation via a critique of Frege's notion of "grounding". Detlefsen looks at what he takes to be Frege's notion of mathematical explanation.

Frege's notion of explanation originates with Aristotle, who spoke of some truths being "prior to" or being the "cause of" other truths. Aristotle's ideal scientific method (as expressed in the *Posterior Analytics*) had all truths derived from the most basic truths. Distinguishing the basic truths from the derivative truths was fundamental to Aristotle's conception of an ideal demonstrative science.

Leibniz too[12] insists that there are some truths which cause other truths, and there is a "natural order of truths". Bolzano too (as we discuss further on) claimed that one virtue of a genuine proof was that the proof presents the "objective ground" of its conclusion.

Detlefsen shows how it is possible to read Frege as offering a concept of explanation in mathematics akin to a causal explanation in science. He then shows how Frege's attempt at this fails. He shows that for Frege explanation is rooted in the idea of "the grounding of propositions". There is a relationship between a proposition p and its ground g, such that "g grounds

---

[12] See *New Essays in Human Understanding*: Book IV, ch. xvii, par 3, and ch. vii, par 9.

p" means that "g explains p". Detlefsen shows that this grounding relation cannot be analogous to the causation relation. If explanation in mathematics is a relation between g and p, it cannot be the same relation that exists between a cause and an effect because in the grounding relation g needs to be necessary for p, but in the causal relation, the cause is not necessary for the effect. Detlefsen exploits a specific fact about causation - namely that causation involves contingency, to counter Frege's notion of explanation in mathematics.

Detlefsen argues that on Frege's account of the grounding relation the proof shows the relation between the grounding propositions and the grounded ones within an *objective hierarchy.* The "grounding relation" can either be explanatory in virtue of one of two things - either the explanatory power rests in the type of implication relation, or it is explanatory in virtue of the relata. Detlefsen argues that the "grounding proofs" are different from non-grounding ones in virtue of their relata. As Frege held a global view of implication, i.e, his account of implication is uniform across all domains, the difference between grounding and non-grounding proofs could not be in the implicational structure, otherwise everything would be explanatory and the grounding proofs would be no different from any other. This would miss Frege's point of an objective hierarchy.

What must the grounding relation consist of? Detlefsen gives two criteria. First, the grounding proposition, g, must imply the grounded proposition, p.

Second, g must in some way be necessary for p (i.e., without g there would be no p). If the relata met those conditions, the relation would be explanatory.

One disanalogy with causation is readily apparent. A cause, c, could not have occurred and effect, e, still could have. So to say that "g is necessary for p" cannot be the same as saying that "c is necessary for e" because c is rarely necessary for e given that most e are multiply realizable.

c's being the cause of e is different from c's being necessary for e because of the following: (a) for c to cause e without c being necessary for e there must be a possible world with c but not e. I.e., a cause must be contingent. (b) e can only occur (as a caused event) without its actual cause if (1) there are events of a type other that c which can be the cause of e, or, (2) if there are events other than c which belong to the same nomic type as c which are subsumable by the same law that subsumes c and e.

Can a parallel analysis to be applied to mathematics? Consider this: if we only had the above requirement (b) but not (a) many events can cause e, and they will occur. This is incompatible with a causal hierarchy because many things *are fully able* to be the cause of c but only one *actually* is. But since many c's can be the cause of e there is no reason to attribute one of them as *the* cause of e rather than any others. Yet a hierarchy of causes like the one Frege sought must specify only one.

So if you have a hierarchy of actual causes with more than one event which can fully be the cause, then you must say that only one cause is actual. Thus you must have access to a conceptual device like (a), a way to sort out the actual world from other worlds.

We cannot have this in mathematics because mathematics contains no contingent propositions. So if g and $g\prime$ can both ground p there is no way to decide which is the *unique ground* for p.

So it appears that there cannot be more than one g that is fully capable of grounding p, but if there is only one, then g is necessary for p (because without g there is no p). So it appears that a grounded proposition can be true only if its one actual ground is.

To have a grounding theory then, you need to interpret necessity in a way that makes this all work out. Detlefsen offers two interpretations, both of which he claims he is unable to square with Frege's original motivation of a global theory of implication. Thus Frege's version of mathematical explanation must fail, at least if he is to maintain consistency with his own view of implication.

Detlefsen's point, made slightly more generally is as follows: Assume you have a foundational program in mathematics. That will give you an objective ordering of truths in a mathematical domain. Then if you take a view of explanation that says that the objects of the lower level (the foundation) explain the objects in the higher levels you will not be able to have a theory of explanation akin to any reasonable theory of causal explanation. However, we should keep in mind that Detlefsen's case is only made by interpreting the grounding relation in causal terms. Detlefsen claims that the value of looking at the grounding relation as analogous to a causal relation is that, as Frege says, it captures the fixed, objective "natural order of truths" and not something rooted in the psychology of its practitioners. But Detlefsen considers only one version of explanation that could make Frege's grounding relation objective and also show *why* it is true. Other accounts of explanation, like the one Kitcher attributes to Bolzano (which we address later on) do not rely on causal notions, or analogies with causation. Thus while it seems that the case may be closed for causal accounts of explanation in mathematics, there may be other interpretations of Frege's grounding relation that allow it to do the work Frege wanted, i.e., situate various propositions within an objective metaphysical hierarchy such that the lower propositions explain the higher ones.

Moreover, exploiting an analogy between causation and necessity is fraught with the additional danger of begging certain questions. Any analogy requires a relevant similarity between the two domains. In this case the analogy would state that mathematics and science have enough in common so that we compare them. But having an account of explanation in common maybe claimed to be the very thing (or one of the things) that makes them analogous. But what we are trying to find out is whether or not they are analogous. Regardless, a case would first have to be made for the similarity of mathematics and science. Furthermore, even if mathematics is continuous with science, or mathematics and science are in some other ways analogous, it does not imply that its methodology need be.

When saying that necessity, or something else, *plays the (functional) role* in mathematics that causation plays in the physical sciences, one needs to show that the structures in which they play the same role are relevantly similar. This is not shown, even by those who take the domains of mathematics and the sciences to be the same.

Therefore for the above reasons we conclude that causation, whatever its merits as a foundation for scientific explanation, is unable to provide a route to explanation in mathematics. There are no mathematical causes, nor do we find anything sufficiently analogous to causation in mathematics which allows us to consider it analogous to causal explanation.

2.2 The Deductive-Nomological model

This brief section explores the applicability of Hempel's Deductive Nomological (D-N) model of explanation to mathematics. We show that while it may at first appear that there are reasons to believe that the prospects for mathematical explanation look good if we take for granted the D-N model, they are in fact quite poor.

In some ways the original positivist model of explanation - Hempel and Oppenheim's (D-N) model[13] may appear well-suited as a candidate for explanation in mathematics. In other ways the D-N model also illuminates the difficulties of the view that there is explanation in mathematics. Proponents of the D-N model maintain that there are two conditions that are separately necessary and jointly sufficient for a successful scientific explanation: (1) explanatory relevance - the sentences that explain have to be relevant to that which is being explained, and (2) testability - the assertions in the explanation must be testable. The D-N model of explanation meets those two requirements by appealing to laws that govern all explainable phenomena, and by asserting that explanation is a matter of deduction involving the laws of nature and particular background data. A D-N explanation is a deductive argument whose conclusion, E, is the statement we are trying to explain (the *explanandum*) and whose premises (the *explanans*) do the explaining. The

---

[13] [Hem65a] contains all relevant papers on his classic account of the D-N model of explanation, especially [HO48]. [Hem66] contains a shorter more popular presentation of the D-N model.

explanans consist of a statement of a general law (or laws), $L_1, L_2, \ldots, L_n$, and other statement(s), $C_1, C_2, \ldots, C_n$, which make assertions about particular facts. The argument is then represented by the following schema:

$L_1, L_2, \ldots, L_n$ General laws
$C_1, C_2, \ldots, C_n$ Particular facts
$\Rightarrow$
E Phenomena to be explained

Unlike other models of explanation which attempt to explain many facts all at once, D-N explanations account for particular facts. They do so by invoking laws, and so this model is often called the "covering-law model" of scientific explanation.

There is no literature that puts forth arguments in support of the use of D-N explanations in mathematics. One possible reason for this is that Hempel, the originator of the model explicitly excludes mathematical cases from falling under the D-N model. He claimed as follows:

> . . . an explanation of why every equilateral triangle is equiangular, or why an integer is divisible by 9 whenever the sum of its digits in decimal representation is so divisible requires an argument whose conclusion expresses the proposition in question, and whose premises include general geometrical or arithmetical statements, not of course empirical laws; nor for that matter, is the explanandum statement an empirical one. This sort of explaining, though rather closely related to the kind with which we are

> concerned, is not meant to be covered by our models. ([Hem63]: 126)

The passage would appear to be specifically excluding to mathematical explanations. But Hempel does not elaborate beyond this passage. Hempel asserts that because both mathematical arguments and scientific explanations primarily involve deductive structures, they are similar. However, because on his view scientific explanations must appeal to empirical laws, and mathematical arguments do not, they are dissimilar; and because of this dissimilarity, mathematics does not contain explanations.

The reason why one might initially judge the D-N model as being as appropriate a theory of explanation for mathematics as it is for the sciences is not difficult to see. Proponents of the D-N model claim on its behalf that it is a superior model of explanation in virtue of its purely deductive structure. So too mathematics which is also wholly a deductive system would be well suited to derive the same philosophical benefit from its very nature as a deductive system. This is what Hempel undoubtedly meant.

The D-N model of scientific explanation was devised as a way to capture the notion of explanation without the metaphysical baggage inherent in earlier models of explanation. All that the D-N model requires is to take particular pieces of [empirical] information $(C_1, C_2, \ldots, C_n)$, specific kinds of generalizations of information about the scientific domain in question (the laws, i.e., $L_1, L_2, \ldots, L_n$) and use [metaphysics-free] deduction to put the two together. Thus we have explanations of particular facts.

It then stands to reason that if you can get explanations out of deductions in science, you should be able to get explanations out of deductions in mathematics. It should matter little that the facts in the mathematical and scientific sentences refer to objects that may be in different ontological categories. The explanations emerge in virtue of the nature of the inferences and the logical form of the explanans sentences that the inferences operate on. There are mathematical sentences that match the logical forms of the sentences used in D-N explanations (of scientific phenomena). The type of inferences also match up, as deduction is uniform across mathematics and science.

Nonetheless, despite the apparent relevance of the D-N model for mathematical explanation there is a strong intuition that the model will not work here.[14] Samuel John Butchart articulates one reason for this. He argues that "This [D-N] account of scientific explanation fails spectacularly in the mathematical case, since *all* mathematical proofs can be thought of as D-N arguments." ([But01]: 328, emphasis in original.)

While we share Butchart's enthusiasm about the failure of D-N explanations in the mathematical case, we do not quite agree with his reasons for believing it fails. We believe that there are more serious reasons why the ac-

---

[14] Kitcher construes Bolzano as presenting a theory of explanation that should apply to science and also mathematics. Per Kitcher's reconstruction of the motivation for Bolzano's particular account of explanation (as well as Kitcher's own subsequent account) was the general inadequacy of D-N explanations in science. See Kitcher's [Kit75]: §V.

count fails. According to Butchart, presumably what makes the D-N model "fail spectacularly" is that if the D-N model is applied to mathematics then there will be no cases in mathematics where the argument is valid and there fails to be an explanation. Lakatos seems to have held a similar view and took all proofs to be explanations. But it is unclear what Lakatos meant by "explanation". It is likely that he meant "explanation" in some pedagogical or psychological sense - certainly he is not referring to D-N explanations.[15]

Lakatos aside, on the D-N account the explanatory model trivializes what it means to explain in mathematics. Perhaps, Butchart could shore up his case by pointing out that even those who believe that there are explanations in mathematics (e.g., Steiner, Kitcher, Mancosu) distinguish explanatory from non-explanatory mathematics. Often the impetus for looking for mathematical explanations comes from the intuition that there are some proofs (in particular) that are "more explanatory" than others (see e.g., Steiner, Detlefsen's Frege, Kitcher's Bolzano, Mancosu's Cournot) and some particular bits of mathematics that explain. That suggests that their intuition is that there are other proofs that are not explanatory, or less explanatory, and there are parts of mathematics that do not explain. So the intuition that all mathematics explains, contradicts the widely-held intuition that there is a restricted class of explanatory mathematics.

But Butchart's reasoning fails to do justice to the D-N model in two ways.

---

[15] See §3.2 below for more on Lakatos.

First, the D-N model will not just accept any old deduction in science as a scientific explanation. The valid inference

Metal M and metal $M\prime$ expanded when heated
Therefore
Metal M expanded when heated

will not count as a D-N explanation of why M expanded. One reason is that the above argument does not invoke any laws, so it does not fit the form of a D-N explanation. We have an analogous situation in mathematics. Putting aside for the moment the question of what a law in mathematics might be, the simple valid logic proof:

$p\&q$
Therefore
p

is analogous to our failed explanation above, and is a perfectly reasonable mathematical proof. The cases can certainly get more complicated, but nonetheless Butchart is wrong to assert that the mathematical case would allow *any mathematical proof* to be a D-N argument. Our above cases are not explanations because D-N arguments must contain sentences of a specific logical form. Why shouldn't proponents of a D-N model for mathematics attempt, as Detlefsen claimed Frege did (see §2.1 above), to articulate a specific logical form for the sentences of the relata of a mathematical D-N explanation?

The second reason Butchard's argument does not do justice to the potential of the D-N model of explanation in mathematics is that there is nothing necessarily problematic about explaining too much in (a domain). In other words, let us suppose that every proof that meets some particular pattern (as we insisted above) explains the thing that it is proving. But a pattern that managed to explain everything it proved would seem to indicate the success of the method of proof as an explanation rather than its failure as an explanatory program. Proponents of the D-N model certainly think that any particular fact that can be phrased as the conclusion of an argument (that somewhere invokes a law that subsumes the fact) is a good explanation. Every particular scientific fact that is or can be nomically subsumed is explained. Why should we count similar success in mathematics against a theory of mathematical explanation? So even if, as Butchart claims, the D-N model would allow everything to be explained, which it wouldn't because not every mathematical argument is an explanation, it would not count against the D-N model being applied to mathematics.

We need a stronger case against applying the D-N model to mathematics. Let us take a closer look at the D-N model. Wesley Salmon ([Sal89]: 24-25) lists the following six theses that are advanced by Hempel and Oppenheim in their discussion of the D-N model: (1) All legitimate explanations are arguments. (2) Every explanation must appeal to at least one law. (3) There is a symmetry between explanation and prediction. (4) Causality does not play a role in explanation. (5) The explanans is literally true. (6) Explanations are amenable to formal (or quasi-formal) modeling.

The above theses all relate to explanation in a way that would be acceptable to those with the metaphysical and ontological scruples of the positivists. But each of these theses is controversial and has been disputed by numerous philosophers. While they have been rejected, their rejection has generally been motivated by the belief that an alternative approach to explanation is correct, and not by a desire to modify the D-N model, and so proponents of the D-N model, still retain and defend all of the above theses.

The famous counter-examples to the D-N approach (see [Sal89]: §2.3.) each exploit one or more of the features of the D-N model mentioned above. For example, Sylvan Bromberger's famous flagpole example exploits thesis (4). The fact that there is no causality constraint in D-N explanations allows for the following absurdity: A flagpole of a certain height stands perpendicular to a level surface. The sun shines brightly from a certain elevation. The pole casts a shadow of a certain length. Given the height of the flagpole, the elevation of the sun (i.e., specific facts) and the law of the rectilinear propagation of light (i.e., a law) we can deduce the length of the flagpole's shadow. However if we took the same law but used the length of the shadow and the elevation of the sun as our specific facts, we end up with a deduction of the height of the flagpole. The D-N model would have to count this latter deduction as a scientific explanation as well. However, the flagpole's height might have been determined by the whims of the local city council. So the height ends up being explained by the laws of light propagation when the true explanation is the whims of the local council. The D-N model is incapable of ascribing the explanation to the whims of the council rather than the laws of light propagation. So the model of explanation fails. It appears that the causality thesis (4) is the culprit here because the light and flagpole do cause the shadow (thereby perhaps explaining the shadow's length), but the light and shadow do not cause the flagpole. Because of counterexamples like the flagpole one, some philosophers insist that there can be no explanation without some appeal to the causal factors involved.

Similarly, other famous counterexamples exploit vulnerabilities in the other theses. Proponents of the D-N model therefore must show how these theses are not vulnerabilities of the D-N model. That is, proponents of the D-N model must show how the counterexamples do not diminish the model's ability to maintain these theses. Let us then see of any of these theses can be defended under a light of mathematical application of the D-N model. If they cannot, then it will be difficult to maintain that the D-N model could be defended for mathematics.

Setting out the philosophical theses as we did above should clarify more about what in the D-N model would be appealing to supporters of explanation in mathematics, and what would not.

In what way are the mathematical cases similar to the scientific cases with respect to the D-N model? *Prima facie* if we took the D-N model of explanation to apply to mathematics we could retain theses (1), (2), (4), (5), and (6). There are many features of the D-N model that the model's proponents take to be part of explanations that would still hold if we applied the D-N model to mathematics. That is, both the mathematical explanations and D-N explanations in science would appear to conform to the general structure of a formalizable argument (theses (1) and (6)) that does not appeal to causes (thesis (4)) and is literally true (thesis (5)). Also, the fact that with the exception of elementary arithmetical identities, all mathematical reasoning involves some general premise, should allow us to also retain thesis (2).

The fact that we have literally true, non-causal, formalizable arguments as explanations in science and also in mathematics provides a strong initial case for the applicability of the D-N model to mathematics.

But let us examine these theses more carefully. Thesis (1) would seem to hold for any kind of D-N argument, even mathematical ones. It might even hold "better" for the mathematical cases than in scientific cases because scientific reasoning is usually not deductive. We should probably demand an argument when we hear otherwise, or when we hear that science generally applies standard D-N-like arguments.

There is much more to say however with respect to thesis (2). Earlier in this section we sidestepped the issue of laws in mathematics, but let us briefly take this up here.

Despite the fact that there are some things in mathematics that we call "laws", there are no laws in mathematics akin to those in science.[16] General premises are not laws. There are of course things we refer to as the "laws of associativity and distributivity", the "law of excluded middle", laws of trigonometry - such as the "law of cosines", and laws of probability like the "law of large numbers"[17]. However, unlike laws of nature, laws in mathematics are *prescriptive*, that is they tell us how to apply certain rules. The rules declare that such and such a mathematical function or procedure is to be applied and manipulated in such and such a way; mathematics may have laws determining the admissible operations on given structures. They somehow summarize what we observe in nature. The world dictates the laws. But the laws of nature (that apply in science) are *descriptive* rather than prescriptive. We declare nothing about nature by fiat. The laws have been determined to hold of the natural world.[18] If we attempt to treat the laws of mathematics mentioned above in the same way we treat scientific laws,

---

[16] Kitcher makes a similar point in [Kit75]: 255.

[17] The law of large numbers is actually ambiguous as to whether it is a mathematical theorem about convergence or an empirical law. As a mathematical law it behaves like other theorems, as an empirical law, it does not belong on the list at all.

[18] We do not need to address the ever-growing debates concerning laws of nature. The usual assumption is that there are laws of nature that are something over and above universal generalizations. Hempel assumed this as well, and to the extent that I use the D-N account, I use Hempel's notion of laws (in [HO48]) as well.

will have the same problems that affected causation in the previous section. Mathematical laws are metaphysically different from the contingent laws of nature and they cannot be treated as if they were on an ontological par. So thesis (2) highlights an important reason why we can consider the scientific version of the D-N model, but not a mathematical version.

Another problem with the D-N account involves the symmetry with prediction, or thesis (3).[19] Prediction is trivial in mathematics. If all theorems deductively follow from other theorems (in whatever way they do), and prediction is just what is logically supposed to be deducible from some data, then all of mathematics is predictive, and to the extent that prediction is symmetrical with explanation, all mathematics is hence explained. But unlike the problem confronted in our earlier criticism of Butchart, prediction makes a piece of mathematics explanatory merely in virtue of its being mathematical.

Of course, ask any mathematician if there is prediction in mathematics and most will say that there is. Most mathematicians would have claimed to have predicted the truth of Fermat's last theorem, or would predict that Goldbach's conjecture will be found to be true. But cases like that are instances of *mathematicians* predicting that something will turn out to be true. They are not instances of the *mathematics itself* predicting the truth of the respective theorems. In science we speak of a theory making predictions, not

---

[19] The symmetry between prediction and explanation is articulated in many places including [Mil41]: Bk III, ch XII §6, p 310, [Hem65b]: 234, [HO48]: 249, and [Pop59]: 10. We return to this argument (with more detail) in §2.3.1 and again in Chapter 4.

the theoretician. When a theory makes a prediction it means that the data, axioms, and conjectures of a theory deductively imply that a certain experiment will have a certain conclusion, or that when the evidence is discovered, it will conform to the deductive prediction of the theory.[20]

In mathematics, if a collection of axioms and theorems deductively imply a result, the result is not a prediction. Prediction in this sense is impossible. The reason for this is in part the temporality of prediction. But even if we ignore this temporality, prediction is the process of being able to take existing data-points and, via an inference or series of inferences, determine a future data-point. In science, doing this is predicting what will happen. For example, when meteorologists predict tomorrow's weather they are inferring from current climactic conditions to future ones. Scientists can also predict what they will find when they will make the appropriate observations (as is often done in astronomy) or they can retrodict what happened in the past (as is often done in the Earth sciences). But in mathematics this process of making the proper inferences is simply the process of doing a proof, and that is all there is in mathematics.

Additionally, in science, prediction is by definition divorced from discovery. In mathematics (on the most liberal account of "prediction") it is discovery. If a theory predicts something, you have not *de facto* discovered that thing. Each prediction is a test of your theory. In mathematics each pre-

---

[20] Michael Levin made this very point in passing, while advancing a very similar thesis for various topics in linguistics. See [Lev77]: 129, 131.

diction is a discovery. It neither corroborates nor disconfirms your "theory". It is a result of it. You have learned something in mathematics rather than predicted it because a prediction is [the ability to] suggest a datum prior to the usual route of discovering it. More precisely, prediction is an epistemic route to some information before you have whatever it is that counts as knowledge of that which was predicted. It is often a conditional belief, given what you know to be the case if you get the result you predict.

That is why it is non-sensical to talk about prediction in mathematics - because deductive inference is the usual route to mathematical knowledge, it is also the only route in mathematics.[21] And prediction is also a matter of deductive inference.

We mentioned metaphysics above. The role of necessity.[22] in mathematics is significant. If science were composed of necessary truths, then it would act like mathematics, where predictions would be as true as the data used to make the prediction. Waiting for experimental confirmation would be uninteresting and unnecessary if your prediction was about a necessary truth.

A way to think about this is by considering a logically omniscient being. A logically omniscient being would have no use for the concept of "prediction" in mathematics. If it knew all the truths of logic there would be nothing to predict. A physically omniscient being can still make use of explanations

---

[21] We can ignore Gödelian views of mathematical intuitions and Ramanujanesque cases of knowledge without impacting our point.

[22] We return to this argument and defend the details more fully in Chapter 4.

to understand what process or data account for which other data, i.e., to "put the pieces of the theory together". A system that does not contain predictions, thus deviates from the traditional explanatory role. Severing the link between explanation and prediction, as you must do in the case of mathematics, certainly makes it more reasonable to wonder what is left of explanation if we say that explanation applies to math too.

We also note that thesis (4) of course applies to a mathematical version of the D-N model[23], and thesis (6) is trivially true of mathematics. But thesis (5) is not obviously applicable to both a science and mathematics version of the D-N model of explanation. Thesis (5)'s applicability to both versions of the D-N model depends on a uniform theory of truth for mathematical statements and scientific ones. Such a model has proven recalcitrant.

So given that when we insert a mathematical example into the D-N model, we cannot retain some of the theses taken so seriously by the D-N model, except for those that end up as trivial, we claim that the D-N model will not be useful as a theory of mathematical explanation.

---

[23] See §2.1 above for a larger discussion of this.

2.3 Unification

The next serious theory of explanation in science we discuss is a unification theory. Unification models initially appear promising as theories of explanation because unification is seen as a desideratum in both science and in mathematics. Unifications display mergers between apparently disparate types of information under one single framework. Unifications satisfy a methodological need for Occamist parsimony, and the unification of knowledge is often taken to be the real motivation for the need for theories of explanation.

The theory in the philosophy of science that schemata for unification can be models of explanation seems like a good place to look for explanations in mathematics as well. Science maintains as one of its goals the reduction of the number of entities our theories need to countenance. For example, we unify strong and weak nuclear forces into one theory of nuclear interactions and we reduce the theory of thermodynamics to statistical mechanics. Many scientists see it as a goal to find a Grand Unified Theory, unifying all scientific statements under one account. The continued existence of *a number* of forces and fundamental types of matter is sometimes seen as an embarrassment to theoretical physics. Attempts at unification are also common in mathematics. For example, Descartes and de Moivre made significant contributions toward unifying geometry and algebra by showing how geometric notions are interpretable in the language of geometry. Tarski attempted to unify, in some way, logic and topology, Russell and Whitehead's *Principia Mathematica* attempted to unify logic and arithmetic by showing that the latter reduces to the former. The Langlands Program seeks to connect number theory with the representation theory of certain groups, and Conway and Norton's "monstrous moonshine" describes (the then surprising) connections between the Monster Group M and certain kinds of modular functions.

In a series of papers, Emily Grosholz, who has explicitly denied that there are mathematical explanations, gives us various accounts of unifications in mathematics. One can argue that like the sought-after unifications in science, there are epistemological gains and, on at least this account, explanatory gains, to be had from (reductive) unification in mathematics.[24]

Because of the variety of contexts in which unifications have been made, unifications allow for an air of ontological egalitarianism; they are easily understood to work in both mathematics and in science. If we have a theory of scientific explanation which mainly involves the unification of domains, then perhaps we can use the same theory in mathematics.

The mathematician W. W. Sawyer wrote that: "One of the most satisfying moments in mathematical history is the instant when it appears that two departments of mathematics until then regarded as separate and unconnected, are in fact disguised forms of one and the same thing." ([Saw55]: 59) He cites the example of the unification of algebra on the one hand with

---

[24] Emily Grosholz has studied many cases of unifications in the history of mathematics. See her [Gro80], [Gro81], [Gro85], [Gro92], and [Gro00]. See also [Stu97].

geometry and trigonometry on the other, the annexation of trigonometry by algebra, and the unification of many kinds of hypergeometric functions.

Nicolas Bourbaki ([Bou50]) has a version of what we may dub "naïve unificationism" in mathematics. They ask if mathematics is a unified discipline. (Is there a "mathematic" or "mathematics"?). They first reject the view that because it uses deduction throughout, mathematics is unified. Using deduction *is* mathematics, it is not a unifying theme *of* mathematics. They propose that mathematics is now unified because it uses the axiomatic method. The axiomatic method provides structure and much more. Because it provides various economies of thought, the axiomatic method serves to unite all of mathematics under one methodology. This view of unification proposes that there is a feature of mathematics that is common to the whole field. Even if we ignore how controversial the axiomatic approach is today, it will still not supply us with an explanatory schema. It merely tells us that all of mathematics is in theory supposed to be axiomatizable. Axioms however do not provide any explanations for the deductions that follow from them.[25]

Michael Friedman offered an early version of a unification model of explanation. Friedman's account of explanation[26] emphasizes that we gain

---

[25] Jeremy Avigad writes "The model of proof standardly used in mathematical logic today is that of formal axiomatic deduction. This formal notion is supposed to provide an explication of the informal notion of proof, one that explains the virtue by which an informal proof is judged to be correct, as well as what it means for a theorem to be a *deductive consequence* of some assumptions." ([Avi06]: 25-26) Clearly Avigad is not referring to the explanatory power of the axioms over the mathematics, but rather to the axiomatic method explaining *our judgement* of the correctness of the proof.

[26] Friedman's account of explanation is found in [Fri74].

scientific understanding by subsuming (or unifying) whole sets of particular instances or hypotheses under a law, or a set of laws under a more encompassing law. There are many examples of this kind of unification in science: Maxwell's unification of the laws governing electric, magnetic, and optical phenomena, for example.

Jamie Tappenden who himself is skeptical of whether or not these unifications can be described as explanations argues ([Tap05] §3), I believe successfully, that Friedman's model cannot be applied to mathematics. He reviews the standard "conjunction argument" against Friedman's approach. We can reduce the number of axioms in any theory generated by a finite number of axioms $A_1 \ldots A_n$, to just one axiom by just taking the conjunction of all the A_ns and forming the axiom $A_1 \& \ldots \& A_n$. But this "trick" represents no improvement in our understanding of the science and is clearly not explanatory. So simplicity cannot be considered only as a function of the number of axioms. The amount of information in each axiom is also something that is seen to be relevant. In light of this objection Friedman advances the notion of an "independently acceptable premise" in an attempt to refine the theory. But Friedman's distinction is not sufficient to counter many straightforward counterexamples. There are no recent attempts to defend Friedman's version of a unification model of scientific explanation.

Tappenden points out a second problem for Friedman's theory that has particular interest to mathematical explanation: "…reducing the number of axioms need not increase understanding, and increasing the number of axioms need not detract from understanding." So if increasing understanding is Friedman's goal, he is on shaky ground in so far as he is presenting a model of scientific explanation.

Tappenden also provides us with a few interesting examples from mathematics that illustrate the problems with Friedman's model. One example is Dedekind and Weber's unified theory of algebraic functions of one variable and algebraic numbers. This unification was mostly accomplished with the former having been seen as geometric and the latter as arithmetical. However, any advantage obtained from this unification is independent of the number of axioms in the unifying theory.

So the idea of unification as explanation is off to a weak start in the unification theories we have seen so far. We will next examine a more sophisticated theory of explanation that was proposed by Philip Kitcher. We will see that that model, despite its staunch defender in mathematics, is also problematic. We note too that many problems for Kitcher's model are also problems for Friedman's model.

2.3.1 Kitcher-style explanations

This section argues that Kitcher's theory of explanation, regardless of its merits as a model for scientific explanation, cannot be usefully applied as a model for explanation in mathematics. We begin by sketching Kitcher's model. We then show that if we allow the model to include mathematical explanations there are counterexamples. We present three such counterexamples from basic algebra, category theory, and logic. The first counterexample displays interesting similarities to aspects of information theory. It shows that Kitcher's theory, when applied to mathematics, resembles Gregory Chaitin's compressability criteria for nonrandomness. I argue that nonrandomness is not identical to explanation. The second counterexample, from category theory, also fits Kitcher's criterion though it does not meet Kitcher's goals for a successful explanation. Third, we briefly discuss Humphreys' mathematical counterexamples to Kitcher's theory of explanation.

Kitcher's theory of explanation[27] is an alternative to causal or deductivist strategies of explanation. It has come to be known as a "unificationist" approach. For unificationists, explanation is a matter of providing unified accounts of ranges of diverse phenomena. Scientific explanations are not to be evaluated piecemeal; rather whole families of explanations are evaluated together.

---

[27] Kitcher's account of explanation, and related discussions, in particular about the mathematical cases can be found in [Kit76], [Kit81b], [Kit81a], and [KS89].

Kitcher stresses that understanding is increased when we unify phenomena under patterns of reasoning. There are certain styles or patterns we utilize to account for all sorts of phenomena. When we find that diverse phenomena can be accounted for by a single method of reasoning, we have thereby united those phenomena in some way - specifically, we have offered explanations for them, and we understand them better as they now fit into familiar patterns. More precisely, Kitcher claims that explanation is a matter of repeatedly deriving descriptions of many phenomena from as few precise argument patterns as possible. The fewer patterns, the more precise they are, and the wider the range of phenomena they subsume, the more unified are our explanations. We increase understanding as we can combine various patterns of reasoning that act as explananda, and unify them under one pattern.

Kitcher's illustrative example is Darwin's theory of evolution which we think of as explanatory because it unifies our beliefs about various biological phenomena. In Kitcher's words, it provides "one (or more generally, a few) pattern(s) of argument which can be used in the derivation of a large number of sentences which we accept." Darwin himself gives explanation-sketches of the characteristics of any current species given the principle of natural selection and premises describing ancestral forms. Using this, Darwin claims, he can account for analogous variation in kindred species, facts of geographical distribution, the greater variability of specific characteristics, extinction, rudimentary organs, etc. ([Kit81b]: 172)

Because of Kitcher's views on the unity of science and mathematics ([Kit83]) he claims that his theory of explanation will not discriminate between cases of scientific explanation and cases of mathematical explanation. Kitcher's theory works in all domains equally.

Kitcher carefully outlines which sets of arguments ought to be accepted for explanatory purposes. We can sketch out the account as follows. Let P stand for a collection of (descriptions of) phenomena. Let S be a collection of argument schemata, each made up of schematic premises and conclusions containing variables which when instantiated yield valid arguments that are instances of the original schema. (So in an argument schema of the kind Darwin had in mind, we would expect that all or many biological phenomena be deduced from one schema.) We can outline the model of explanation (simplifying slightly) as follows:

> Definition: S is a *systematization* of P if some members of P are derivable from S.
>
> Definition: $S\prime$ is a *better systematization* than $S\prime\prime$ if either a) n members of P are derivable from $S\prime$ and m members are derivable from $S\prime\prime$ and $n > m$, or b) They both derive the same number of elements of P but $S\prime$ has fewer members than $S\prime\prime$.
>
> Note: (2a) and (2b) may have to be weighed against each other so that any $S\prime$ with the same number of distinct patterns and that implies the same accepted statements as P contains argument patterns that are less "stringent" than the patterns in $S\prime\prime$. (Stringency is a largely syntactic notion having to do with

> the number of non-logical symbols in a symbolic formula that may be permissibly substituted for.)
>
> Definition: S *unifies* P if S is the best available systematization of P.
>
> Definition: Given some unification of a set of phenomena, P, by S, where E is an instance of a member of S, E *explains* R if R is a member of P and R can be derived from E.

As Robert Klee puts it, this formulation is all in service of enabling us to "*count* argument patterns, accepted truths, and substitution places in symbolic formulae." The goal is to "[e]xplain as much of the unfamiliar using as little of the familiar as you possibly can get away with."([Kle97]: 119)

In the course of elaborating the unificationist account of explanation, Kitcher lists a number of examples from mathematics (though he gives no details) ([KS89]: 424-426). The first example is the claim that Bolzano's proof for the intermediate zero theorem needs to be a "properly grounded proof" and that this need for a properly grounded proof as opposed to a "regular" proof emerged out of explanatory considerations.[28] Kitcher's sec-

---

[28] A few historical caveats, the first about Kitcher and the second about Bolzano: (1) considering the dates of publication of the respective papers, Kitcher seems to have used his analysis of Bolzano as his model for explanation in general. So considerations in mathematical explanation seemed to be driving Kitcher's views on Scientific explanation. (2) Bolzano is generally taken to be a foundationalist (see his [Bol10]: part 1, esp. §8) and his work is taken to be a precursor of Frege's logicism. But his [Bol10] and [Bol72] clearly spell out a special system of proof that appeals to what he calls the *Abfolge* relationship. The *Abfolge* relation, unlike the deducability relation, holds only between true propositions. These relations serves to "ground" the "deduced" propositions in the propositions that they are deduced from. This distinction parallels an Aristotelian distinction between a demonstration and a *reasoned* demonstration. Bolzano further utilizes this distinction in [Bol04] and most significantly in discovering his proof of the intermediate zero theorem in [Bol17]. Kitcher's analysis of Bolzano's *Abfolge* relationship, and his claim that Bolzano

ond example involves the question of the acceptance of axioms of finite group theory on the grounds that one set of axioms is more "natural" than the others. Kitcher claims that "more natural" in this context is more explanatory, and that the more explanatory axioms were selected.

The third example invokes Lagrange's and Galois' investigations in to the explanation of the unsolvability of the quintic equation.[29] In 1821 Niels Henrik Abel proved that the quintic equation cannot be solved by algebraic means. As Ian Stewart puts it: Abel's proof "was rather mysterious and indirect. It proved that no general solution is possible, but it did not really explain *why*." ([Ste07]: xi) Galois ultimately discovered that the quintic was unsolvable because the equation did not have the right kind of symmetries. Presumably what Kitcher had in mind is that the problem of the quintic was mysterious even after Abel showed that it could not be solved. Galois demystified the problem by unifying the problem of the quintic under a rubric that included the unsolvability of any equation that displayed the wrong kind of symmetries.

In an exposition of Kitcher's theory of explanation in science, Edwin Hung gives a simple mathematical example of what he claims is Kitcher's idea "that unification seems to be the application of a general idea or formula to a variety of instances." ([Hun97]: 188) Hung asks us to consider the following

---

had explanatory considerations in mind when using this distinction is in [Kit75]. Mancosu ([Man99a]) addresses this example. See also our discussion of this in §3.3.)

[29] A quintic equation is a polynomial equation of degree five. It is generally of the form $ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$, where a, b, c, d, e, f are members of a field and $a \neq 0$.

sequence of numbers:

$$0, 3, 8, 15, 24, 35, 48, 63, 80, \ldots \tag{2.1}$$

At first glance they appear like numbers that ascend with random gaps between them. Is there some unifying principle under which they can be subsumed so that the gaps appear less arbitrary or less puzzling? The answer is yes, there is. Take the formula

$$m = (n^2 - 1), \; \textit{where } n \textit{ is an integer.} \tag{2.2}$$

We can see that (2.1) can be generated when we substitute the integers $1, 2, 3, \ldots\ n$. Performing this substitution eliminates the puzzlement about the sequence. So the formula unifies the sequence. Note that it unifies the sequence as a whole, not just some particular member of the sequence.

But from an explanation-seeking point of view, this analysis seems inadequate. The sequence (2.1) is generated by the formula (2.2), and it generates the *entire* sequence, and it is the smallest of all the rival formulae that anyone can think of that can generate that sequence in that number system . . . But is it explanatory? Kitcher must say yes. But it seems hard to justify the fact that a sequence is explained merely in virtue of having a formula of a certain kind that can generate it.

Perhaps if this were an empirical investigation, this would sound more plausible as an instance of explanation. If we found a set of empirical data-points that corresponded to the sequence (2.1) and speculated that there was a natural law that covered the phenomenon of the form the sequence (2.2), we might be justified in saying that the law explained the phenomenon. However as a relationship between formula and sequence, it is hard to understand how this is an explanation.

Beyond our intuition that (2.2) does not explain (2.1), perhaps a brief foray into information theory will shed some light on why we have the intuition that the formula (2.2) does not explain.

We notice that there is an interesting relationship between Kitcher's theory of explanation and Gregory Chaitin's notion of compressability, at least in some mathematical cases. Chaitin ([Cha75]) would approach our sequence by stating that the sequence (2.1) is not random because it is capable of being compressed into a shorter sentence, the sequence (2.2), that expresses the same information. Had (2.1) been random it would be incompressable, as the shortest possible expression of a random sequence is the sequence itself. And if random sequence S could only be explained by S, we then have a model of explanation which allows S to be explained by S, which is unacceptable. Or, we would have to claim that random sequences, in virtue of the fact that they could only be explained by themselves, have no legitimate explanations. But this is ad hoc. We would want an argument why some sequences have explanations and some do not.

(2.2) thereby acts as a brief description of (2.1). It is actually the shortest description we know of. But descriptions, as we know, are not explanations. Historically, one of the central tasks of the study of scientific explanation is to distinguish explanation from description. Explanation must be something over and above description, because descriptions state the facts and explanations account for those facts. Asserting that there are clouds and there is rain may describe a situation, but without a discussion of the relationship between the two, we do not have an explanation. So giving a description, like we do in the mathematical case, is not the same as giving an explanation.

In general, one can relate noncompressability and randomness with explanation to see why we would treat mathematical sequences differently than we would treat a sequence of scientific data-points. Given a set of data-points about some natural phenomena, one way of saying that the data is inexplicable is by claiming that there is no law that can subsume all of them - that is to say that they are random. If that is the case, then there is no way to predict what the next data-point in that series will be. You must simply wait for (or find) the next datum. Conversely, if the data-points are nonrandom, the next data-point is in principle predictable. The algorithm that predicts the next data-point can be said to explain or factor into an explanation of the set of data-points.

When considering mathematics we do not have this intuition because (to the extent that we use the word) "prediction" in mathematics is a misnomer. We can continue to generate the set of numbers, but that is not predicting them.[30] The prediction/explanation symmetry that Hempel asserts is important here. Hempel claimed that prediction and explanation are two sides of the same coin. In the case of a set of scientific data-points, say those we get in measuring radioactive decay, it is plausible to say that the law that can be formulated as a mathematical formula that describes the rate of decay explains the decay because it allows us to predict when the sample will lose its next subatomic particle or its average rate of decay. We can wait for the next instance of decay and observe it. In the mathematical case we do not make such predictions. We do not literally predict what the next number in the sequence (2.1) is. To say that the next number is 99 is not to have made a successful prediction. To say that it *will be* 99 is nonsensical. It *is* [timelessly] 99. Even if you claim that mathematics is invented (as opposed to discovered) it is hard to see how one can argue that the next number in the sequence isn't 99 until someone invents it. So if Kitcher is correct and prediction is related to explanation, we have another reason to reject Kitcher's notion of explanation in mathematics.

If we were playing a game, and a computer was slowly displaying numbers on a screen and we saw 0, then 3, then 8, . . . , then 63, then 80 and we were

---

[30] As an aside we note the famous line attributed to Paul Dirac that "I consider that I understand an equation when I can predict the properties of its solutions, without actually solving it." (Quoted in Frank Wilczek and Betsy Devine. *Longing for the Harmonies: themes and variations from modern physics*. New York: Norton; p 102.) Dirac could not literally be thinking of predicting solutions that would come to be. Dirac was considering whether he could get part of the solution before *he* solves the equation in detail. He was certainly not considering the question of whether he could solve part of the answer before the rest of the answer came to be [true].

asked to predict what number the computer will display next, we might say 99. This would probably be an accurate prediction, but it would be a prediction about some artifact, specifically it will predict what the machine will display next. It is not a prediction about the mathematics. The mathematics might provide the reason for the prediction about the machine, but it is not a prediction about the mathematics itself, any more than a prediction about the weather is about the algorithm that does the predicting. Predicting rain is about what the clouds will do given the atmospheric conditions, it is not about the mathematics used to describe the clouds.

If you want to argue that explanation (for Kitcher) *is* compressability, you must then show that the compressed version of what we have ultimately serves an explanatory role, e.g., it increases understanding or, more importantly, it better *accounts for the phenomena.* Clearly there are aesthetic and technological advantages to compression, but not explanatory advantages.

Given these considerations about prediction, compression and randomness it seems clear that it is not correct to say that every mathematical case that meets Kitcher's requirement is explanatory. Moreover, there are plenty of non-random facts, both scientific and mathematical, that no one would consider explanatory. Accounting for a phenomenon is different from accounting for the fact that a phenomenon is not random. Kitcher's criteria seem to lack the ability to distinguish between these in many cases. So we have our first counterexample to Kitcher's theory of explanation.

Our next counterexample to Kitcher's theory of explanation is from category theory. Consider the following: we can represent "the" Cartesian product of sets S and T as follows: For every set $S = \{a, b, c\}$ and $T = \{x, y\}$, via the pairing axiom, we get

C 1 $S \times T = \{(a, x), (b, x), (c, x), (a, y), (b, y), (c, y)\}$.

Of course, we can also represent the Cartesian product $S \times T$ as

C 2 $S \times T = \{[a, x], [b, x], [c, x], [a, y], [b, y], [c, y]\}$

or

C 3 $S \times T = \{\{\{a\}, \{a, x\}\}, \{\{b\}, \{b, x\}\}, \{\{c\}, \{c, x\}\}, \{\{a\}, \{a, y\}\}, \{\{b\}, \{b, y\}\}, \{\{c\}, \{c, y\}\}\}$

...where in each case there are mappings $(a, x) \longmapsto \{\{a\}, \{a, x\}\}$ ...such that **C1** and **C2**, **...** are isomorphic. Whereas these are isomorphic, they are not identical. This raises the question "what is *the* Cartesian product of $S \times T$?"

Category theory however allows us to stay neutral with regard to the semantics of the objects we are working with and can thus represent *a* Cartesian product as follows: Given S and T, a Cartesian product of S and T is an object C and two maps:

$$S \xleftarrow{\pi_1} C \xrightarrow{\pi_2} T \tag{2.3}$$

such that if there is any other object, D, mapping to S and T then there is a unique map from D to C making the following triangles commute:

$$\begin{tikzcd}
S & D \arrow[l, "f"'] \arrow[r, "g"] \arrow[d, dashed] & T \\
 & C \arrow[ul, "\pi_1"] \arrow[ur, "\pi_2"'] &
\end{tikzcd}$$

Now,

> *If C and C' both satisfy the previous requirement for S and T then there is an isomorphism from C to C':*

$$\begin{tikzcd}
 & C \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] \arrow[dd, dashed, bend right, "h"'] & \\
S & & T \\
 & C' \arrow[ul, "\pi_1^1"] \arrow[ur, "\pi_2^1"'] \arrow[uu, dashed, bend right, "g"'] &
\end{tikzcd}$$

By "isomorphism" we mean that there is a map $h : C \longrightarrow C'$ and a map $g : C' \longrightarrow C$ such that $h \circ g = Id$ and $g \circ h = Id$. These maps must be equal to the Id because of the uniqueness of the definition of the Cartesian product.

So each of the sets given in **C1**, **C2**, and **C3** are isomorphisms.

Given this, we can prove the theorem that the Cartesian product is associative up to isomorphism, or

$$S \times (T \times W) \cong (S \times T) \times W. \tag{2.4}$$

The proof is as follows: Since there is a map $\pi_1^{S\times(T\times W)}$ from $S \times (T \times W)$ to S and $\pi_2^{S\times(T\times W)}$ from $S \times (T \times W)$ to $(T \times W)$ and there are two maps from $(T \times W)$, $\pi_1^{T\times W}$ to T and $\pi_2^{T\times W}$ to W, and there are mappings from $S\times(T\times W)$ to T and to W. There are two maps from $(S\times T)\times W$; $\pi_1^{(S\times T)\times W}$ to $(S\times T)$ and $\pi_2^{(S\times T)\times W}$ to W, and $(S\times T)$ has a map $\pi_1^{S\times T}$, to S and another map $\pi_2^{S\times T}$ to T. So, since there are two mappings from $(T \times W)$, one to T and one to W, there is a unique map (1) from $(S \times T) \times W$ to $T \times W$.

Since there are two mappings from $S \times (T \times W)$, one to S and another to W, and two maps from $(S \times T) \times W$, one to S and another to W, then there is a unique map (2) from $(S \times T) \times W$ to $S \times (T \times W)$.

Since there are two mappings from $S \times (T \times W)$, one to S and another to T, and there is a mapping from $S \times T$ to S and another from $S \times T$ to T, then there is a unique mapping (3) from $S \times (T \times W)$ to $S \times T$.

Finally, since there are two maps from $S\times(T\times W)$, one to S and another to W and there are two maps from $(S \times T) \times W$, one to S and another to W, then there is a unique map (4) from $S \times (T \times W)$ to $(S \times T) \times W$. And since there are unique maps (2) and (4), $S \times (T \times W) \cong (S \times T) \times W$. QED.

$$\begin{tikzcd}
 & & S \times (T \times W) \arrow[dll, "\pi_1^{S\times(T\times W)}"'] \arrow[dr, "\pi_2^{S\times(T\times W)}"] \arrow[dddl, dashed, "(3)"] \arrow[dddd, dashed, bend right, "(4)"'] & & \\
S & & & T \times W \arrow[dl, "\pi_1^{T\times W}"'] \arrow[dr, "\pi_2^{T\times W}"] & \\
 & & T & & W \\
 & S \times T \arrow[uul, "\pi_1^{S\times T}"] \arrow[ur, "\pi_2^{S\times T}"'] & & & \\
 & & (S \times T) \times W \arrow[ul, "\pi_1^{(S\times T)\times W}"] \arrow[uurr, "\pi_2^{(S\times T)\times W}"'] \arrow[uuur, dashed, "(1)"] \arrow[uuuu, dashed, bend right, "(2)"'] & &
\end{tikzcd}$$

We have shown that the Cartesian product is associative up to isomorphism. What is important here, however, is to note that we did not specify what S, T, or W stands for. As a matter of fact, they can stand for sets, Hopf algebras, Lie groups, rings, topological spaces, metric spaces, or a myriad of other mathematical objects. For all these different objects we can specify what the maps are. For example, for topological spaces, we will assume that all the maps, including the πs are continuous maps. So we have shown that the Cartesian product for topological spaces is associative.

The proof gives us a wide variety of theorems proved using only one argument pattern. We have associativity proofs for a wide variety of mathematical objects subsumed under one proof. It provides an extremely compact framework from which we can derive the greatest number of these kinds of results. Thus category theory ought to explain associativity. In Kitcher's terminology, the proof above is a best unification of all the argument patterns that imply the associativity of all the mathematical objects we listed.

Yet it is hard to see how our understanding of associativity is increased by the above proof. It is also very hard to say that we have now given an explanation of anything. What did this "explanatory proof" explain? It is difficult to imagine how anything as semantically inert as category theory can explain anything. Or more precisely, how do we get intuitively explanatory proofs about sets or spaces from this proof? Who would say that after seeing this they now understand associativity of Cartesian products of metric spaces, or that *this* is how you *really* account for the associativity of the Cartesian product of Lie groups. So in virtue of what is deducing associativity from the fact that $S \times T$ maps to S and T explanatory?

Paolo Mancosu quotes the mathematician Solomon Feferman, specifically referring to category theorists to the effect that "many category theorists find a mathematical result intelligible only when it is formulated within the framework of category theory." ([Man00]: 106) The suggestion (especially in this context) is that *this* is the only way *they* understand whereas other mathematicians do not. Perhaps understanding varies with the particular mathematician, and each set of mathematicians has a different method for understanding mathematics. In any case if Feferman is suggesting that category theorists only understand mathematical results expressed in category-theoretic language, then there are no explanations for them from other disciplines - a very counterintuitive suggestion especially since none of the literature on mathematical explanation uses examples from category theory, and all the proponents of mathematical explanation claim or imply that there were explanations prior to the advent of category theory. More likely Feferman is suggesting that explanation is relative to explainee - in this case category theorists. But this violates our explicit condition (set out in Chapter 1) that for an explanation to be philosophically interesting it cannot only be an explanation relative to some explainee, it must be objective. It must account for the phenomena regardless of who the explanation is being recounted to. The whole point of a model or theory of explanation is to avoid that problem.

Furthermore, after seeing this proof we still have no clue as to *how* algebra, topology, set theory, . . . are united, or *why*. We have a proof. We have eliminated the need for diverse proofs and proof procedures for a variety of fields, but what has been explained? How is our understanding of any of these fields increased?

Moreover, here the abstract nature of the theory works to its detriment: a more abstract theory in mathematics can serve as a hindrance to understanding. The more abstract a theory is, the less obvious it is what the theory has to do with the actual "phenomena", in our case the associativity, or what the phenomena have to do with each other (e.g., the associativity of the Cartesian product of topological spaces and the associativity of the Cartesian Product of Hopf algebras). In this case, the more abstract category-theoretic proof seems completely unrelated to the question of topological spaces, yet it can serve as an easy proof of the associativity of product spaces.

Moreover, and perhaps most importantly for Kitcher, mathematical practice has not treated category theory as an explanatory framework within which it will handle associativity proofs. This is a particularly significant difficulty for Kitcher because much of the motivation for his general philosophy of mathematics is that it is in conformity with mathematical practice. We return to this point at the end of the section. But first we have additional counterexamples.

Paul Humphreys gives two related counterexamples to Kitcher's theory of explanation. Humphreys ([Hum93]) challenges Kitcher's contention that we get greater understanding out of Kitcher's (and Friedman's ([Fri74]) unification theory. Humphreys recognizes and endorses the fact that unification is involved in understanding and explanation. He does however criticize the ability of Kitcher's formal model to handle mathematical cases. Kitcher claims that explanations that conform to his model increase our understanding. Humphreys' first example exploits the stringency condition (3) in Kitcher's model of explanation. He shows that there can be two axiom schemas for propositional logic whose only relevant difference is that one has greater unifying power (as per Kitcher's definition) yet seems no more (or less) explanatory, and certainly provides no more (or less) insight or understanding into propositional logic than the other.

Humphreys' second example is more striking. It uses the axiom proposed by Jean Nicod as the sole axiom needed for a complete axiomitization of propositional logic. Using a single connective, the Scheffer stroke (|), and a single rule of inference (from $A|(B|C)$ and A, C follows) and the following axiom:

$$(A|[B|C])|[D|(D|D)]|[(E|B)|((A|E)|(A|E))] \tag{2.5}$$

we can derive all the theorems in propositional logic, allowing us to unify quite a bit from one axiom and one rule.[31] It also meets all of Kitcher's criteria for a good explanation, but it does not produce any gain in understanding over rival axiomatizations which are (slightly) less parsimonious. To the contrary. This one axiom and its rule of inference is more difficult to understand than its less parsimonious rivals. Moreover, such extreme reductions in logic have had little influence on mathematical practice (see also [San97]: 24), which would suggest that it has not been viewed by mathematicians as having explanatory import. Nicod's axiom remains a clever

---

[31] Similar examples exist for group theory. See `ftp://info.mcs.anl.gov/pub/tech_reports/reports/P901.pdf`. Accessed on 30 April 2007.

curiosity.

The final issue we raise about Kitcher's theory, is that there are a number of general technical challenges that the model faces. Some of these difficulties involve problems of weighing the various criteria and choosing the best argument pattern - e.g., we pit the fecundity of a systematization against the economy of argument patterns, or we can consider the tension between the two stringency criteria ((2a) and (2b)). However, if we are in principle unable to decide, when faced with rival options, which set is the best generating set for a systematization, then the account is unusable.

Taking a suggestion from Benacerraf's Ernie and Johnny story about the nature of numbers, we note the following problem. Given some conception of mathematics which countenances sets as foundational, Kitcher's account is at odds with an accepted constraint on explanation. Take rival conceptions of sets, say those illustrated by **C1**, **C2**, and **C3**, which can each stand as a foundation for all (or even some part) of mathematics. Presumably, each set theory, given the same set of inference rules, can generate the same mathematics. The difficulty then for Kitcher is that we are faced with an infinite number of equally good explanations for the same phenomena - an untenable position for a theory of explanation to be in.

This argument holds regardless of whether or not you hold that mathematics has foundations, or what you think those foundations might be. As long as there can be rival isomorphic axiomatizations within mathematics, we face the problem of the impossibility of ordering them well. This is unacceptable for any explanatory schema.

Though Kitcher's theory seems to cover a large number of cases that are explanatory, it has been criticized as also subsuming cases that are not clearly explanatory. Many of the scientific cases are explanatory. Accounting for all sorts of biological and economic phenomena by invoking game theory certainly seems explanatory. However, it is not at all clear that representing the development of a physical system as a path in phase space, and expressing physical laws as constraints on paths in phase space, is really explanatory. But Kitcher must count it as explanatory. Many people see chaos dynamics as unifying water wheels, weather, and predation patterns. It is certainly arguable that the similarly chaotic paths of these systems in their associated phase spaces is not why they behave as they do. But Kitcher has to count it all as explanatory.

The mathematical cases are all explanatory using Kitcher's theory but do not seem to meet our intuitions about what should and should not count as explanations. So Kitcher can hardly have shown that there is such a thing as mathematical explanation when the mathematical cases are all so like the countercases.

That is our case against applying Kitcher's model of explanation to mathematics. As a matter of mathematical practice Lawvere's early category theoretic proof uniting diagonalization arguments ([Law69]) was mostly ignored for quite some time. Colin McLarty cites the mathematician R. Goldblatt displaying a not atypical attitude[32] toward category theory:

> In explaining 'the style I [Goldblatt] have adopted' [Goldblatt] deplores modern mathematical writing which gives abstract definitions before it '*reveals the original motivation*,' so that 'the student is not actually *shown* the genesis of concepts - how and why they evolved - and is thereby taught nothing about mechanisms of creative thinking.' 'All of this,' [Goldblatt] says, 'seems to me particularly dangerous in the case of category theory, *a discipline that has more than once been referred to as "abstract nonsense"*' ([McL90]: 352. First and third emphases added.).

David Sandborg cites an unpublished manuscript[33] of J. P. Marquis (who has written extensively on philosophical ramifications of category theory), which would seem to support the main argument and primary example in this chapter. I quote Sandborg here in full:

> Marquis doubts the existence of mathematical explanations in "Are there Explanations in Mathematics." Like Resnik and Kushner and following van Fraassen, he takes as his starting point the hypothesis that explanations are answers to why-questions. He suggests that mathematicians do not actually answer why-questions. When they describe themselves as explaining, they instead answer "how possible" questions. His analysis suggests that proofs in category theory should, by the standards of Steiner's or

---

[32] The mathematician Miles Reid likewise parenthetically remarks that the study of category theory for its own sake is "surely the most sterile of all intellectual pursuits" (quoted in [Cor03]: 228).

[33] I have been unable to obtain this.

> Kitcher's theories, be paradigmatic of explanatory proofs. However, such proofs do not appear to be particularly explanatory. Though he is skeptical of the existence of mathematical explanations in the sense of answering why-questions, he is hopeful that progress can be made in understanding mathematical understanding. ([San97]: 28-29)

I take it that I have found an example which Marquis would support.

We have seen in this section that Kitcher's theory of explanation is susceptible to a number of mathematical counterexamples. We have shown using standard algebra, category theory and logic that there are cases where a clear unification of domains is achieved, or there are cases which unify much information under a single argument pattern, yet there is no increase in explanation. It is for these reasons that we claim that Kitcher's model of explanation cannot stand as a model for mathematical explanation.

2.4 Why-question accounts

In this section I examine the why-question account of explanation developed by Bas van Fraassen and its applicability to mathematics. I first outline van Fraassen's theory of explanation, and then review David Sandborg's assessment of the why-question approach for mathematics. Sandborg's example is presented and criticized. Then we examine Sandborg's arguments against the why-question account and present our own criticisms. Finally, we use this chapter to address a related question about the methodology employed by Sandborg and others in the theory of mathematical explanations.

Introduction

In the course of rejecting Steiner's model[34] of mathematical explanation, Resnik and Kushner mention a why-question approach to mathematical explanation. They claim that thinking of explanation in terms of why-questions better captures the intuition behind their own thoughts about how mathematics is explained ([RK87]: 153). David Sandborg ([San98]) offers a number of criticisms of the why-question account of explanation as it might be applied to mathematics. To date, Sandborg's is the first and only critical examination of the application of a why-question account to mathematics.

The why-question account is sometimes called a contextualist account of explanation, or an erotetic account. Unlike [other] epistemic accounts of explanation, the why-question account does not view explanations as objective

---

[34] See §3.1 below.

relations that hold between statements about facts in the world, or that explanations are independent of the contexts in which they are requested and provided. The why-question account factors in the pragmatics of language as much as it does the syntax and semantics. The pragmatics of language reflect the practical circumstances in which we put language to use, that is they reflect the conditions and context that make some statements appropriate or meaningful. This contrasts with the accounts of explanation that only require an evaluation of the syntax and semantics of the relata in question. The syntax and semantics are respectively the linguistic rules that determine whether or not a string of words are grammatical or not, and what their meanings are. If an explanation has a pragmatic element, then we can only tell whether or not it succeeded if we are given an account of the human context in addition to an account of the syntax and semantics. So on this account explanations are ternary relations holding between theories, facts, *and contexts*, as opposed to what van Fraassen takes to be the "received view" where an explanation is a binary relation between theory and fact. We discuss this more shortly.

Van Fraassen cites a number of reasons to favor his approach to explanation including the inability of other models of explanation to account for rejections of explanation requests, and the failure of other models to handle famous counterexamples like the asymmetry problem. His model, he claims, does not face these difficulties.

A second reason for adopting a why-question account of explanation has less to do with the details of the model and more to do with the metaphysical scientific anti-realism which the model easily accommodates. The why-question model was designed around van Fraassen's own metaphysical Constructive Empiricism spelled out in his [Fra80]. Constructive Empiricism is a theory that opposes both realism and positivism. It claims that contemporary scientific realism is predicated on a basic misunderstanding of the nature of explanation. On this view, explanations are interpreted as answers to whatever question a particular speaker is interested in at some particular time, given a particular set of knowledge and interest. This is in contrast with the positivist conception of explanation as objective relations which hold for all people at all times. A theory of explanation for scientific realists would make sense with a theory of explanation that has an explanation relating the real objects to be explained.

A theory of explanation designed with constructive empiricism in mind sounds like an unlikely place to look for a theory of mathematical explanation. Constructive empiricism generally denies the existence of unobservables, presumably including such abstracta as mathematical objects. Though its detractors argue against constructive empiricism on the grounds that there is no clear line between observable and unobservable, *any line* between observables and unobservables would put mathematical objects on the unobservable side.[35] And if numbers and other unobservables are not taken to exist, then certainly they are not in need of explaining on this view. Thus the why-question approach initially seems to be an inappropriate place to look for a theory of mathematical explanation.

Nonetheless despite the fact that no philosopher has explicitly advocated on behalf of a why-question approach to mathematical explanation, there are a number of reasons one may want to do so. For one, a why-question account should be able to handle any area of human inquiry in which we can phrase a why-question. We certainly have many questions about mathematics that can be phrased as why-questions. For example upon learning about the logics between classical and intuitionistic, one may be tempted to ask "Why are there exactly seven intermediate logics for which Craig's Interpolation Theorem holds?"

Van Fraassen draws no distinctions between different domains in which explanations are given. "To ask that their explanations be scientific is only to demand that they rely on scientific theories and experimentation, not old wives' tales." (120) There is no obvious reason to think that van Fraassen would not extend this to mathematics. Mathematical explanations are ones that rely on mathematical structures and methods, and not science or mysticism.

---

[35] Penelope Maddy's view ([Mad90]) that sets are observable notwithstanding. Her view is only for sets though, and not other mathematical entities which are thus far considered unobservable by almost all philosophers.

So there is reason to believe that the why-question approach will not yield a model for mathematics. But intuitively one might think that it is a simple matter to fit mathematics into the why-question account. However, when we try, the account fails for a number of reasons. We will show first that various parts of the why-question account become trivialized when we plug in mathematical topic sentences. And secondly we will show that the why-question account does not contribute anything to mathematics or the projects of mathematicians.

The Why-Question Account

Unlike other accounts, the why-question account does not assume that there are any explanations *per se*.[36] Rather, explanations are answers to why-questions. Why questions are questions of the form "why x?" And since explanations are answers, they are successful explanations in so far as they answer their respective why-questions. There are three components to an analysis of the why-question: (1) A topic, (2) a contrast class, and (3) a relevance relation. The topic is what the question is about. So when one asks "why P", the topic is P - the subject of the question. The "contrast class" is everything else the question could be about. So X is a set of propositions $\{P_1, P_2, P_3, \dots P_n\}$, where the topic is some P_k in X and the contrast class is

---

[36] Hannson ([Han06]), Garfinkel ([Gar81]) and others address and support why-question accounts of explanation. The why-question account described here follows Chapter 5 of van Fraassen's [Fra80]. Van Fraassen does not offer the only plausible why-question account, but he does offer the most developed and widely-discussed one. Our discussion should apply to any similar non-trivial why-question account. (This will be explained more in this section.) The exposition also relies heavily on the discussion in [San98].

all the other P_is that are alternatives to P that the questioner might consider. The contrast class clarifies the fact that when one asks "why P?", what he or she is really asking is "why P_k and not some other P?" The relevance relation, R, is a binary relation that relates the topic and the contrast class to an answer that the questioner allows by making explicit the contextual feature. The question is thus be denoted by $Q = < P_k, X, R >$.

Take the following example adapted from van Fraassen ([Fra80]: 127). Consider the following question:

(1) Why did Alice drink the wine?

There are at least three ways this question can be interpreted.

(2) Why did *Alice* drink the wine?
(3) Why did Alice *drink* the wine?
(4) Why did Alice drink the *wine*?

(2), (3), and (4) are three different questions in that they anticipate different answers. They do so because each of those questions has a different contrast class. For example, the contrast class for (2) may contain Bob and Charlie. That is, the question may be asking

(2′) why did Alice, and not Bob, drink the wine?

or

(2′′) Why did Alice, and not Charlie, drink the wine?

Similarly for (3) whose contrast class may include throwing the wine in her boyfriend's face, pouring it as a libation to her god, or resisting the temptation to drink. The contrast class for (4) may include grape juice, cola, water, or whiskey. But despite differing contrast classes, (2), (3), and (4) all have the same form and content because they all have the same *topic*, which is that "Alice drank the wine".

The *context* serves to disambiguate sentences of type (1). (1) can be disambiguated by asking a question of the form (2′) or (2′′) - wherein we explicitly mention the contrast class. In cases where the members of the contrast class are not explicitly mentioned, we rely on the context to disambiguate the sentence. For example (1) might be part of a larger text. Consider

> (5) Alice just left the Alcoholics Anonymous meeting. She seemed quite determined to give up drinking alcohol for good. So, why did Alice drink the wine?

The context could be *understood* and not explicitly stated in the discourse or conversation. If two devout Muslims (who eschew alcohol) are having a conversation about Alice, also known to them as a devout Muslim, then (1) might mean:

> (3′) Why did Alice drink the wine instead of resisting the temptation to drink alcohol altogether?

Further contextual information is provided by the relevance relation, R. R relates the question to the answer, and constrains admissible answers by specifying what counts as explanatorily relevant for Q. The relevance relation constrains admissible answers by distinguishing between different senses of the question. In most cases, it is unlikely that an answer to (1) will be acceptable merely by appealing to the human-bio-mechanics of swallowing accompanied by a discourse on the fluid dynamics of wine. R would rule out such an answer so that it will be clear that the questioner is looking for a social or psychological answer and not a biological or physical one. R may specify for example that the only answers that a questioner finds applicable must appeal to Alice's religious beliefs. In response to a question about why some particular species is monogamous for example, one can appeal to the brain chemistry of the species (if that is indeed what causes members of the species to "mate for life") or one can give an answer that appeals to the evolutionary advantages of monogamy for this particular species. R will specify which type of answer is appropriate. The relation thus binds the topic/contrast-class pairs to proposed answers. The theory of why-questions thus serves to take the ordinary (ambiguous) natural-language why-question and translate it into a (precise) logical why-question. It makes the question precise by using contextual features - mostly contained in the relevance relation and the contrast class.

Now that we have the logical why-question, we can see how the theory provides for the evaluation of the answer. There are two parts. First, given some why-question $Q =< P_k, X, R >$ we analyze the following to get a *direct answer*:

$$P_k \text{ in contrast to the rest of } X, \text{ because } A \tag{2.6}$$

where A is the reason for, or *core* of the answer to, Q. In such an answer the following four things are true: (1) P_k is true, (2) All other elements of X are not true, (3) A is true, and (4) A bears R to $< P_k >$ ([Fra80]: 144). Direct answers serve to distinguish answers from non-answers and can distinguish between say, the brain chemistry and functional answers in our monogamy query by revealing the question's presuppositions. A why-question presupposes its topic is true, as is at least one proposition bearing the relevance relation to the topic and contrast class. If a why-question has a false presupposition, the answer will reflect this, perhaps by offering a correction, e.g., "P_k is not true" instead of a direct answer, because no direct answer will be true.

There is also a second part to the evaluation of the answer. Even if the above four conditions for a direct answer are met, the answer may still not be *telling*. That is, while direct answers will tease out the right *kind* of answer, it does not guarantee the correct answer.[37] There may be rival direct answers that are superior on factual grounds. To get an answer that is also telling, given our background knowledge K, and a restricted subset of K, $K(Q)$,

---

[37] We note that e.g., the famous syphilis-paresis case may have no telling answers on this account.

the answer must (1) be probable in light of K, (2) probabilistically favor the topic over other members of the contrast class relative to $K(Q)$, and (3) be comparatively better in these regards than other potential answers to Q. This will generate a telling answer.

There are a number of problems with this assessment of telling answers. For our purposes (1) is unproblematic, and the problems with (3) do not concern us. But why must a telling answer probabilistically favor the topic over the rest of X relative to $K(Q)$ and not all of K? And how do we select the content of $K(Q)$ from K?

The reason why an answer must appeal to $K(Q)$ is that a why-question (and its asker) presuppose that the topic is true. So given some question "why B and not some $C \dots N$?": (a) K will imply B, and (b) K will also imply that $C \dots N$ are not true. But exactly (a) and (b) are irrelevant to how favorable the answer is to the topic. Whatever knowledge we used to get the fact that B is true and $C \dots N$ are not, is exactly not what we are asking for when we want an explanation for why B and not $C \dots N$. We want an answer that will have a probabilistically favorable impact on the topic given what we know, and not one that just repeats what we know or assumed. So the evaluation of the answer uses the $K(Q)$ that constitutes the general theory about the phenomena and other known facts that do not imply B. Van Fraassen is somewhat vague about how we choose $K(Q)$ saying only that "it must be a further contextual factor" ([Fra80]: 147).

So to sum up, the theory forces us to spell out the logically precise why-question and its evaluation. The context determines what question is really being asked, and the precise question will determine how the answer is evaluated. The relevance relation depends on the context of the question, and the theory of telling answers depends on the contrast class and the $K(Q)$. That is how explanatory evaluations are necessarily context dependent on this theory of explanation.

Before we explore the theory of why-questions further, let us make two brief digressions into David Sandborg's work on this topic, and then we return with an assessment of the why-question approach for mathematics.

Sandborg's discussion - his example

Sandborg begins his discussion of the why-question approach ([San98][38]) with an example of what he takes to be a mathematical explanation. This example is taken from George Polya. Polya wants to prove the following theorem: *If the terms of the sequence* $a_1, a_2, a_3 \ldots$ *are non-negative real numbers, not all equal to 0, then*:

$$\sum_{n=1}^{\infty}(a_1a_2a_3\ldots a_n)^{1/n} < e\sum_{n=1}^{\infty}a_n. \tag{2.7}$$

Polya marshals a straightforward algebraic proof: Define an auxiliary sequence $c_1, c_2, c_3 \ldots$ by the formula

$$c_1c_2c_3\ldots c_n = (n+1)^n \text{ for } n = 1, 2, 3, \ldots \tag{2.8}$$

---

[38] This is derived from Chapter 9 of [San97].

Then, by straightforward algebra we get a series of inequalities from which our theorem follows.[39] The proof is as follows: Given (a) the definition of our auxiliary sequence above and (b) the inequality between the arithmetic mean (on the right side of 2.7) and geometric mean (on the left side of 2.7), and (c) the fact that the sequence defining e, the general term of which is $[(k+1)/k]^k$, is increasing:

1. $\sum_{n=1}^{\infty}(a_1a_2a_3\ldots a_n)^{1/n} = \sum_{n=1}^{\infty}(a_1a_2\ldots a_n)^{1/n}\frac{(c_1c_2\ldots c_n)^{1/n}}{n+1}$
2. by (a) $= \sum_{n=1}^{\infty}\frac{(a_1c_1a_2c_2\ldots a_nc_n)^{1/n}}{n+1}$
3. by (b) $\leq \sum_{n=1}^{\infty}\frac{a_1c_1+a_2c_2+\ldots+a_nc_n}{n(n+1)}$
4. by rearranging terms $= \sum_{k=1}^{\infty}a_kc_k\sum_{n=k}^{\infty}\frac{1}{n(n+1)}$
5. since $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ $= \sum_{k=1}^{\infty}a_kc_k\sum_{n=k}^{\infty}\left(\frac{1}{n} - \frac{1}{n+1}\right)$
6. expanding 2^{nd} sum $= \sum_{k=1}^{\infty}a_kc_k\frac{1}{k}$
7. by (a) $= \sum_{k=1}^{\infty}a_k\frac{(k+1)^k}{k^{k-1}}\frac{1}{k}$
8. by (c) $< e\sum_{k=1}^{\infty}a_k$

We note that the proof is mathematically adequate, and the c_i sequence is essential to the proof. But despite the proof's adequacy, Polya claims, and Sandborg agrees, that it is unclear why any sequence, and this one in particular, is introduced in the first place. In Polya's words, the sequence appears as if it were a "rabbit pulled from a hat" or is a "*Deus ex Machina*" ([Pol54]: 147). Because a reader is liable to be puzzled by the introduction of this sequence, the proof is deemed insufficient. To remedy this, Polya inserts extra information which he claims makes the derivation "more understandable", or as Sandborg puts it, "explanatory".

To make the derivation more understandable, Polya ([Pol54]: XVI.6) adds

---

[39] Complete details of this proof and its history can be found in [San97]: ch 7.

some information that might motivate the decision to use the c_i sequence. Polya begins by asserting that theorem (2.7) is pretty surprising in the first place, and would be less so if we knew how it was, or could be, discovered.[40] To start, Polya tells us that theorem (2.7) is needed to prove another theorem: *if the series with positive terms* $\sum_{n=1}^{\infty} a_n$ *is convergent, the series* $\sum_{n=1}^{\infty}(a_1 a_2 a_3 \ldots a_n)^{\frac{1}{n}}$ *is small.* Polya expects that keeping this fact in mind will lead one in the direction of the correct proof. So Polya tries invoking the inequality between the geometric and arithmetic means:

$$(a_1 a_2 a_3 \ldots a_n)^{\frac{1}{n}} \leq \frac{(a_1 + a_2 + a_3 + \ldots + a_n)}{n} \tag{2.9}$$

which shows that $(a_1 a_2 a_3 \ldots a_n)^{\frac{1}{n}}$ is small when $\frac{(a_1+a_2+a_3+\ldots+a_n)}{n}$ is not large (assuming nothing can be both large and small).

But when we apply the inequality and collect terms naturally as such:

$$\begin{aligned} \textstyle\sum_{n=1}^{\infty}(a_1 a_2 a_3 \ldots a_n)^{\frac{1}{n}} &\leq \textstyle\sum_{n=1}^{\infty} \frac{(a_1+a_2+a_3+\ldots+a_n)}{n} \\ &= \textstyle\sum_{k=1}^{\infty} a_k \sum_{n=k}^{\infty} \frac{1}{n} \end{aligned}$$

the right hand sum is *divergent*, not convergent as we would have hoped. So we have not succeeded in proving what we need to prove. But since this way of collecting terms in so natural, it is worth trying to salvage what we can from it. Polya traces the failure to find a proof to the fact that since the series $\sum_{n=1}^{\infty} a_n$ converges, a_n is small when n is large, and since earlier terms will be larger than later terms in the $a_1, a_2, \ldots a_n$ sequence, the two sides of equation

---

[40] See Chapter 4 below for our discussion of the theory of explanation and reducing surprise.

(2.9) will be unequal - in our case very much so. So a modified approach is warranted. Polya recommends balancing the two sides of equation (2.9) by multiplying the a_i by some *increasing* factor. This now becomes the impetus for the c_i sequence in the proof. Polya first tries $1a_1, 2a_2, 3a_3, \ldots na_n$, and then after some rumination tries $1^\lambda a_1, 2^\lambda a_2, 3^\lambda a_3, \ldots n^\lambda a_n$, with λ left as indeterminate for the meanwhile, as a value we can fill in when we figure out what value will serve our needs.

After some calculation, Polya finds that we can only get an approximate result, but one close enough to the original term to still be useful: e.g., $e^\lambda \lambda^{-1} \sum_{k=1}^{\infty} a_k$. Choosing a value of $\lambda = 1$ will keep $e^\lambda \lambda^{-1}$ to a minimum, and perhaps give us the value we need to prove theorem (2.7).

However, setting $\lambda = 1$ still leads Polya to a term he cannot calculate, though he realizes that a sequence with similar growth behavior may work. So he considers using a different *sequence*, perhaps one *close to* $1, 2, 3, \ldots$ which is asymptotically equivalent to it at the limit. Call this sequence c_n, and go through now familiar steps:

$$\begin{aligned}
\textstyle\sum_{n=1}^{\infty} (a_1 a_2 a_3 \ldots a_n)^{1/n} &= \textstyle\sum_{n=1}^{\infty} (a_1 a_2 \ldots a_n)^{1/n} \frac{(c_1 c_2 \ldots c_n)^{1/n}}{(c_1 c_2 \ldots c_n)^{1/n}} \\
&= \textstyle\sum_{n=1}^{\infty} \frac{(a_1 c_1 a_2 c_2 a_3 c_3 \ldots a_n c_n)^{\frac{1}{n}}}{(c_1 c_2 c_3 \ldots c_n)^{\frac{1}{n}}} \\
&\leq \textstyle\sum_{n=1}^{\infty} \frac{a_1 c_1 + a_2 c_2 + a_3 c_3 \ldots + a_n c_n}{n (c_1 c_2 c_3 \ldots c_n)^{\frac{1}{n}}} \\
&= \textstyle\sum_{k=1}^{\infty} a_k c_k \sum_{n=k}^{\infty} \frac{1}{n (c_1 c_2 c_3 \ldots c_n)^{\frac{1}{n}}}
\end{aligned}$$

But how do we choose a c_n such that it allows us to evaluate $\sum_{n=k}^{\infty} \frac{1}{n (c_1 c_2 \ldots c_n)^{\frac{1}{n}}}$? Polya then recalls that $\sum \frac{1}{n(n+1)} = \sum \left(\frac{1}{n} - \frac{1}{n+1} \right)$ and $\sum_{n=k}^{\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \frac{1}{n}$. If $c_1 c_2 c_3 \ldots c_n = (n+1)^n$ Polya can insert the sequence into the proof and complete it. Beside completing the proof it also clarifies the origin and genesis of the c_i sequence, and tells us how we could have taken the original step in the proof. So what Sandborg calls "Polya's explanation" is the additional material that leads us in a fairly intuitive way from the motivation[41] behind the theorem to the c_i sequence which lets us begin the proof. This material is extraneous from the perspective of justifying the theorem so its sole function is to provide explanatory material.

Analysis of the example

Let us look at this example more carefully. Our first question is: is this a genuine example of mathematical explanation? Sandborg asserts that it is. But we are skeptical. We can show that this example violates a number of conditions for a genuine explanation that we laid out earlier. We can show that Polya actually offers us an explanation that is subjective, pedagogical, and only speaks to the context discovery. In §1.5 we spelled out the uses of "explanation" which are often conflated with the kind of explanations that philosophers are interested in. Given that it is clear that Polya intended these non-philosophical uses of "explanation", there is no reason to assume that Polya would additionally claim that he is giving an explanation of mathematical facts of the sort that interests philosophers. Therefore, Sandborg has no warrant to claim that Polya is giving us explanations. We have three

---

[41] It is not uncommon for mathematicians to point out that there is some unmotivated step in a proof. But unless the mathematician is specifically making a pedagogical point, the motivation will be left to the resourcefulness of the reader. See e.g., [Men97]: 37, n†.

reasons for asserting this. The first is that Sandborg misreads Polya's questions; they are not why-questions analyzable on a theory of mathematical why-questions. The second reason is that Polya does not appear to be giving us an explanation. And the third reason is that even philosophers who claim that Polya is giving us an explanation do not seem to take him seriously with respect to his own analysis of his "explanations".

First, the why-questions Sandborg attributes to Polya are not the questions that Polya is asking. As we will discuss shortly, Sandborg identifies the following two questions that Polya is addressing in this example: (1) "why do we introduce an auxiliary sequence into the proof?", and (2) "why did we use the particular sequence that we used?". The answer to both of those questions should be pretty obvious after seeing the proof: we used them because they work; they make the proof go through.

This may be a trivial answer, but it is the correct answer to both the above questions. So what of Polya's perplexity? Do we just dismiss it? Of course not. Polya is asking important questions. They are (1′) "What might lead someone to the idea that an auxiliary sequence will be useful for a proof of equation (2.7)?", and (2′) "What might lead someone to the idea that we should use this particular sequence to prove (2.7)?".

Neither (1′) nor (2′) are mathematical why-questions, nor are they requests for mathematical explanations. They are both psychological questions regarding how human mathematicians and mathematics students could have come up with certain ideas. So our first claim about Sandborg's example is that a closer inspection or Polya's questions reveal them not to be requests for mathematical explanations after all.

Second, *prima facie* the example does not appear to be explanatory. The extra material we are given to get to step (1) of the proof looks like a set of hints in a solution manual. Polya is talking us through the steps as if to say "How do you solve this problem? Try this. Try that. Then you may get the first step."[42] Polya is certainly not explicitly offering or attempting to offer a theory or an example of explanation. Polya claims that he is presenting a heuristic. A heuristic is a device for solving problems. And there is no obvious philosophical road from a heurisitic to an explanation. The as yet unmet burden of proof thus lies with Sandborg who claims that Polya is giving us more than a heuristic. For the reasons we mentioned in §1.5, we will not call something an "explanation" if its only claim to being an explanation is that it provides heuristic or pedagogical assistance for some

---

[42] Compare our example with the following dialogue (from [Par87]):

> q: Do you know the factorisation of 143?
> r: Not off hand.
> q: Is 11 a prime?
> r: (After thinking a little) Yes.
> q: Is 13 a prime?
> r: Yes,
> q: How much is 11 times 13?
> r: Let us see; 11 times 10 is 110. 11 times 3 is 33. 110 plus 33 is 143.
> Oh, I see.
> q: Can you factorise 143 now?
> r: Of course, it is 11 times 13.

theorem or proof.

Given how readable Polya's work is, it is not surprising that some philosophers think that Polya was providing examples of mathematical explanation. Polya's work has become an inkblot where each reader sees his own favorite view well described therein. Sandborg and Steiner[43] see examples of explanation in Polya's mathematics. Corfield ([Cor03]: Ch. 5) sees a Bayesian theory of confirmation in Polya's outline of degrees of certainty in mathematics. Teachers of mathematics see pedagogical strategies (e.g., [Bro56]), psychologists see it as "a source of suggestion for research on higher mental processes" ([Bus56]: 166) . . . and mathematicians see interesting proofs. Despite numerous reviews the community of mathematicians failed to notice any explanations, leaving us with further reason to doubt that either Polya gives us examples of explanation, or even that mathematicians are interested in explanations at all.

Much of Polya's written work, going back to his now classic *How to Solve it* ([Pol57]), is about the nature of heuristics and pedagogy. In the absence of a theory of explanation, solving a problem is not the same as explaining it no matter how you solve it, and Polya's subsequent work on induction and analogy in mathematics also fails to further the goal of providing explanations. There is undoubtedly mathematical and pedagogical value to Polya's work, but if there is any *philosophical* value at all to Polya's work, the only

---

[43] See §3.1 below for our discussion of Steiner's views.

real case made for it is by Corfield on behalf of a proto-Bayesenistic account described in [Pol54].

In our particular case we are given a specific problem to solve and an approach to solving it. The approach accounts for (and satisfies) a certain perplexity about how we could have thought of the initial step of the given proof on our own. But as a philosophical problem it is still not close to saying that Polya gave us a logic of confirmation or discovery. But even if he did, that is not our present philosophical concern; the idea that Polya was providing explanations is.

As a matter of fact, this particular example was originally published as a "classroom note" ([Pol49]) where Polya is quite specific that the value of his proof is in its ability to "clarify a few not quite trivial points of classroom technique." (685) In the book, where Polya revises this proof, he reiterates this claim. When Polya reflects on his own method he claims that the difference between the two proofs he offers is the difference between proofs and plausibilities. The former giving demonstrative conclusions, the latter heurisitic motives. The latter can be used by an ambitious teacher, Polya claims, to distinguish between a more and less reasonable guess. So we have no reason to think that we are being given anything beyond a useful pedagogical technique.

The third reason we have to think that Polya did not give us an example of mathematical explanation is that Polya's methodology is not taken seriously as a model of explanation even by those who claim that he is giving us explanations. As we mentioned, both Steiner and Sandborg claim that Polya gave us examples of mathematical explanation. But Polya did not merely give us examples. Polya talked us through those examples step-by-step. Sandborg ([San97]) even recapitulates and clarifies Polya's analysis. But neither Sandborg nor Steiner give any indication that they take Polya's analysis of his own process by which he came to these alleged explanations to be a model of explanation. But if it is so obvious that these are explanations, and Polya clearly gives us detailed break-downs of the process of obtaining them, or at least the process by which he claims they can be obtained, then Polya's discussion ought to be seen as an ideal (or at least initial) candidate for a model of mathematical explanation. But Steiner does not even consider it, but rather uses one of Polya's examples as a foil for his own model of explanation. Sandborg who does discuss Polya's analysis, discusses it only in service of the why-question approach to explanation. And Corfield, who does discuss Polya's methodology in detail, and is aware of (and probably sympathetic with) the question of mathematical explanation, uses Polya's work only for other philosophical purposes.

If Polya didn't think his example is an example of mathematical explanation and the mathematical community did not mention it, and none of the philosophical commentators took his methodology seriously enough as a model of explanation, then it is unclear where the intuition that Polya's extra material is a genuine example of mathematical explanation came from.

Arguments against the why-question account for mathematics

Let us now examine the arguments against a why-question model of mathematical explanation. Sandborg is sympathetic with the idea that there are mathematical explanations and claims that the example above from Polya is an illustrative case of one. Sandborg claims that the example is *prima facie* analyzable in terms of why-questions. How? As we mentioned above, Polya can be said to be answering two different why-questions (with different contrast classes) about the initial proof of equation (2.7). The first why-question is: why is it appropriate to introduce the c_i sequence into the proof? The second question is: why of all sequences, introduce that particular one? Assume, despite our arguments above that these are not the why-questions that Polya asks, that these *are* the appropriate why-questions to ask in this context. The answer to the first question is: to replace a divergent series with a convergent one. The second question has two answers: (a) because our particular sequence exhibited a favorable growth behavior, and (b) because it allows for simplification of a crucial term in the derivation.

Sandborg criticizes an analysis of the example in terms of the why-question model on the grounds that the analysis of the question and answer does not correctly account for what rings true about Polya's "explanation" (612).

The why-question approach to mathematical explanation is susceptible to a number of criticisms. First, one can show that given some example of a mathematical explanation it cannot be accounted for by the why-question approach. (Or rather: given some example provided by one who does have the intuition that there are mathematical explanations, the explanation is not analyzable in terms of the why-question approach.) Second, one can show that given the constraints of the why-question model, it cannot be a model for mathematics.

Sandborg's first argument against van Fraassen's theory of explanatory evaluation is that the model cannot account for mathematical answers to why-questions in which the members of the contrast class are mutually exclusive. Here we first present Sandborg's argument together with our diagnosis of its soundness.[44]

Sandborg asks us to consider a why-question that has a contrast class of mutually exclusive members such as "why does $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \ldots$ converge to $\frac{\pi}{4}$ rather than some other real number?" The problem is that under van Fraassen's theory, any proof of the topic will be a telling answer because regardless of the answer it would be judged as maximally probable - i.e., proven (probability of 1 that it is correct), and the answer follows from accepted mathematical premises. All other members of the contrast class will have a probability of 0 because they are all incorrect answers. No other answer can favor the topic better or screen it off. Therefore any proof of the

---

[44] More details can be found in the [San98].

topic is trivially a telling answer on van Fraassen's theory.

As an aside we note the following about this last example. Regardless of how we understand this natural language why-question, whether it is as: "why does $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}\ldots$ converge to $\frac{\pi}{4}$ rather than some other real number?", or "why does $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}\ldots$ converge to $\frac{\pi}{4}$ rather than some other series (that does not converge to $\frac{\pi}{4}$) converge to $\frac{\pi}{4}$?", or "why does $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}\ldots$ converge to $\frac{\pi}{4}$ rather than diverge?" the answer ends up being the same. We would offer a proof that $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}\ldots$ converges to $\frac{\pi}{4}$. Regardless of the contrast class, there is still only one type of acceptable answer. So the theory of why-questions ends up having a particularly uninteresting theory of mathematical, logically precise why-questions for a large class of cases.

Returning to the problem of the proof probabilistically favoring the topic, it should come as no surprise that all mathematical *proofs* are trivially telling. Telling answers rely completely on the background knowledge of the questioner and the kind of answer being solicited. Provided that the questioner knows enough about mathematics to grasp the proof, there is no reason any proof should prove unsatisfying if you just wanted to know why the proof is valid or the theorem is true. In mathematics there are no other kinds of answers. And while there can be different proofs of the same theorem, they will all be the same kind, they will be mathematical proofs. This is true regardless of how you restrict $K(Q)$, just as long as $K(Q)$ still allows for a proof of the topic. And since the proof follows from $K(Q)$, it defeats the point of restricting K. If we cannot restrict K, then we are back to the problem we mentioned in the exposition of the theory. Moreover, the theory of telling answers does not even seem to have the resources to distinguish between different types of proofs say, between algebraic and geometric proofs; though even if it did, one could not plausibly argue that a proof that followed from the subset of what we know about geometry is more explanatory than a proof that followed from the restricted subset of our knowledge that contained what we know about algebra. It is also unreasonable to argue that geometric proofs are more explanatory than algebraic ones. Algebraic or geometric proofs are merely styles of proof that a questioner may be more or less familiar with. These kinds of proofs thus provide answers that the questioner may be more or less comfortable with. Here, a lack of understanding does not make an answer inappropriate in the same way that an answer that appeals to molecular biology is inappropriate when it is clear that we are seeking a functional answer to a question about the evolutionary history of some animal. In this latter case the answer references the wrong biological theory - e.g., some theory about bio-chemistry, instead of the theory of evolution.

In mathematics it seems impossible to to refer to the wrong theory. First, the word “theory” in “number theory” and “set theory” is different from the word “theory” in “evolutionary theory” and “cell theory”. Though there are no clear accepted definitions of “theory” in mathematics or science, in the former case "theory" refers, say in the case of model theory, to sets of axioms, rules of inference, and the theorems they produce, or in the case of set theory to a body of knowledge within mathematics. In the latter case, on the other hand, "theory" refers to a complicated set of hypothesized rules that frame the (sometimes causal) relationships between different pieces of empirical data together with the tests of these theories, the inferences and generalizations from the data, and the explanations that all of these provide. Scientific theories are open (among other things) to verification, complete or partial revision, new evidence, correction, and statistical analysis of the likelihood of success; mathematical theories are not. In science we may be looking for confirmation of, support from, or consistency with some scientific theory when asking a why-question. We can thus respond to some why-questions with answers that appeal to a given theory knowing that it is what the relevance relation demanded. Or we may have a pragmatic context that demonstrates the need for some particular type or level of answer being necessary for the proper response to the question. We can have no such needs in the mathematical case because the fate of say, algebra, rises and falls with the success and failure of geometry. They are not in any meaningful way independent theories.

Sandborg further argues that if one believes that some proofs explain or some explain better, the fact that on the why-question account all mathematical proofs of a theorem have a probability of 1 or 0 of being valid is also problematic. Sandborg claims that in science some explanations seem better than others, therefore a comparable claim for mathematics is plausible. But this is misleading. All the major accounts of scientific explanation take each explanation to either hold as *the* explanation or make way for one that does.[45] That means that the "better" explanation will replace all other explanations, relegating all the latter to the status of non-explanations. It is only in the sense that some explanations are somehow clearer than others is it the case that we judge explanations to be better or worse.

It is only because Sandborg merges these various senses of explanation that the previous question makes sense and as such he suggests that a revision of the why-question model is in order in light of the Polya example. To judge the Polya case as explanatory Sandborg claims that we may have to allow for proofs to be somehow more or less explanatory depending on whether or not the c_i sequence is *uniquely* picked out, as *that* would make the proof truly explanatory – by answering the why-question as specifically as possible. (Recall that Polya's exposition did not originally provide a *unique* answer to what the c_i sequence should be.) The answer that uniquely picks out the correct c_i sequence would be the most explanatory, if there is such an answer. So since Polya's "explanation" does not appear to initially uniquely favor the c_i sequence, the proof is not a telling answer, and only some hypothetical

---

[45] At the very least, we note that while there are many theories of the nature of scientific explanation, there are relatively few accounts of the nature of better and worse explanations. Some philosophers however (e.g., Noretta Koertge ([Koe92])) have listed this among the problems of accounts of scientific explanation.

explanation that did uniquely pick out the sequence is a genuine explanation. But if that were true, that only an elaboration that uniquely picked out the c_i sequence is genuinely explanatory, then we would need to account for the initial intuition that had Sandborg pick out Polya's elaboration on the proof as an explanation in the first place.

We must also consider the possibility that there is no answer that uniquely picks out the correct sequence. In that case there are some theorems that either (1) have no explanations, (2) have no good explanations, or (3) that Sandborg's example is completely off, and the real explanation for theorem (2.7) is a whole different proof, or something else entirely. If (1) then we have a significant disanalogy with science where everything is presumed to have some explanation.[46] (2) just sounds odd. It would be difficult to make sense of the claim that there is something that has many explanations but that (in principle) they are all not very good. Part of our judgment stems from a perplexity about what a bad explanation might be. Even if we knew what an explanation or a good explanation is, there is nothing that tells us what a bad one should look like. Finally, (3) undermines the original claim that our proof is explanatory in the first place. If this is a bad example of explanation, then we ask again, whence the original intuition that the new proof explains?

So our first criticism of the why-question model in a mathematical context

---

[46] See ch. 4 below.

is that in the mathematical cases all the proofs end up trivially telling. So if the theory of telling answers and the contrast class are both trivialized, the only part of the explanation remaining not trivialized is the relevance relation. That is all that is left to distinguish explanatory from non-explanatory answers.

Sandborg also points out that the reason we do not consider some proofs to be explanatory usually has nothing to do with the relevance relation. This is especially for reasons that the questioner can specify in advance (615). As Polya's proof illustrates, we never know in advance what might be useful to the proof. And if the relevance relation is the last remaining non-trivial part of the theory, then there will be some cases that we have further reason to claim is unanalyzeable on this account.

So Sandborg's claim is that a mathematical example (like the one he cites from Polya) will not necessarily be explanatory with respect to the why-questions that motivated the example in the first place. The problem is that the relation between the why-question and the answer is so loose it allows for many answers, including some that are clearly not explanatory. No one holds that all proofs are explanatory, as making all proofs explanatory would trivialize the problem of mathematical explanation.

We alluded earlier to another of Sandborg's arguments against the why-question model. Van Fraassen regards an explanation as a presentation of descriptive information relevant to the topic. The information in an explanation is similar in kind to a description. This description in turn finds its way into telling answers. An answer is evaluated based on its contribution to favoring the topic in light of $K(Q)$, a restricted subset of our background knowledge. But in mathematical cases, K is either axioms or perhaps the deductive closure of some axioms. The topic is then [in K or] implied by K. So K must be restricted to some $K(Q)$. So we must then consider a smaller set of axioms or their deductive closure. But if we did that the answer would be a proposition which together with $K(Q)$ implies the topic. This would be a missing axiom or some weaker statement that implies the theorem. (615) But van Fraassen's theory, on a mathematical account, would suggest that the point of a mathematical explanation is to provide an analysis of *why* a theorem is true. One way mathematicians may provide such an analysis is by employing what is known as reverse mathematics - showing which axioms a theorem depends on. But there does not seem to be anything explanatory about reverse mathematics, and no one has suggested there is.

Moreover, proofs are said to be non-amplitive. That means that a proof adds no new information - descriptive or otherwise. Rather it draws out information already said to be “in the axioms”. Explanations in science, on the other hand, do give us additional information (see Chapter 4). So how could a proof, which tells us nothing new, serve as an explanation? Yet of those who claim that there are explanations in mathematics, most claim that explanations could be proofs.[47]

The fact that the theory of telling answers has a probabilistic component should have made us immediately suspicious. Since theorems with proofs will almost always be judged by mathematicians to have a probability 1 of being true,[48] that is they are all judged to have the same probability, there is no way to evaluate answers against each other.

A further argument (not due to Sandborg) appeals to a more general understanding of the why-question approach to explanation. As Sandborg ([San98]) points out, the inherent topic neutrality of the why-question approach should imply that the why-question model does not discriminate between mathematics and science. The model should apply to both in the same way. Presumably this is only obvious if you take the epistemic (and ontological) concerns of mathematics and science to be equally accommodated by one methodology. But the model's applicability to mathematics is not obvious to everyone. Addressing the why-question approach to explanation, Alex Rosenberg claims that "[t]he pragmatics of language is presumably something we can ignore in mathematical proof, but not . . . in scientific explanation" ([Ros00]: 37). Presumably he claims this because the pragmatic

---

[47] Steiner's model (see§3.1 below) describes explanatory proofs, though Steiner does not rule out other kinds of explanations. Kitcher claimed that Bolzano's proof of the Intermediate Value Theorem was motivated by explanatory considerations (see §2.3.1 above). Mancosu (see §3.2, claim 5 below) claims that mathematical explanations take on a variety of forms, not excluding proofs . . .

[48] Even if you take Corfield's Bayesian notion of confirmation seriously, the probabilities will generally not (without significantly modifying the theory) range over already proven theorems, only mathematical conjectures.

context does not figure into mathematical methodology; and to the extent that it does, it is not at all obvious how it helps to explain mathematics. So for Rosenberg the pragmatic part inherently precludes the model from being applicable to mathematics. A theory of explanation that relies on the human context would not seem to accommodate a system such as mathematics whose content does not take pragmatics into account at all. An emphasis on pragmatics is certainly not a part of mathematical practice. Also given that this model of explanation was given by van Fraassen in the context of elucidating his Constructive Empiricism - his ontological and epistemic views of science and he did not provide an account of mathematics, it is not clear that van Fraassen would have intended his model to be used here. This model is extremely useful in furthering the goals of Constructive Empiricism. But Constructive Empiricism is certainly tied to scientific goals.

Moreover, even the view that mathematics answers why-questions is not obvious. As we saw above[49] J. P. Marquis thinks that mathematics does not answer why-questions, rather he thinks it answers "how-possible" questions which seem to call for instructions rather than explanations.[50] So from the beginning, a model that is predicated exclusively on an analysis of the syntax, semantics, and pragmatics of why-questions seems inapplicable.

---

[49] See section §2.3.1.

[50] How-actually questions can invoke many kinds of responses. One of Salmon's objections to van Fraassen's model is sometimes how-actually questions in addition to why-questions are also requests for explanation. Take a question of the form "How did there come to be non-flying land mammals in New Zealand, an island far from any land mass?"

Salmon ([Sal89]) pointed out that in addition to asking whether all explanations can be given in terms of why-questions, we can also ask whether all why-questions are requests for explanations. There are many questions that look like why-questions that simply do not call for explanations. Salmon mentions questions like "why do bad things keep happening to me?", which has the surface grammar of a request for an explanation is actually meant to invoke sympathy. But here we contend that mathematical cases provide better examples of why-questions that are not requests for explanation.

A number of objections against van Fraassen's account have emerged in the philosophy of science. Some of these objections are far more relevant when applied to mathematics rather than to science. Consider a variation of an objection noted by Edwin Hung ([Hun97]: 180): Would the why-question model of explanation, if it was indeed a model of mathematical explanation, further any mathematical goals? Is this theory of explanation relevant to mathematical methodology? These questions challenge the relevance of the model relative to a perceived need of a discipline to provide explanations. we have seen some who have argued[51] that explanation is a goal of mathematics, and moreover it is a mathematical goal of mathematicians. Moreover, the search for a model of explanation only makes sense in the context of such a goal. But if we know that explanations conform to the why-questions model do we now know something about the method mathematicians use?

---

[51] See e.g., [San97] and [Man00].

The model of mathematical explanation we use should tell us about the role why-questions play in mathematics. But this model does not do that at all. The model we choose of mathematical explanation should tell us what sort of things in mathematics need explanations. But as the Polya example illustrates, the model gives us no insight into what requires explaining beyond "those things that seem unsatisfying to the student of mathematics".

Those parts of mathematics that leave students unsatisfied may or may not drive mathematics. The model offers us little guidance about which explanatory considerations if any would drive mathematics. "Why is 17 prime and 18 not?" does not drive mathematical research; "why is $n^2-79n+1601$ prime for $n \leq 79$ and not for $n = 80$?" did. One would expect that if the why-question model of explanation was appropriate for mathematics then it would "guide" us to the explanatory questions. But it does not. So if the answer to the second question is explanatory, why isn't the answer to the first also? The first answer is merely a definition, and how can a definition be an explanation?

So given our reasons above - that mathematical cases tend to trivialize the model, and that the why-question model adds little of value to our understanding of mathematics, we find that the why-question approach to explanation is inappropriate as a model of mathematical explanation.

The two-proof methodology

Sandborg's example also raises an important methodological point which

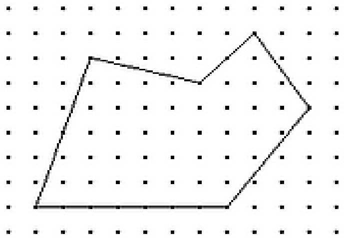

Figure 2.2: A Lattice Polygon

is worth a brief discussion here. The methodological point concerns the proper method of selecting examples to consider as mathematical explanations. In Sandborg's [San97] he argues that mathematicians have genuine mathematical concerns beyond the justification of theorems, and one of those mathematical concerns is to provide explanations. Sandborg's argument proceeds by tracing the history of Pick's theorem. Pick's Theorem states that given i, the number of interior points of a simple lattice polygon and b the number of points on its boundary, then the area A is given by the formula: $A = i + \frac{b}{2} - 1$. In Figure 2.2, $i = 39$, $b = 14$ so $A = 39 + \frac{14}{2} - 1 = 45$ (square units). Sandborg assembles 12 proofs of the theorem and shows that the authors of many of the successive proofs in the history of Pick's theorem are trying to improve on one or another of the previous proofs in some particular way - in some cases by inserting explanatory material into the proofs or by giving explanatory proofs. The case of Pick's Theorem is thus methodologically similar to Sanborg's example above of Polya's theorem. In both cases we are given a proof which is judged by some mathematicians[52] to be in some way inadequate because it fails to explain. We shall see in §3.1 and §3.2 that Mark Steiner and Paolo Mancosu respectively also use the same strategy for demonstrating the existence of mathematical explanations. Jamie Tappenden has done so as well.[53] By doing so, Sandborg and the other philosophers we mention are endorsing this methodology.

To clarify, the strategy that some philosophers use to discover explanations in mathematics is as follows: find a proof of x in the mathematical literature. Next find a subsequent proof of x which was given by a mathematician who was aware of the original proof of x and perhaps expressed some kind of dissatisfaction with the [style of the] original proof. Then, claim that the second proof was given for "explanatory reasons", and argue that this shows two things: First that we have a genuine example of mathematical explanation, and second that mathematicians (and the mathematical community) are interested in offering mathematical explanations.

Let us postpone discussion of the second claim about the mathematicians and their community for later.[54] But does a "two-proof example" show examples of, or interest in, mathematical explanations? I argue that it does

---

[52] See §3.2: claim 2 below for our discussion of mathematicians' judgments.

[53] Tappenden ([Tap05]: 149-150) explores the case of Riemann's "conceptual" versus Weierstrass' "computational" approach to proof methods in complex analysis. Weierstrass "aimed at finding explicit representations of functions and explicit algorithms to compute their values . . . Riemann . . . aimed to describe functions in terms of general properties, and to prove indirect function existence results that need not be tied to explicit representations."

[54] See §3.2: claim 1 below.

not. More importantly I argue that as a matter of philosophical methodology, it is a weak maneuver. First, for reasons I outline in the next chapter[55], we want a *model* of explanation to exist prior to our having individual examples of intuitive explanations. So without a model of explanation it is hard to know what counts as a real example of mathematical explanation.

In the absence of a model however we would prefer examples that *prima facie* look like explanations. It seems unreasonable to say that we only know an explanation in hindsight - after we see that it does something that another proof of the same theorem fails to accomplish. One reason for this is that when we examine scientific explanations we rarely find failed explanations cited merely to bring out the explanatoriness of a successful explanation. So by comparison with scientific explanation, using and exhibiting two proofs of the same theorem to illustrate mathematical explanation seems out of place.

It would certainly be a much clearer test of the explanatoriness of some piece of mathematics if we were just shown a mathematical example and we found it to be intuitively explanatory. But the two-proof method interferes with our intuitions. In all the two-proof cases in the philosophical literature, many people would say that after looking at the first proof, the second one looks better. It is almost as if our intuitions were primed to appreciate the advantages of the second proof in light of the poverty of the first. But although we judge it as better, we have no reason to judge it as explanatory.

---

[55] See §3.2: claim 4 below.

By having us judge the proof as better we put ourselves in a worse position to judge whether or not it is explanatory.

There is an important reason why one may wish to exhibit two proofs. The pedagogical point of the example, if one is being made at all (and it is pretty clear that one is being made in Sandborg's example), would be incomprehensible without a clear discussion of both proofs. Polya would have a difficult time convincing someone that his proof was in some special way "motivated", or better to teach with, without having the original "unmotivated" proof to compare it with. On the other hand, the explanatory point being made by Polya, if he is indeed making one, should not need us to see both proofs to appreciate the explanation. If the second proof is explanatory (or explains) then the original unexplanatory proof (or proof without the explanatory material) serves no purpose - certainly not to further an explanation of the explananda. The extra proof is thus unnecessary.

Now we can see why there is harm in citing two proofs of the same theorem as part of the methodology of looking for mathematical explanations. The proof that shows that some theorem T is true should also be used to explain T, if it is possible to do so. The proof that establishes for us *that T is true* should also show us *why.* If we use separate proofs we run the risk of conflating the pedagogical point with the explanatory (philosophical) one. The explanation or the explanatory proof should not need the existence of the unexplanatory proof. If the proof explains, one would assume that it would explain simpliciter and we would notice the explanation, not just when we are confronted with an unexplanatory version of the same proof, or an unexplanatory proof of the same theorem. As we just mentioned, it is certainly the case in scientific explanation that a physicist, for example, would not feel compelled to offer say, a Newtonian explanation of some phenomenon so that the relativistic explanation appears more obvious. Contrasting non-explanations are only offered when the scientist wants to show that some other particular [kind of] explanation is mistaken, not when the scientist wishes to exhibit the the correctness of his or her own view. Scientists mention rival explanations, for example, in discussions of the debate between versions of the theory of evolution that advocate phyletic gradualism versus punctuated equilibrium.

Explanations, should they exist in mathematics, ought to *prima facie* respond directly to some perplexity and appear as explanations to the light of our intuitions, not merely when they are proofs that are contrasted with “non-explanatory” or “less-explanatory” proofs. It is one thing to argue that there are explanations in mathematics, á là Steiner, Kitcher, or van Fraassen. It is another to claim that the way to discover them is by comparing proofs.

2.5 The familiar

Another view we want to consider is the view that explanation is the reduction to the familiar.[56] That this is an important style of explanation in science was suggested by, among others, Norman Robert Campbell (see, e.g., [Cam57]: 114). On this account to explain is to reduce complicated and mysterious phenomena to phenomena that are more tractable and commonplace. On this theory, explanation consists in breaking some large-scale phenomena down to ever-smaller parts and showing how each part is non-mysterious and well understood, and how each part is accounted for until the smallest possible level of resolution. Then, it is claimed, the larger phenomena are really just the sum of the smaller ones, each of which needs no explanation, or is taken to be explanatorily primitive, or self-evident. At this point we can claim explanatory success because we have said all there is to say about the larger phenomena.

Mancosu cites Karl Menger's intuitionistic-formalistic dictionary as an example of this kind of explanation in mathematics ([Man00]: 106). Brouwer's work on intuitionistic set theory was introduced in intuitionistic terminology. But his work was puzzling to many classical mathematicians until Menger published a "dictionary" that showed that Brower's terminology corresponded to classical concepts. For example, a "species" in Brouwer's terminology was a "set" in traditional terminology.

---

[56] For this see [SEG$^+$92], p 14.

However, Mancosu's example cannot be right, as dictionaries work both ways. Classical mathematicians (like Hausdorff) may have learned how to read intuitionistic texts from this dictionary. But if there were intuitionistic mathematicians, presumably they can benefit from the dictionary by learning how to read classical mathematics. So if a translation from definition to definienda is explanatory, then so is a translation from definienda to definition. But then we are in the position where x explains y, and y explains x. This is unacceptable in a theory of explanation.

However, this view of explanation in science was already rejected by Hempel ([Hem66]: 83). Wesley Salmon ([Sal89]: xii) also states that "in company with most thoughtful philosophers, I reject the view that scientific explanation involves reduction of the unfamiliar to the familiar...". It is a view he attributes to "some scientists" ([Sal89]: 6). One reason we might reject this view of explanation is that sometimes our greatest explanatory success involves reducing concepts or phenomena to the *un*familiar. Heat, for example, is more familiar to us than the concept of mean kinetic energy. Yet we explain heat by appealing to the concept of mean kinetic energy. So we explain the familiar with an appeal to the unfamiliar.

Elsewhere, while discussing various kinds of explanation, in the context of scientific explanation, Wesley C. Salmon says

> ...someone might ask for an explanation of a mathematical proof; an appropriate response would be to fill in additional steps to show how one gets from one step to another in the original demon-

> stration. ([SEG$^+$92]: 8)

Salmon does not take there to be mathematical explanations. But one way of considering the question is along the lines that Salmon seems to suggest. Could a mathematical explanation be that which is gained when we fill in the steps of a proof and show how to get from one step to the next? No. A simple example ought to suffice to show that explanations do not come about in mathematics via reductions of this sort.

Leslie Lamport [Lam95] suggests a method to help clarify mathematical proofs and decrease the likelihood of mistakes. He also alludes to the pedagogical and epistemological benefits as well as the advantages in obtained in logical rigor. He gives the example of the classical proof of the irrationality of $\sqrt{2}$. The theorem to be proven is: There does not exist r in $\mathbb{Q}$ such that $r^2 = 2$, where $\mathbb{Q}$ denotes the set of rationals.

What we might generally expect as a proof of the irrationality of $\sqrt{2}$ is a proof sketch that looks something like the following:

> **Theorem.** There does not exist r in $\mathbb{Q}$ such that $r^2 = 2$.
>
> Proof Sketch: We assume $r^2 = 2$ for $r \in \mathbb{Q}$ and obtain a contradiction. Writing $r = m/n$, where m and n have no common divisors (Step 1), we deduce from $(m/n)^2 = 2$ and the lemma that both m and n must be divisible by 2 (Steps 2 and 3).
>
> ASSUME: 1. $r \in \mathbb{Q}$
> 2. $r^2 = 2$
>
> PROVE: False

1. Choose m, n in $\mathbb{Z}$ such that
    1. gcd(m,n)=1
    2. r=(m/n)
2. 2 divides m.
3. 2 divides n.
4. Q.E.D.

There is, however, a more fine-grained way to express this proof. Using what Lamport (oddly) calls "natural deduction" we produce a structured proof: we get an extremely detailed, fine grained proof of the same theorem that looks like this:

> **Theorem.** There does not exist r in $\mathbb{Q}$ such that $r^2 = 2$.
>
> Proof Sketch: We assume $r^2 = 2$ for $r \in \mathbb{Q}$ and obtain a contradiction. Writing $r = m/n$, where m and n have no common divisors (step <1>1), we deduce from $(m/n)^2 = 2$ and the lemma that both m and n must be divisible by 2 (<1>2 and <1>3).
>
> Assume: 1. $r \in \mathbb{Q}$
> 2. $r^2 = 2$
>
> Prove: False
>
> <1> 1. Choose m, n in $\mathbb{Z}$ such that
> 1. gcd(m, n) = 1
> 2. r = (m / n)
>
> <2>1. Choose p, q in $\mathbb{Z}$ such that $q \neq 0$ and $r = p/q$.
> Proof: By assumption <0>:1.
>
> Let $m \equiv p/gcd(p,q)$
> $n \equiv q/gcd(p,q)$

<2>2. $m, n \in \mathbb{Z}$

PROOF: <2>1 and definition of m and n.

<2>3. $r = m/n$

PROOF: $m/n = \frac{p/gcd(p,q)}{q/gcd(p,q)}$ [definition of m and n]

$= p/q$ [Simple algebra]

$= r$ [By $< 2 > 1$]

<2>4. gcd(m,n) = 1

PROOF: By definition of the gcd, it suffices to:

ASSUME: 1. s divides m

2. s divides n

PROVE: $s = \pm 1$

<3> 1. $s \cdot gcd(p, q)$ divides q

PROOF <2>:1 and the definition of m.

<3>2. $s \cdot gcd(p, q)$ divides q.

PROOF: <2>:2 and the definition of n.

<3>3. Q. E. D.

PROOF: <3>1, <3>2, and the definition of gcd.

<2>5 Q. E. D.

<1>2. 2 divides m

<2>1 $m^2 = 2n^2$

PROOF: <1>1.1 implies $(m/n)^2 = 2$.

<2>2. Q.E.D.

PROOF: By <2>1 and the lemma.

<1>3. 2 divides n.

<2>1. Choose p in $\mathbb{Z}$ such that $m = 2p$.

PROOF: By <1>2.

<2>2 $n^2 = 2p^2$

PROOF: $2 = (m/n)^2$ [<1>1.2 and <0>:2]

$= (2p/n)^2$ [<2>1]

$= 4p^2/n^2$ [Algebra]

From which the result follows easily by algebra.

<2>3 Q.E.D. Proof: By <2>2 and the lemma.

<1>4 Q.E.D. Proof: <1>1.1 <1>2 <1>3 and the definition of gcd.

(Note that here "$\equiv$" means "equal by definition".[57])

Reducing a bit of mathematics to its most primitive and presumably most familiar, or intuitive mathematical statements, in the above case is very similar to what we might expect from a reduction in science where we look for the most "fundamental" laws, or types of laws. This happens when we fill in the steps to the proof, and thereby reducing each individual step to a more "primitive" step. However, the proof is much longer (because it fills in *all* the steps), less perspicuous[58], does not aid our understanding, adds little that a mathematician would find useful, and can certainly not be judged to be explanatory - certainly not more than the first proof. The latter proof is less readable and intuitively does not explain the theorem, in any way, any better than the former. It is for these reasons we claim that reduction to the familiar in mathematics is not a way to offer mathematical explanations.

We also note that most ways of using "familiar" will have a subjective account. Something can be familiar to one mathematician and not another. Though in §1.5 we take pains to distance ourselves from such a view of explanation we address it here because (a) we we mentioned, this view is addressed - however hastily it is dismissed - in the literature on mathematical

---

[57] The above proof also corrects a typo in $< 1 > 1.2$.

[58] Perspicuity was an important consideration for Wittgenstein. See [Wit83].

explanation, and (b) because it is conceivably a potential candidate for a condition on explanation, and should be acknowledged and responded to in that respect.

Chapter 3

Explanations without science

"Daddy, are you lost?" I asked tenderly.
"Shut up," he explained.[1]

The previous chapter looked at theories of scientific explanations and the possibility of applying them to mathematics. We saw the familiar theories of Hempel, Salmon, Kitcher, van Fraassen, and others. We showed how the theories we are familiar with from the philosophy of science will not produce explanations in mathematics. In this chapter we examine a theory of explanation that has no counterpart in the philosophy of science and discussions of mathematical explanation that are independent of science and its philosophy. We examine a theory that can only work in mathematical contexts. We first look at Mark Steiner's model of explanation in mathematics. Second, though he presents no formal model of mathematical explanation (and believes there is not *just one*), we examine Paolo Mancosu's various claims, which amount to an argument for explanation in mathematics.

---

[1] Ring Lardner, *The Young Immigrants.*

This chapter concludes the “negative” part of our argument where we argue against the extant views of mathematical explanation in the literature as well as the views that merely advocate that there are mathematical explanations.

3.1 Steiner's account

This section evaluates Mark Steiner's account of mathematical explanation. We begin by examining Steiner's argument step-by-step, challenging various steps, and conclude with some general arguments against the account. Steiner claims that there is explanation in mathematics. He argues that explanation is not abstraction, nor is it generality. Rather, he claims, in mathematics explanation involves a certain kind of proof. We find Steiner's account unsatisfying for a number of reasons.

Steiner uses various examples at each stage of his argument to show that there are intuitions about mathematical proofs that need to be accounted for, and that those intuitions are only accounted for by appealing to the fact that there is explanation in mathematics. Following the order of Steiner's argument, the first problem we present for his account is that all the examples he gives which are supposed to motivate our intuitions for explanation in mathematics can be accounted for by appealing to considerations of ease, psychology, and pedagogy. We provide alternative interpretations of the intuitions behind each example, showing that we do not need to appeal to a concept of explanation for any of them. By doing this we show that while the examples can be construed as supporting his theory of explanation, they can more naturally be interpreted as exhibiting something else. Thus we conclude that all the motivating examples can be, in Steiner's terminology, "explained away" as psychological artifacts. We are left with no motivation for (or analysis of) explanation.

We then present a second challenge to Steiner's account. We offered alternative and seemingly more natural interpretations of the intuitions behind Steiner's examples. Nonetheless, Steiner could still claim that if he presents a model which can subsume all the examples, then despite the fact that all the examples are accounted for in other ways, it cannot be a coincidence that one model captures all the cases. His account could even claim the benefit of parsimony, i.e., it appeals to one concept of explanation instead of a variety of psychological ones. So we must examine Steiner's account of explanation more minutely.

Assuming that Steiner's examples are representative of a unique kind of proof, Resnik and Kushner, and Mancosu both challenge the model on the grounds that it does not hold up as a complete account of explanation. So I consider Resnik and Kushner's counterexamples to Steiner's model which seek to show a) there are cases that should count as explanatory but are not covered by Steiner's model, and b) Steiner counts things as explanatory that are clearly not.

Following the discussion of some technical problems with Steiner's account we have independent arguments against Steiner's argument for mathematical explanation, as well as a counterexample.

First let's look at Steiner's discussion: Steiner addresses mathematical explanation in a series of papers. His theory is first developed in [Ste78a], and then used in various ways in [Ste78b], [Ste79], [Ste83a], [Ste83b], and in [Ste00]. In [Ste78b] Steiner uses mathematical explanation as part of an argument for mathematical realism. In [Ste83b] Steiner puts forth an (implausible, I think) argument that Imre Lakatos was a kind of mathematical realist. Steiner uses the existence of mathematical explanations to argue against Lakatos' fallabilist view of the history of mathematics. Lakatos claims that mathematics only provides tentative results. Steiner shows that modern proofs are explanatory in a way that some earlier proofs are not because modern proofs are explanatory in Steiner's sense. Thus there is progress in mathematics - contrary to Lakatos' view that mathematics is merely an ever-continuing series of proofs and refutations.

Steiner gives the only account of mathematical explanation specifically designed to account for the intuition that mathematics countenances explanations that differ from those in other fields that use explanations. Beyond suggesting an example, Steiner does not elaborate on explanations *per se*, rather he argues that there is a class of *proofs* that are explanatory. So, for Steiner, there is not necessarily a unified concept of explanation that can be applied across science and mathematics. This is interesting because typically the motivation for generating a theory of explanation in mathematics is part of a desire to show that mathematics and science are somehow unified. Thus there is an incentive to offer a unified account of explanation in mathematics and science.

For Steiner there is no *model of explanation* that applies to mathematics; rather there is a subset of proofs that explain their theorems. These proofs have specific properties that identify them as explanatory. Steiner's argument for this (from the initial discussion in [Ste78a]) consists of four steps, as follows:

(1) Mathematicians routinely distinguish between explanatory and non-explanatory proofs.

(2) Explanatoriness cannot be identified with abstractness; i.e., the intuition that there are more or less explanatory mathematics cannot be accounted for by saying that some mathematics is abstract, and *as such* we value it above less abstract mathematics because it is explanatory.

(3) Generality cannot be identified with explanation either; i.e., the intuition that there is more and less explanatory mathematics cannot be accounted for by demonstrating that some mathematics is more general, and *as such* explanatory.

(4) The examples in Steiner's discussion of more "abstract" and more "general" proofs that motivate his dismissal of abstraction and generality as theories of explanation can be used as examples of explanatory proofs. We can account for the intuitions behind all the motivating examples by subsuming them within a single model of mathematical explanation.

That is the gist of the argument. Let us examine it step-by-step. Step

(1) asserts that there is an intuition that mathematicians have that routinely distinguishes between explanatory and non-explanatory proofs. This of course remains to be proven. What is uncontroversial however, is that mathematicians often divide proofs into two categories: those they prefer more and those they prefer less.[2] G. H. Hardy, expressing disdain for some mathematics, famously said "There is no permanent place in the world for ugly mathematics." ([Har67]: 1) So for Hardy at least, there is a predilection for aesthetic mathematics (whatever that might be) over ugly mathematics. The mathematician Helmut Hasse once said that "truth is necessary, but not sufficient for real mathematics - what is also needed is beautiful form and organic harmony" (quoted in [Seg03]: 4). Proofs are valued or disliked for a variety of reasons. Exhaustively enumerating these reasons is beyond the scope of this work, but there are various aesthetic and formal criteria that can be marshaled to evaluate the "quality" of a proof independent of its formal validity. David Sandborg's epistemic virtue theory (in his [San97]) enumerates some of the criteria that might factor into a mathematician's judgement about what is to be preferred or shunned in a proof. John Daw-

---

[2] See §3.2 below for some historical debates over the superiority of particular kinds of proofs and styles in mathematics. Mancosu ([Man96]) sees such a division between "causal" and "non-causal" proofs in the sixteenth and seventeenth centuries. Mazzotti ([Maz98]) sees such debates over the preference for analytic versus synthetic proofs as spill-over from religious and political concerns in the Kingdom of Naples. Hadamard ([Had45]: 107) discusses mathematical preferences reflecting political divisions, especially in Felix Klein ([Kle11]: Lecture VI). Mathematical preferences of that period also reflected racial stereotypes ([Row86]) ultimately culminating with the mathematical views of Ludwig Bieberbach (see [Seg03]).

son ([Daw06]) too thinks that there are virtues of proof attractive enough to motivate mathematicians to prove a theorem a second time using a different proof or style of proof.[3] Sandborg cites explanation as one of those virtues. It is telling that Dawson does not cite explanatoriness, despite the fact that in the very article in question he cites some of the literature on mathematical explanation. Dawson does claim that some proofs are more insightful than others in various ways, but stops short of claiming that we reprove some theorems because the known proofs fail to explain. Steiner of course must still establish that explanation is a desideratum in mathematics. Presumably if he finds a model of explanation that subsumes all the cases where a proof is judged superior to its rivals, and there is no more natural interpretation, then his claim will have been substantiated. The remainder of his argument attempts this.

Step (1) thus depends on mathematicians being able to discern explanatory import in their own mathematics. We shall take up this issue again later in this section. But for now we can note that Steiner himself notices that Polya's *Induction and Analogy in Mathematics* ([Pol54]: vol. I) is "a gold mine of examples of mathematical explanation [yet] Polya himself does not discuss the notion" ([Ste78a]: 138). So we wonder just how clear a notion mathematicians have of this elusive concept of explanation if even a mathe-

---

[3] See also Jean Dhombres' [Dho93] for an interesting account of why and how the seventeenth century mathematician Gregory of Saint-Vincent made a practice of offering multiple proofs of the same theorem.

matician as sensitive to methodological concerns as Polya fails to notice that his own work is "a gold mine" of explanations. Like Moliere's character who failed to realize he had always been speaking prose, would Polya be surprised to discover he had been providing explanations all along? If he would be, what becomes of Steiner's contention that mathematicians routinely distinguish proofs that explain from proofs that merely demonstrate?

Step (2) of Steiner's argument denies that abstraction can be identified with explanation.[4] To argue for (2) Steiner gives the following example (one he reuses in [Ste83b]): Take the theorem

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = n(n+1)/2 \tag{3.1}$$

which can be proved by straightforward mathematical induction. The base case

$$\sum_{i=1}^{1} i = 1(1+1)/2 = 1 \tag{3.2}$$

with the induction step

$$n(n+1)/2 + (n+1) = (n+1)(n+2)/2 \tag{3.3}$$

gives us the theorem.

This inductive proof of theorem (3.1) is then contrasted with two proofs that are "more illuminating" and "more explanatory" respectively. The former "more illuminating" proof involves the Gaussian strategy of adding a

---

[4] Steiner claims that Georg Kreisel has endorsed generality/abstraction as the criterion for explanatoriness. In a footnote in the same paper however, he claims that Kreisel did not mean to be taken literally.

sequence to a copy of itself in the reverse order so that the terms match up. Add as follows:

$$\begin{array}{lllll} 1 & +\,2 & +\,3 & +\cdots & +\,n \\ n & +\,(n-1) & +\,(n-2) & +\cdots & +\,1 \\ \hline \end{array}$$

$$(n+1)+(n+1)+(n+1)+\cdots+(n+1)=n(n+1)$$

then divide the sum by 2. Steiner takes this proof to be more abstract since when fully formalized (using techniques from Quine's *Set Theory and Logic* ([Qui78])) it quantifies over series[5] of *sequences* of numbers, whereas the earlier inductive proof merely quantifies over the numbers. Quantification over series of numbers, is more abstract than quantification over the numbers themselves, as any higher-order quantification is more abstract than a lower-order quantification.

Steiner claims that the following "explanatory" proof[6] is *even more explanatory* than the previous proof though not as abstract. It is a purely geometric point-counting argument. Take the nxn lattice of Figure 3.1: Each side has n points. Dividing the lattice in half on its diagonal we get an isosceles right triangle whose two equal sides are the edges of the lattice. The triangle contains $1+2+3+\cdots+n$ points. Keeping in mind that we

---

[5] A series is represented as a sum of a sequence of terms, a list of numbers with addition operations between them, and is thereby a set of a certain sort.

[6] Pythagoras may be the original inspiration for the use of this proof in this context. Compare Aristotle's *Physics* 203a2, and the discussion of the Pythagoreans in [Whe60]: 203-204.

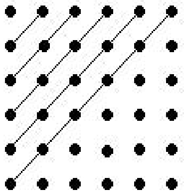

Figure 3.1: Lattice

counted the diagonal (n) twice, we get

$$2\sum_{i=1}^{n} i = n^2 + n \tag{3.4}$$

from which our theorem follows easily. This last proof is said to be more explanatory. But if it is more explanatory, yet not more abstract, then abstraction is not identical to explanation. Rather, Steiner claims, it is not the abstractness, but the fact that there is a formalizable "pictorial aspect" to the proof that accounts for its explanatoriness.[7]

Before I go on to discuss steps (3) and (4) of Steiner's argument I wish to address a number of problems with step (2). First, as Paolo Mancosu points out in a footnote ([Man01]: 113, n11) there is no consensus of intuitions about this case. Gila Hannah for example, finds the Gaussian proof explanatory,

---

[7] See James Robert Brown ([Bro97]: esp. §9) for more on the suggestion that pictures are explanatory.

but not the inductive one. James Robert Brown ([Bro99]: 42) takes the inductive proof to be more "insightful and explanatory" because "induction – the passage from n to $n+1$ – more than any other feature, best characterizes the natural numbers.

Second, while acknowledging that the term needs clarification, Steiner gives us a rather odd definition of "abstract". Steiner claims that abstraction "surely increases in the ascent from first to higher 'order' arithmetic." (136) So a proof invoking or quantifying over a higher-order object is a more abstract proof. But it is hard to see why talking about say, the number 5 is not to invoke something abstract, whereas when we talk about $\{5\}$ we do; or why $\{\{5\}\}$ is more abstract than $\{5\}$ ....

The definition of "abstract" is crucial here. It should first be noted that there is no agreed upon definition of "abstract". And, while we cannot offer a complete definition, we can see that even a partial definition will muddy the conceptual waters and tell us what abstraction is not. Mathematicians frequently refer to something mathematical being abstract, or more abstract. Hilbert's geometry, for example, is said to be abstract, Euclid's is not. Or at least Hilbert's geometry is taken to be *more* abstract than Euclid's. Hilbert's geometry is said to be more abstract, because it does not make reference to some properties of the "entities" in question (e.g., "points", "lines", "planes", etc). Hilbert's geometry doesn't even need the "entities". As he famously remarked that one need only invoke the relations between the "entities" that geometry deals with; the entities could just as easily be tables, chairs, and beer mugs and the relations between them would still hold. The abstract character of Hilbert's geometry is not due to the order of the entities in an arithmetic hierarchy, but rather in the fact that the theory recognizes fewer properties of the geometric "objects" as relevant in identifying them.[8] Steiner does agree that to make a clear case for identifying abstraction with explanation, we would need a definition of abstraction, but he takes the fact that he shows that abstraction is not explanation as releasing him from an obligation to provide such a definition. However, it is hard to see how Steiner's rejection can be warranted in the absence of the definition. Without a definition, the question of which of two objects is more abstract is not settled, and for all we know all the intuitions about explanatory proofs can be accounted for by invoking abstraction. Therefore, we cannot determine whether explanation can or cannot be identified with abstraction. And since, according to Steiner, abstraction is a plausible candidate for explanation, we are still short a good argument for why we should reject it, or why we shouldn't give abstraction a closer look in this context.

Third, here we are concerned with how abstract the proof is, not how abstract the objects within the proof are. If a proof α invoked objects that are more abstract than the objects invoked in a proof β, α is then not more abstract than β. A more-abstract proof is different from a proof involving

---

[8] Stephen Pollard has a much more formal and extensive discussion of "abstraction" in his [Pol87].

more-abstract objects.

Fourth, with reference to Steiner's example, there is a very obvious reason why we might find the dot-counting proof preferable to the Gaussian proof, and the Gaussian proof preferable to the inductive one, that does not appeal to Steiner's notion of explanation. We can appeal to psychological factors to account for the judgement that the geometric proof is more enlightening than the Gaussian one and the Gaussian one more enlightening than the inductive one. As a matter of fact, all of Steiner's examples can be accounted for by appeal to very straight-forward psychological factors.

In all of Steiner's cases, what he calls "more explanatory" is just simpler, easier, or learned earlier in a mathematics education. Look back at the first proof. The first proof involves mathematical induction. Mathematical induction is a concept that one does not find at all intuitive until after one is taught it relatively late, perhaps in college. Certainly it is taught after basic arithmetic. The Gaussian trick was thought up by Gauss, who as the story goes, was seven years old at the time. Clearly Gauss was a prodigy, but today the "trick" can still be appreciated by someone whose mathematical knowledge is not nearly extensive enough to include the principle of, or reasoning behind, mathematical induction. All that is required is an appreciation of addition and dividing in half. And simpler than both of those proofs is the one that uses lines and dots and does not require much numerical ability at all. The dot-counting proof merely requires some appreciation for spatial relationships and the ability to count dots. This is clearly easier (for most people) than manipulating numbers. One need not espouse a Kantian epistemology to appreciate that a spatial intuition seems to be epistemically, pedagogically, and psychologically prior to the ability to manipulate numbers. Certainly counting "concrete" (or easily visualizable) objects (like dots) is prior to adding them up, which in turn is prior to mathematical induction. So the explanatory order here seems to be reduced to the pedagogical order. And while we are hesitant to say that there is a "presupposition relation" here such that knowledge of induction presupposes knowledge of arithmetic, it is hard to imagine a successful mathematics education that teaches mathematical induction before addition, or that can teach addition before counting.[9] However if there were such a successful mathematics education that taught induction prior to arithmetic and counting, we may surmise that for those students the inductive proof would seem more intuitive or explanatory.

We have thus far shown that the examples that motivate Steiner to claim that there is more and less explanatory mathematics can be otherwise accounted for. Thus we are left with no motivation for the mathematicians' intuition that there is explanation in mathematics. Also if there is explanation in mathematics, we still don't know if it can be identified with abstraction,

---

[9] In another case of pedagogical phylogeny recapitulating a subject's ontogeny, the pedagogical order here resembles the historical order. Morris Kline writes: "Mathematics is . . . also a cumulative development, that is, newer creations are built logically upon older ones, so that one must usually understand older results to master newer ones." ([Kli67]: 11)

as we have no a clear analysis of what abstraction amounts to.

Next we come to step (3) of Steiner's argument that the intuition that there are explanations in mathematics cannot be accounted for by identifying explanatoriness with generality. Generality sounds like a plausible candidate for explanation; many unification accounts take generality seriously. Philip Kitcher ([Kit75]) attributes to Bolzano the claim that one of the criteria for an explanatory proof is that the premises must be at least as general as the conclusion, though Kitcher himself ultimately rejects this criterion. Steiner gives some examples to motivate the idea that generality is explanation. But before turning to them we must define the terms "more general" and "general". Steiner suggests (137) that "a partial criterion emerges from considering that some proofs prove more than others." Let us go along with this for now.

Steiner gives three examples involving generality and explanation. The first (which is repeated in this context by [But01]: 328) is the Prime Number Theorem, which concerns the distribution of the primes on the number line. It states that

$$\pi(x) \backsim \frac{x}{\log x} \tag{3.5}$$

where $\pi(x)$ is the number of primes less than x. Steiner addresses two known proofs of the Prime Number Theorem, an analytic proof and a so-called elementary one. Originally conjectured by Gauss, the Prime Number Theorem was eventually proved independently and simultaneously by Hadamard and de la Vallée Poussin. These proofs are extremely difficult and involve the theory of the functions of a complex variable, specifically the Riemann Zeta Function, which has no apparent connection with the Prime Number Theorem. Because of the lack of an apparent connection between the proof and the theorem, an "elementary" proof was sought that did not involve complex analysis. Ultimately one was given simultaneously by Selberg and Erdős that made use of only arithmetic and combinatoric notions. The Selberg-Erdős proof was seen by some mathematicians to "be more in accord with Ockham's razor" ([Ben99]: 231) and as "more elementary" and "more natural" because the solution does not involve concepts that do not occur in the statement of the problem ([Ben99]: 62).

Despite the discomfort of the mathematical community with the original analytic proof of the Prime Number Theorem, Steiner judges (and claims Solomon Feferman might agree) that it is more explanatory. Why? Because the analytic proof proves more, i.e., it is more general. How does it prove more? Because for each given prime number x, the analytic proof gives a better estimate for how much $\pi(x)$ deviates from $\frac{x}{\log(x)}$. The error term in the analytic proof depends on what is known about the zero-free region of the Riemann zeta function within the critical strip.[10] As our knowledge of the

---

[10] The Riemann zeta function is, for some complex number z, $\zeta(z) = 1 + \frac{1}{2^z} + \frac{1}{3^z} + \frac{1}{4^z} + \cdots$. Riemann conjectured that the "roots" of the zeta function - the complex numbers at which the zeta function equals zero - all have real part $= \frac{1}{2}$. "Geometrically they lie on the line 'real part of $z = \frac{1}{2}$' - i.e., a line parallel to the imaginary axis and $\frac{1}{2}$ unit to the right of it." This proof of the Prime Number Theorem depends on the fact that all the zeros are between the imaginary axis and the line $x = 1$. "To prove that they all lie exactly

size of this region increases, the error term decreases, thereby generating more precise values than the Selberg-Erdős proof can produce for $\pi(x) \backsim \frac{x}{\log(x)}$ for each particular prime x.

So the judgement about the comparative value of the two proofs is not related to their respective abilities to explain, but rather to how precisely each one answers questions about the distribution of primes. The nature of the distribution is naturally what is sought in this proof. The analytic proof and the elementary proof both deliver the Prime Number Theorem, but the analytic proof is able to generate finer values *for any given prime number.*

However, while precision is valued by mathematicians and scientists alike, I argue that it should not to be identified with explanation. Science also countenances precision as a virtue. It is hard to imagine how it could not. Though it is an acknowledged role in both the Hempelian D-N model of explanation and in contrastive accounts, we have no account of scientific explanation which has it as its sole aim to increase precision, or sees precision as part of the explanatory project.

Take the case of the trancendental number π. Whereas there is a long history behind behind the computation of the value of π (see [Bec71], [LBB04], [Bla97], [Cas88]), it has never been claimed (to my knowledge) that $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \cdots$ is more explanatory than $(\frac{4}{3})^4$ just because

---

on $x = \frac{1}{2}$ would imply even more precise conclusions about the distribution of prime numbers." Hardy proved that there are infinitely many zeros of the zeta function on the line $x = \frac{1}{2}$, but we still do not know if they are *all* there. ([DH81]: ch. 8).

$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \cdots$ is able to generate a better estimation of the ratio of the circumference to the diameter of a circle than $(\frac{4}{3})^4$. 3 is not less explanatory than 3.14159... it is just less precise. Thus there is a very good reason - that has nothing to do with explanatoriness - why any mathematician would prefer $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \cdots$ to $(\frac{4}{3})^4$ when trying to compute the value of π.

Precision for that matter is also not generality. In both proofs of the Prime Number Theorem we get a distance from $\pi(x)$ to $x/\log(x)$ for each x. But neither proof can be said to prove more than the other. The analytic proof delivers more precise values, not more values.

G. H. Hardy, who is frequently cited by Steiner (though not as giving "explanatory proofs") died just before the Erdős-Selberg elementary proof was published, so he could not comment on it. But, it is interesting to note in this context that he believed that the prime number theorem was a "deep" result that could only be proven with "deep" methods, i.e., complex analysis. In a 1921 lecture he claimed that

> ...We have certain views about the logic of the theory; we think that some theorems, as we say 'lie deep' and others nearer to the surface. If anyone produces an elementary proof of the prime number theorem, he will show that these views are wrong, that the subject does not hang together in the way we have supposed, and that it is time for the books to be cast aside and for the theory to be rewritten. ([Boh52]: 129)

Hardy was expressing an intuition about the analytic proof of the Prime Number Theorem. For him, the motivation for preferring the analytic proof

- so much so that he thought an elementary one was not even possible - was a symmetry between the "depth" of the proof and the "depth" of the theorem. So we see that there are alternative ways to account for the intuition that the analytic proof is preferable without invoking explanation or even generality. Hardy even seemed to have an *a priori* preference for the analytic proof. Or, at least he spoke as if other proofs were out of the realm of phenomena that were in line with his normal intuitions about the rest of mathematics.

So it seems as if the prime number theorem example is a non-starter. We have on the one hand a reason to think that the elementary proof is valuable - namely that it does not invoke complex analysis, though on the other hand the analytic proof provides more precise results. From this example it is also unclear if either proof is more general, or even what Steiner's criterion for generality is.

Steiner's second example involving the generality criterion uses the proof that $\sqrt{2}$ is irrational. The standard Pythagorean proof is if $a^2 = 2b^2$ with a/b in lowest terms, then a^2 and thus a must be even. Thus a^2 must be a multiple of 4, and b^2 and thus b, be multiples of two. Since therefore $a^2 = 2b^2$ implies that both a and b must be even, it contradicts our stipulation that they are reduced to lowest terms. Thus it can never be true that $a^2 = 2b^2$. Q.E.D.

The key point in the above proof, says Steiner, is that if a^2 is even, so is a. To verify this, take an arbitrary odd number, $2q + 1$ and square it. The result must be odd. "Indeed for each prime p, one can separately verify that if p divides a^2 it must divide a also, though the proofs become more and more complex (where $p = 5$, for example, one must square $5q + 1$, $5q + 2$, $5q + 3$, and $5q + 4$ and show that in no case is the result divisible by 5)." (138)

But, Steiner claims, a better proof comes from using the Fundamental Theorem of Arithmetic - that each number has a unique prime power expansion. We can argue that in the prime power expansion of a^2, 2 will appear with an even exponent (double the exponent of the expansion of a), while in $2b^2$, 2's exponent must be odd. So a^2 never equals $2b^2$.

Now, consider the more general result that $a^2 = nb^2$ implies that n is a perfect square. Hardy proves this by assuming a/b is in lowest terms. If a prime p divides b, and thus b^2, it must also divide a^2 and thus a. This is a contradiction. So no prime divides b, and b must be 1, and n must be a perfect square. When $n = 2$ we get Pythagoras' result above, that $\sqrt{2}$ is irrational. Hardy's result is more general because it is not restricted to the case where $n = 2$.

But the more general result is not more explanatory. The proof involving the Fundamental Theorem of Arithmetic is more explanatory, but not more general. Steiner thus argues that the proof invoking the Fundamental Theorem of Arithmetic cannot be more explanatory *because* it is more general, rather on the assumption that it is more explanatory, it must be more explanatory for some other reason.

There are problems with this example. First, we are not given the criterion that was enlisted to make the judgement that Hardy's proof is less explanatory. Later we find out that it isn't derived from an explanatory proof. But asserting that outright would beg the question in Steiner's favor. We propose that there is a simple reason for the judgement that the proof involving the Fundamental Theorem of Arithmetic is superior. The reason is that the proof involving the Fundamental Theorem of Arithmetic makes use of unique prime power expansions (e.g., the prime power expansion of 756 is uniquely $2^2 * 3^3 * 7^1$). But looking back at Hardy's proof we see that in verifying the results for each prime we are required to multiply polynomials, clearly a much more difficult procedure than multiplying integers. Steiner himself notes the complexity of verifying the results in Hardy's proof (138).

So again, Steiner's intuition that one proof explains more than another proof can be accounted for without appealing to any criteria other than simplicity and ease. Multiplication of integers is clearly simpler and easier than multiplication of polynomials.

Though Hardy's proof is more general, it is not as simple as the proof that uses the Fundamental Theorem of Arithmetic. So Steiner is left with a proof that is more general and another that is easier. However, if Hardy's proof is not explanatory we still have nothing motivating the judgement that some proofs explain.

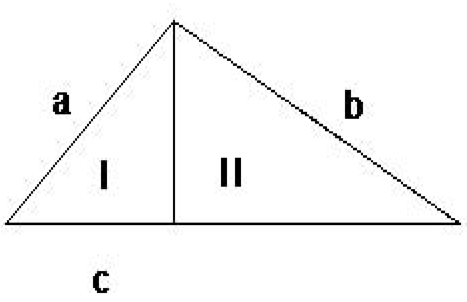

Figure 3.2: Right Triangle

Steiner's third example, using the Pythagorean Theorem, is taken from Polya's *Induction and Analogy in Mathematics* ([Pol54] vol. I). First, we know that the areas of similar plane figures are similar to each other as the squares of their corresponding sides. In particular, any set of three similar figures constructed on the right triangle of Figure 3.2 has areas which can be represented as ka^2, kb^2, kc^2, as in Figure 3.3.

Now, if we can find a trio of similar figures constructed on the sides of the triangle in which the sum of the figures on sides a and b were equal to the area of the figure on side c, then we would have $ka^2 + kb^2 = kc^2$, from which the Pythagorean Theorem would immediately follow.

Thus (surprisingly) the Pythagorean Theorem is equivalent to a generalization of the theorem, and a generalization to any of its instances. But triangles I and II above are obviously similar to each other and to the whole

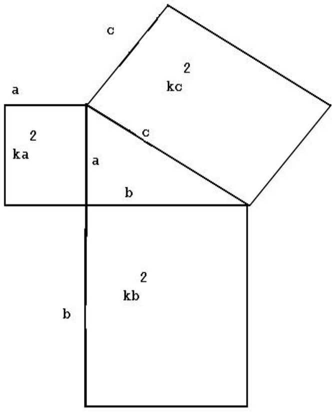

Figure 3.3: Squares of Sides of Right Triangle

triangle; and the whole triangle can be regarded as being constructed on its own hypotenuse, c as in Figure 3.4. Clearly I+II= the whole triangle III, so we have our similar trio, QED.

In this case it becomes clearer what “generality” is. It is deducing a theorem as an instance of a generalization, or as a corollary of a stronger theorem. But generating an instance of a generalization cannot be explanation. We can see this now as the definition of generalization becomes apparent. This type of generalization is the same as the “explanation” we offered in Hardy’s proof of the irrationality of $\sqrt{2}$; and Steiner refused to accept Hardy’s proof as explanatory.

This last case, by the way, does not come with proofs which we can contrast. It merely takes for granted that we see it as “the most explanatory”. Also, despite the fact that there are literally hundreds of extant proofs of the

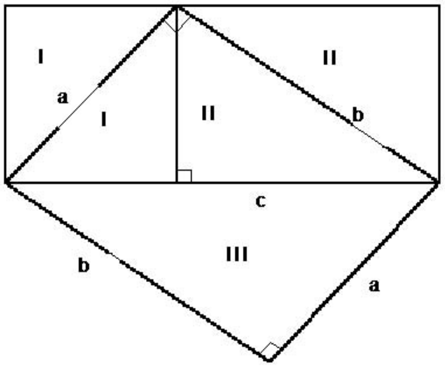

Figure 3.4: Triangle Constructed on its Own Hypotenuse

Pythagorean Theorem, we are not shown the difference between explanatory and non-explanatory ones, or general and non-general ones.

In this respect we agree with Steiner. The concept of generality itself should not be identified with explanation. Generality in mathematics can perhaps be identified with “more widely applicable in mathematics”, not “more explanatory”. Consider, for example part of the definition of a ring. A ring is an algebraic structure that generalizes some of the algebraic properties of some numbers. A ring has the property that “multiplication distributes over addition”. This property is more general than proposition “$x(y + z) = xy + xz$”, which in turn is more general than the proposition: “5(3+7)=5*3+5*7”. The fact that multiplication distributes over addition is more widely applicable (in mathematics) than “$x(y + z) = xy + xz$”, and that is far more widely applicable than a specific numerical equality which is only useful if we are interested in one particular case. The second is applicable to denumerably many arithmetical cases. The first is applicable to all the numerical cases, and also in many contexts that involve a pair of non-numerical operators. Applicability to more cases is a more natural way of understanding what we mean by the ability to generalize.

Thus far we have shown that all of Steiner's examples are judged superior in virtue of their simplicity. I have not given any formal definition of simplicity. However, I believe I have presented reasons to judge each of the above cases as examples of psychologically satisfying proofs rather than explanatory ones, or (putting it another way) proofs whose explanatory virtues lie in their pedagogical value and not in their ability to offer an objective account of some mathematical "phenomena". At the very least, there is no *a priori* reason to choose explanatoriness as the reason for the judgement of the superiority of some proofs over others. And since the concept of simplicity in mathematics is an uncontroversially psychological one and I earlier argued that explanation in the sense of interest to us is objective, or at least has an objective dimension, we should conclude that we are dealing with simplicity, not explanation, in Steiner's cases. However, Steiner can still claim that if he can produce a model of explanation that subsumes all the examples which are judged to be superior, then he has found a rival interpretation to ours that ought to be taken seriously. So let us now turn to Steiner's model of explanation itself.

Step (4) of Steiner's argument is the presentation of his own model of mathematical explanation. He argues that his model captures the intuitions behind judgements of explanation in our above examples - the intuitions which abstraction and generality failed to capture.

Following Aristotle, Steiner claims that something is ordinarily considered explained if it is shown that it *must* be the case, i.e., shown to be necessary. But in mathematics everything is taken to be necessary. To replace necessity, Steiner uses something analogous - characterizing properties. As he phrases it: "My view exploits the idea that to explain the behavior of an entity, one deduces the behavior from the essence or nature of the entity." A characterizing property is unique to a given entity or structure within a family or domain of such entities or structures. An explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem in such a way that it is evident from the proof that the result depends on the property ([Ste78a]: 143).

An explanatory proof always uses a characterizing property of an entity or structure mentioned in the theorem. Moreover, it must be evident that the result depends on the property, where "depends on" is defined as follows: if we substitute another entity in the family which does not have that property, the proof will fail, but if we suitably 'deform' the proof, while at the same time holding the 'proof-idea' constant, we get a proof of a related theorem.

The first point that is noted by some commentators is that the terms "family", "domain", and "structure" are all left undefined. Hafner and Mancosu ([HM05]: 233) complain about the difficulties in assessing characterizing properties. Weber and Verhoeven claim the terms are "vague". ([WV02]: 299). Resnik and Kushner ([RK87]:146-7) stress that Steiner's ideas are framed in terms of the "undefined ideas of a family of mathematical entities and of proof deformation". Resnik and Kushner point out a number of potential problems from the outset. First, they claim that there may be difficulties in determining family lines or the limit of proof deformation. Second, Steiner states that the explanatory proof must refer to the characterizing property. Resnik and Kushner claim that the explicitness of a reference to a characterizing property in a proof is a matter of style. Third, they claim that Steiner's criteria of "making it evident" is relative to a particular mathematician. Finally they argue, the boundaries of a proof are not well defined.

As far as the charge that some terms are undefined, we can offer some preliminary analysis, though it may not satisfy Resnik and Kushner but should be sufficient to work with for the moment. First let us look at the word "family". We can say that for Steiner, many theorems have explanatory proofs. Every explanatory proof can be deformed to give proofs of related theorems. In contrast, non-explanatory proofs cannot. Each deformation yields a proof of a related theorem. A family of proofs is then the outcome of these deformations. Of course, now we need an analysis of "deformation". We will come back to that shortly.

But first, we can see what Steiner means by a "characterizing property" from a remark in one of his later works ([Ste00]). Characterizing properties are the properties that generate explanatory insight. He gives an example in which characterizing properties are distinguished from non characterizing properties. The number 3 for example can be represented as the complex vector $3 + 0i$ or it can be $\{\{\{\emptyset\}\}\}$, or $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$. The first representation yields "explanatory proofs" the later two do not. The "explanatory proofs" or "explanatory insight" he is referring to comes from the fact that we have embedded a real number in the complex plane. His example involves the theorem that the region of convergence of a power series in a complex variable is always a disk. This theorem is used to explain the breakdown of the equality $\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \ldots$ when $x > 1$.

Steiner claims that this breakdown is "mysterious" when we only consider the real numbers. The function $\frac{1}{1+x^2}$ is defined, continuous, and infinitely differentiable for every real number x. So why is its expansion in a power series not as well behaved? The mystery is dispelled (that is to say the explanatory insight revealed) when we take the reals as a part of the complex plane. $\frac{1}{1+z^2}$ has a denominator of zero for $z = \pm i$. "This singularity affects all the points outside the unit circle, including all the reals greater than $+1$ or less than -1." That is, the region of convergence of a power series on a complex plane is the interior of a circle that touches $(1, 0)$, $(0, -1i)$, $(-1, 0)$, and $(0, +1i)$ as in Figure 3.5.

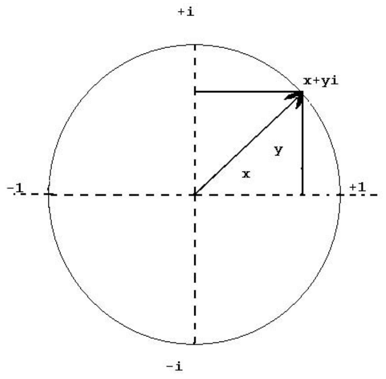

Figure 3.5: Region of Convergence of a Power Series on a Complex Plane

With respect to this characterization of the reals Steiner claims he sees

> no reason except dogmatism, not to accept . . . at face value: the embedding of the reals in a complex plane yields explanatory proofs of otherwise unexplained facts about the real numbers. The explanatory power of such proofs depends on our investing the complex numbers with properties they were never perceived as having before: length and direction. ([Ste00]: 141)

The set theoretic definitions of the number 3 are, in contrast "entirely unilluminating" - they don't allow any new explanatory proofs.[11] Steiner's definition of a characterizing property, however, is reminiscent of a definition Faraday once gave for a magnet as "the habitation of lines of force". It is not an account that provides a penetrating insight into the essence of the magnetic state. ([Wil66]: 113) It is similarly uninformative to characterize

---

[11] Steiner offers this example in the service of an argument for mathematical realism. Freidrich Waismann independently used the same example to advance the same purpose. See [Wai82]: 29-31.

a characterizing property as something that generates explanatory insight, when explanatory insight is defined in terms of characterizing properties.

We may reply on Steiner's behalf that we use "characterizing property" in its traditional sense. That is, C is a characterizing property for a sort of object s if something is of sort s iff it is C. Thus, a characterizing property is a necessary and sufficient condition for s. For Steiner then, suppose there is a family of sorts s_1, s_2, s_n. A characterizing property of s_i would be necessary and sufficient for being s_i and not s_j.

But is a characterizing property for s_i in general a necessary and sufficient condition for s_i? It seems that Steiner's notion of characterizing property is intensionally weaker than necessary and sufficient condition. It is merely a property that is related to the result via a proof such that if the property changes in a certain way, then the result correspondingly changes. The reason Steiner may not want to define characterizing property as a necessary and sufficient condition is because there are certainly many necessary and sufficient conditions for many mathematical objects that are not as easy to "deform" into "related" objects that generate "related" proofs. And without being able to characterize proofs as being deformable in certain ways the model of explanation does not hold up.

So Steiner still owes us a definition of "characterizing property" that is not beholden to the notion of "explanatory insight" and is different from necessary and sufficient conditions. The notion he owes us must be "deformable" in the right kind of way. So let us turn to the notion of "deform" where we will see that an explication of that term is also difficult. Moreover, it is unclear how to reconcile characterizing properties as necessary and sufficient conditions with Steiner's use of characterizing certain kinds of properties like the symmetry properties or geometric properties we encounter in Steiner's examples.

Weber and Verhoeven attempt to explicate Steiner's usage of "deform". In §6 of their [WV02] they assume that there are a variety of kinds of explanation in mathematics. They assume that a variant of Steiner's theory is one of those kinds. Though it is not at all their stated goal, their paper spells out what it means (or should mean) to "deform" a proof. In doing so Weber and Verhoeven end up with a modified version of Steiner's theory that is clearer about what proof deformation is.

For Weber and Verhoeven (and they claim for Resnik and Kushner too) the problem with Steiner's view is that it is confined to questions of the form

> (C) Why do mathematical objects of class X have property Q,
> while those of class Y have property $Q\prime$?

and questions like (C) are not precise enough to generate proper explanations. They prefer the following account, (B):

> (B) The couple (P_1, P_2) explains why mathematical objects of class X have property Q, while those of class Y have property $Q\prime$ if and only if

> (1)P_1 is a proof for the theorem that mathematical objects of class X have property Q;
>
> (2) P_2 is a proof for the theorem that mathematical objects of class Y have property $Q\prime$;
>
> (3) P_1 and P_2 use the same axioms and theorems as premises;
>
> (4) P_1 and P_2 have the same logical rules;
>
> (5) P_1 uses a defining property of X, but not Y, as a premise; and
>
> (6) P_2 uses a defining property of Y, but not of X[12], as a premise. ([WV02]: 301)

This formulation above is preferable because it accounts for the following types of questions:

> (P) Why do mathematical objects of class X have property Q, but not property $Q\prime$?

which they claim is more precisely what Steiner needs. Steiner needs this, they claim, because merely stating that a proof is to be deformed is not sufficient. For Weber and Verhoeven, the deformation part of a Steiner-type explanation consists of actually showing the "before" and "after" of the deformation. So a deformation, at least for Weber and Verhoeven is a pair[13] of proofs that exhibit an initial proof and the result of the deformation. And while this is not quite a definition of "deformation" it does clarify what one should look like.

---

[12] This corrects for what is obviously a typographical error in the original paper.

[13] [WV02] uses "couple" instead of "pair".

Note too that Weber and Verhoeven's conditions (3) and (4) above spell out Steiner's notion of "keeping the proof idea constant". It is however an inadequate definition, as one often needs more than the same axioms and rules of logic. The order of the steps of the proof, for example, is important.

Weber and Verhoeven cite the pedagogical advantage of the unifications enabled by questions in the form of (C). They also claim however that the Lakatosian pedagogical strategy of proofs and the strategy's ability to aid in the discovery of new theorems is an advantage of taking questions of the form of (P) seriously.

It should be clear by now how all of Steiner's definitions form an interconnected web in which each term is definable only by reference to all the others. One consequence of our inability to clearly and independently explicate the notion of "family", "structure", "characterizing property", . . . is that we sacrifice a firm grip on the explanans and the explananda (which the definitions should have provided), for a clear articulation of the explanation relation. For Steiner, to explain is akin to deriving something from "necessary" premises. But no particular fact and no general laws are being explained, and no particular argument is doing the explaining. Rather a "family" of proofs explain a group of related theorems. Thus in Steiner's model, we have a clear notion of the relation between the explanans and explanandum, but a less clear sense of the nature of the proofs doing the explaining and the theorems being explained. Steiner may have recourse to the reply here that "explains" is a relation between a *particular proof* and a *particular theorem*, but one that can only be defined by existential generalization on families (E.g., "there is a family of proofs, and characterizing properties…") But this reply is only partial. Steiner is still left with the task of spelling out how we can generate such a family or recognize the characterizing property.

With the definitions in place, we come back to Steiner's model of mathematical explanation itself. Two of Steiner's examples should suffice to illustrate his point. First, the proof mentioned above that $a^2 = 2b^2$ which uses the prime power expansions of a, b, and 2 is explanatory in Steiner's sense since the prime power expansion of a number is a characterizing property. Moreover, when you substitute 4 or another square for 2, the prime power expansion of 4 (or the other square) will contain 2 raised to an even power allowing

$$a^2 = 2b^2. \tag{3.6}$$

This also generates more general theorems: the square root of n is either an integer or irrational, and the same goes for the pth root. The *generalizable proof* that produces these theorems thus explains. (144)

Steiner's second example involves the explanatory proofs of $\sum_{i=1}^{n} i = n(n+1)/2$. In the Gaussian proof we "characterize the symmetry properties" of the sum $1+2+\cdots+n$. In the geometric proof, though Steiner is not exactly clear how, we "characterize the geometrical properties". Thus the proofs both proceed by utilizing characterizing properties - which non-explanatory proofs do not have. The inductive proof does not contain any characterizing property that is mentioned in the theorem. Moreover, in the explanatory proofs we can vary the symmetry or geometry and get new proofs of new theorems. (145)

Resnik and Kushner, whose primary target seems to be Steiner's realism, present counterexamples to Steiner's account of explanation in mathematics. First, they show that there is a proof that meets Steiner's criteria but fails to explain and then a proof that is intuitively explanatory, but fails to meet Steiner's criteria. Resnik and Kushner begin with Steiner's own intuitively non-explanatory Pythagorean proof of the irrationality of $\sqrt{2}$. They show that there is a characterizing property and the proof is deformable and should thus make the proof explanatory, despite Steiner's claim that it is not. Recall that it depends on the fact that if $a^2 = 2b^2$ then 2 divides a^2. But the same is true for any positive integer substituted for 2. By deforming the proof as such we get the same explanatory results that made the proof with the prime factorization explanatory. The characterizing property here is: "being the least integer x such that any integer that x divides is also divisible by 2.[14] ([RK87]: 147) So we have a deformable proof that meets Steiner's criteria but Steiner himself considers it unexplanatory. Steiner might reply to this that the characterizing property Resnik and Kushner cite involves reference

---

[14] Resnik and Kushner add "Moreover, all but perfect squares divide a^2 only if they divide by a." However, as Rohit Parikh pointed out to me, the dividend must be a product of distinct primes for this to work.

to a particular number, and is thus not a full-fledged property. This reply would be very *ad hoc*, as we are given no prior restrictions on admissible properties. Nor are there any obvious reasons, except to stave off *ad hoc* problems, to place restrictions on what counts as a mathematical property.

Resnik and Kushner then (p 148-150) present examples which they take to be intuitively explanatory, but which Steiner's account would not accept as explanatory. One example they use is a version of the Intermediate Value Theorem: if a real valued function f is continuous on the closed interval $[a, b]$ and if $f(a) < c < f(b)$, then there is an x in $[a, b]$ for which $f(x) = c$. (See Figure 3.6 (from [RK87]).)

Consider the set A of all the t in $[a, b]$ for which $f(t) < c$. A contains a and is bounded above by b; so, but by the continuity of the function, A has a least upper bound x with $x \leq b$.

Since f is continuous on $[a, b]$, $f(x)$ is defined. We prove by contradiction that $f(x) = c$. Suppose that $f(x) < c$. Because f is continuous we can pick a point y to the right of x for which we have $f(y) < c$. But this contradicts the fact that we have isolated all the points y to the left of x. Also, if $f(x) > c$ then for each point t in $[a, b]$ to the right of x we have $f(t) > c$. But again, by the continuity of f there is a point $y\prime$ to the *left* of x such that for all t in $[y\prime, b], f(t) > c$. This contradicts the fact that x is the leftmost point to isolate the points of $[a, b]$ with values less than c.

As far as this proof goes, Resnik and Kushner state that

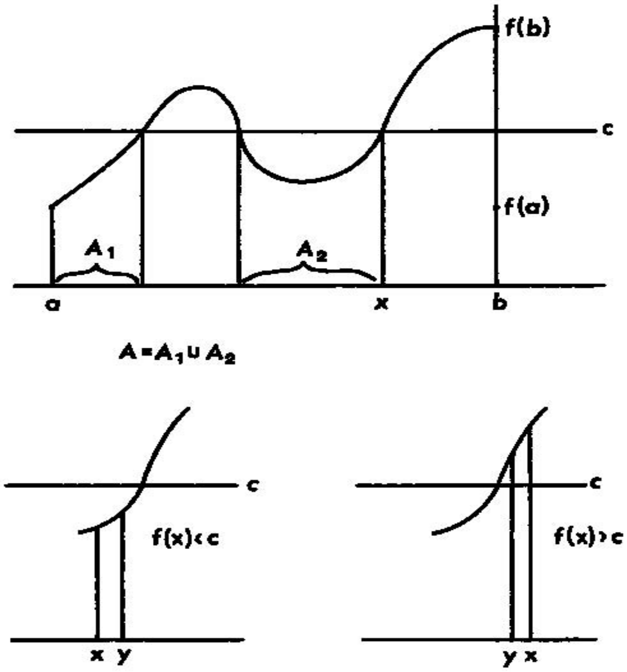

Figure 3.6: Intermediate Value Theorem

> we find it hard to see how someone could understand this proof and yet ask *why* the theorem is true (or what makes it true). The proof not only demonstrates how each element of the theorem is necessary to the validity of the proof but also what role each feature of the function and the interval plays in 'making' the theorem true. Moreover, it is easy to see that the theorem fails to hold if we drop any of its conditions.

Hafner and Mancosu ([HM05]: 223), while not supporting Steiner's account, criticize Resnik and Kushner on two accounts. (Elsewhere ([Man00]: 106) Mancosu says that Resnik and Kushner "concoct two examples aimed at showing that there are proofs that explain but do not meet Steiner's criteria...") Both of Hafner and Mancosu's criticisms would seem to count against Steiner's account and most accounts of explanation in mathematics (including their own notions too, as we shall see later on). They argue that 1) there is no justification for arguing that the examples that Resnik and Kushner give are explanatory, and 2) that "we are not given any hint of what exactly the explanatory feature(s) of this proof are supposed to consist in." Perhaps this is really only one objection: what principled reason was given for choosing some proof as explanatory?

This is a very powerful objection against Resnik and Kushner. It would be a great mistake to simply call some arbitrary proof "explanatory" and criticize a model of explanation for not accommodating it. But we have seen that Steiner gives no principled reason to motivate his examples either, and when we inspected his examples closely, the reason they were deemed explanatory seemed to boil down to mere psychological considerations that do not intrinsically distinguish types of proof. So the objection, if it is to count against Resnik and Kushner's objection, would have to count against Steiner's as well.

A route that Hafner and Mancosu could have taken in criticizing Resnik and Kushner's examples would have been to complain that Resnik and Kusher's second counterexample was a proof by contradiction. Indirect proofs are often taken by mathematicians to be inferior.[15] This would have put the burden on Resnik and Kushner to find a comparably intuitive"explanatory" proof that was a direct proof. Surprisingly though, despite Mancosu's[16] and Hafner's[17] sensitivity to these issues they do not mention it. Samuel John Butchart however argues against the idea that a proof by *reductio* precludes its being explanatory. He argues as follows[18]: take two *reductio* proofs for the irrationality of $\sqrt{2}$. One proof is clearly more explanatory than the other, yet they are both *reductio* proofs. Thus being a proof by *reductio* does not rule out being explanatory. Butchart would still have to meet Mancosu's demand that he find a principled way of distinguishing explanatory from

---

[15] See e.g., the works of Heyting in [BP83] regarding the intuitionist stance on indirect proofs. For a clear statement against proof by *reductio* see [Goo65].

[16] See [Man91] and [Man96] especially ch. 4 where he discusses at length the attitude of numerous mathematicians toward proofs by contradiction.

[17] See Hafner's [Haf99] for his discussion of Bolzano's criticism of indirect proofs.

[18] See [But01] ch. 6 §2 for the details. This example is similarly used by Davis, Hersh, and Marchisotto in the study guide to [DH81] (pp. 331-333) to make a similar argument.

non-explanatory proofs, but on the assumption that Hafner and Mancosu, and Butchart share that there are explanations in mathematics, and that they can be identified intuitively, it is reasonable to take it for granted for now that explanation is not a matter of the difference between a direct and indirect proof.

We take these considerations to all count against Steiner's view that there are explanations in mathematics, and even if there are explanations, that they take the form of Steiner's deformable proofs. But there are further problems with Steiner's account. Steiner begins his argument by claiming that mathematicians have an intuition that mathematics provides explanations. He then swiftly dismisses their notions of explanation and replaces them with his own.

So what Steiner has actually shown if he is right is that Feferman (and Kreisel) do not have intuitions that there is explanation in mathematics. They have an intuition about something in mathematics that divides proofs or other parts of mathematics into two classes - the part that has something valuable and the part that does not. But the class of valuable mathematics is not obviously the class of explanatory mathematics. After all, Steiner just showed that the mathematicians' intuitions did not conform to a sense of explanation. Rather the intuitions conformed to some idea of more intuitive, abstract, interesting, or whatever.

Thus, Steiner's model is unmotivated unless he can show how the mathematicians *he cites* as motivation for investigating mathematical explanation are intentionally giving examples that as such conform to his model when they use the term "explanation". Steiner does not do this. As Resnik and Kushner "find it remarkable that …Feferman proposed an intuitively explanatory proof to Steiner with the challenge to find the characterizing property used." (146)[19]

Steiner, instead of working with the intuitions of the mathematicians, in effect argues that the examples that appear to conform to their intuitions simply fail as explanations. This shows that his model does not account for what *mathematicians* might really mean when they talk about explanation.

We do not require or expect that mathematicians are the best judges of the philosophical aspects of mathematics. However, if the sole motivation for a defense of a particular account of explanation in mathematics is the intuitions of the mathematicians, and the defense and account imply that these intuitions are faulty or misleading, this does not suggest that there must

---

[19] One can make the same argument about the following recollection of Richard Feynman's conversation with Enrico Fermi where the problem in question was the difficulty and complexity of the calculations:

> I had been doing some calculations and gotten some results. The calculations were so elaborate it was very difficult. Now, usually, I was the expert at this, I could always tell you what the answer was going to look like or when I got it I could explain why. But this thing was so complicated I couldn't explain *why* it was like that. So I said to Fermi that I was doing this problem and I started to calculate-he said, wait, before you tell me the results, let me think. It's going to come out like this (he was right), and it's going to come out like this because of so and so. And there's a perfectly obvious explanation. ([Fey99]: 85-86, *emphasis* in original.)

be another model of explanation. Rather it suggests that there is none, or at least that we have no reason to assume there is mathematical explanation. And without motivation from mathematicians' intuitions Steiner's premise (1) cannot contribute to the argument.

We have suggested alternate reasons for evaluating some mathematics as "explanatory", and we do account for the philosophical judgement that some mathematics explains and they appeal to criteria that we assume even Steiner would concede have merit.

We also have the following problem. In [Ste83b] Steiner addresses Lakatos' famous discussion of the history of the formula $V - E + F = 2$ (from [Lak76]). That is, for regular polyhedra, the number of vertices minus the number of edges plus the number of faces is equal to 2. Steiner dismisses Euler's classic proof as non-explanatory, while the modern proof is praised for providing explanations. Euler's proof is clearly simple, elegant, and I daresay, quite beautiful. But it is not deformable, as it only applies to regular closed shapes. The "deformable" proof allows for the varying of one parameter, namely the shape, to get a corresponding proof of what $V - E + F$ is for that shape, in whatever dimension. But in this case Steiner provides no evidence that anyone finds this latter proof more explanatory (or explanatory at all). *Prima facie* it would seem that he would be very hard pressed to. Euler's elegant proof is a few lines long, and simple to comprehend, and upon seeing it, it takes little mathematical sophistication to see why and how it works. The deformable proof goes on for many pages, and gives little or no insight about *why* each theorem that comes out of it works. It could hardly be called explanatory unless you have some prior theory of explanation that includes this type of case - as Steiner has. But Steiner maintained in his "Mathematical explanation" that the justification for assuming that there is mathematical explanation is that there is a shared intuition (among mathematicians) that there are some proofs that are more explanatory than others. This is clearly not the case here.

There is a bait-and-switch here. Once we've "bought in" to Steiner's notion that there are explanatorily better and worse proofs, because of the intuitively superior proofs given in his "Mathematical explanation", we are expected to count less intuitive proofs as explanatory despite the fact that no one would judge (or at least no one has judged) these latter proofs to be explanatory.

A further problem with Steiner's account is that he initially supplies us with intuitions about which proofs are intuitively *more and less explanatory*; e.g., the Gaussian proof is more explanatory than the inductive one, but less explanatory than the "dots proof". But then all these intuitions are promptly abandoned in favor of the Characterizing Property model.

With respect to the Characterizing Property model, Steiner is clear that his "proposal is an attempt at explicating mathematical explanation, not *relative* explanatory value, as the previous criteria.". Then he goes on to tell us that the previous examples are all analyzable in terms of his new criteria. This is certainly going against what he tells us is the mathematical practice of assessing relative explanatory weight (as with the generalizability or abstractness criteria). So again we are "baited" with one notion of explanation and then given another. The initial motivating intuition included the idea that explanations are ordered (i.e., there are more and less explanatory ones). With that in mind, we are then given an account that takes explanations for granted and abandons what we think they must be like.

Lastly, looking back to our counterexample to Kitcher's theory in §2.2, we can show how it also serves as a counterexample to Steiner's theory as well. Recall that we offered a category theoretic proof of the associativity of the cartesian product of a wide range of objects. The characterizing property there of $S \times T$ in the proof that $S \times (T \times W) \cong (S \times T) \times W$ is the fact that it has maps to S and to T, and it is the "best fitting" such object. The proof can easily be deformed to other categories. In our example S, T, and W could represent sets. If so, then our proof is about the associativity of the product operation on sets. But, as we said in §2.3.1, we could just as easily have "deformed" this proof so that it is about Hopf algebras, Lie groups, rings, topological spaces, metric spaces, or many other such mathematical entities, and the deformation will net us proofs of related theorems. We deform the proof by making S, T, and W stand for different mathematical objects. Namely, the new proofs will give us associativity for the Cartesian products of those mathematical structures. In sum, we have all the ingredients for a proof that meets Steiner's criteria for an explanatory proof.

The particularly interesting features of that proof are the same ones that make it so deformable - i.e., that the proof is so abstract. It is also what seems to make it non-explanatory. We look at the proof in wonder of the fact that the operations of say, topological spaces, associate by just stipulating that S, T, or W are sets, spaces, algebras or whatever. So it is hard to imagine Steiner getting many mathematicians to agree that this example is more explanatory (or just explanatory) than the earlier proofs for the associativity of the Cartesian products in the respective branches of mathematics (category theorists notwithstanding).

At the end, Steiner's examples appear unmotivated. There are also counterexamples to Steiner's model he is left to deal with. Also, there are two general problems that seem to count against Steiner's model. First, we have expressed out concerns that the various concepts that Steiner relies on may not be entirely coherent in mathematics.

Second, Steiner's model trades on the following analogy between scientific explanation and mathematical explanation. In empirical cases we inquire what is special about particular contingent facts that brought about some event E rather than some other event F. In mathematical cases we want to know what is special about the assumptions made in some particular mathematical proof that allows us to derive p from the proof, rather than q.

This analogy is intriguing, but it risks missing the main point of scientific explanation. In a scientific explanation there are contingent features of the actual world. If there were no contingent facts in nature we would not even consider this kind of analysis of why some empirical phenomena was observed adequate as an account of explanation. We only consider this kind of explanation because in the actual world, things could have been otherwise. If you push the analogy too far you lose that which we originally found valuable in the empirical cases. And so there are more general reasons we may want to reject Steiner's whole approach to mathematical explanation.

We take all of the above considerations to count against Steiner's model of explanation in mathematics. Together these arguments suggest that Steiner's model is not a viable candidate for mathematical explanation.

3.2 Mancosu on mathematical explanation

Paolo Mancosu's work[20] restarted the discussion about explanation in mathematics. Unlike Steiner, Mancosu does not argue in favor of any particular model of mathematical explanation. Nor does he directly *argue for* the existence of explanations in mathematics. He advocates a number of related theses about mathematical explanation and for a certain kind of approach to the study of mathematical explanation. In this section I extract and analyze five claims concerning mathematical explanation that emerge from Mancosu's work.

> Claim (1) Explanation was and is a mathematical concern for many mathematicians.
>
> Claim (2) The evidence for whether and where there is explanation in mathematics ought to come solely from the practice of mathematics. The intuitions of philosophers (or anyone else) are inconsequential.
>
> Claim (3) Some mathematicians (and philosophers) are "h-inductivists". H-inductivism is a position that suggests strong analogies between science and mathematics, which raises the possibility that both make use of explanations.

---

[20] Mancosu's relevant work on mathematical explanation is his [Man96], [Man99a], [Man00], [Man01], [Man08b], [Man08a], and together with his student Johannes Hafner, [HM05].

> Claim (4) The only acceptable method for discovering the existence and nature of mathematical explanation is by examining *cases* of explanatory mathematics, as opposed to *theories* of explanation.
>
> Claim (5) Mathematical explanations are not homogeneous. There are many types of explanations in mathematics.

Taken together, these claims could amount to an argument for mathematical explanation and also begin to describe some constraints restricting what mathematical explanations can look like. Though the claims seem to play separate roles in Mancosu's advocacy of mathematical explanation, they could be consolidated as an apparently valid argument for a claim he ultimately will take for granted. The argument is as follows: If mathematicians, who are the only ones qualified to assess explanation in mathematics, believe that there are explanations, and they believe both that there are a variety of types of mathematical explanations, and that there are some similarities between mathematics and natural science, and they provide examples, then we can conclude that there are indeed a variety of types of explanations in mathematics. After Mancosu argues for these claims in a number of papers, Hafner and Mancosu conclude that the fact that "mathematicians seek explanations in their ordinary practice and cherish different types of explanations is for us, after working on this topic for so long, so obvious as to require almost no proof." ([HM05]: 118)

In what follows we show how the five claims above, are for the most part, unsubstantiated or mistaken, thereby leaving the conclusion - the existence of mathematical explanation - without a sound argument. Let us examine each of the above claims in turn.

Claim 1: Mancosu's first claim is that explanation was and still is a concern for many mathematicians. His argument begins with the examination of seventeenth century mathematics, mathematical practice, and philosophy of mathematics.[21] Much of seventeenth century mathematics and the accompanying philosophical speculation about mathematics was motivated by the rediscovery of Proclus' mathematical works ([Man00]: 108). As we mentioned in §2.1, Proclus, one of the earliest mathematicians to comment on Euclid's work, specifically states that mathematics addresses why-questions. Mancosu treats Proclus' statement as evidence that explanation was a concern of the mathematicians of Proclus' era. Proclus takes it for granted that the Aristotelian requirement - that we ought to look for the "why" of things - extends to mathematics. He also assumes that we should take the Aristotelian distinction between "demonstrations of the fact" and "demonstrations of the reasoned fact" seriously in mathematics, and *that* is the difference between explanatory and non-explanatory mathematical demonstrations. Aristotle

---

[21] See Mancosu's [MV90], [Man91], [MV91], [Man92], [Man96], and [Man99b] for the relevant historical aspects of the mathematics behind Mancosu's claims about the seventeenth century.

seems to have held a similar view.[22]

Though Proclus's mathematical works were not widely known during the middle ages, some medieval mathematicians[23] seemed also to think about mathematics in light of the above mentioned Aristotelian distinctions. Mancosu claims that the question that the Renaissance mathematicians called the *quaestio di certitudine mathematicarum* can be identified as the predecessor of our question of the existence of explanation in mathematics. A *quaestio* is a feature of Scholastic dialectical analysis wherein a problematic statement is posed, then its possible clarification, with citation of authorities, and then a solution ([Cro97]: 151). This particular *quaestio* asked: what constitutes certainty in the abstract sciences? Due to its Aristotelian legacy this *quaestio* was widely discussed throughout the Renaissance ([Man96]: ch. 1 and the Appendix[24]). Scholars such as Piccolomini, Pereyra, and the Coimbran Jesuits took the position that mathematics is not a science, while in response, thinkers such as Barozzi, Biancani, Barrow, and Wallace took up the cause in favor of mathematics being a science ([Man00]: 110).[25] Biancani went as far

---

[22] Jonathan Barnes ([Ari75]: 110, admittedly not the most orthodox of commentators) cites Proclus as support for the view that Aristotle intended his view of explanation to include mathematics.

[23] See for example, Gersonides' [Ger99]: Book 6, §1, ch. 6.

[24] The Appendix is a translation of Biancani's *De Mathematicarum Natura Dissertatio* of 1615.

[25] Depending on one's interpretation of the work of the above mathematicians, our discussion of causal-talk may be an equivocation on the early modern notion of cause. Medieval philosophers, like Aquinas, sometimes distinguished between the four Aristotelian causes (metaphysical causes) and "demonstrational causes" (usually taken to be epistemic causes); the latter having to do with what causes one to believe in x as opposed to causing x. See especially the case of Richard Rufus in [WA96].

as declaring that mathematics employs formal and material causes ([Man96]: 191). Gassendi, the Port Royal logicians, and others would later embrace the case against mathematics being counted among the sciences. Concurrently, they were involved in blurring the boundaries between epistemology and logic. Gassendi seemed to deny that *any* mathematics could count as an Aristotelian science and the Port Royal logicians complained about the focus on certainty in mathematics over proof and justification. The latter however claimed that there were different kinds of proof: some are geometric proofs with which "our mind is satisfied" and others which were merely demonstrations which cause our mind to "still seek greater knowledge" [about them] ([Man00]: 112). Similarly, direct proofs were "clearer and easier" than indirect proofs. Guldin in his *Centrobaryca* strove to reconstruct classical mathematics so that it did not employ double-reductio proofs and only made use of what was then known as "causal proofs" (though this program did not ultimately succeed, nor was it consistently applied ([Man00]: 114)).

Mancosu sees these discussions as evidence that some mathematicians distinguish between different grades of proof and that these distinctions differentiate between explanatory and non-explanatory mathematics. Mancosu further argues that the history of mathematics, particularly in the seventeenth century displays many genuinely philosophical ideas that speak to modern philosophical concerns. Beside the issue of explanation, these include the nature of infinity, geometry, and the foundational debate in mathematics

([Man99b]).

Mancosu ([Man99a]) argues that interest in explanations is in evidence even after the Renaissance in Bolzano and Cournot's mathematical writings. Bolzano divides proofs into two categories. The first are what Bolzano calls *Abfolge* proofs. They are proofs that are sound as well as valid. Bolzano does not clearly spell out what he means by this kind of proof, though he gives numerous examples. We can derive some of the properties that Bolzano had in mind for them. For example the *Abfolge* relation cannot hold between two false propositions. These *Abfolge* proofs (and only these) serve to "ground" the propositions that they prove. They are said to obey the *grund-Folge* relation. The second category of proof for Bolzano are proofs that are valid but do not obey other constraints placed on the *Abfolge* proofs. Mancosu engages Bolzano's criteria for explanations and concludes that Bolzano's distinction that some proofs are "*Abfolge*" or "grounded" and others are not, cannot be equivalent to the distinction between explanatory and non-explanatory proofs.[26] Despite the fact that Mancosu does not see Bolzano's program ultimately giving us a theory of explanation, he does take the fact that Bolzano was concerned to attempt such a distinction as proof that Bolzano took mathematics to provide explanations.

Cournot too, Mancosu argues, believed that the main goal of the philosophy of mathematics is (or should be) to find explanations, or what Cournot

---

[26] Kitcher ([Kit75]) however deems proofs that are in line with Bolzano's "*Grund-Folge*" relation to be explanatory.

called the “the rational order” or “the true reason” of a theorem. These were notions that Cournot took to be something more than proof. However, Mancosu argues that like Bolzano, Cournot does not provide us with a viable candidate for a contemporary account of explanation in mathematics, especially given the fact that even Cournot himself believed that his own intuition about “the true reason” of some theorems is not formalizable. Nonetheless, Cournot’s interest in finding “the rational order” is evidence that he was interested in finding explanations.

Mancosu also provides two more recent examples (“case studies”) from the history of mathematics to show how a mathematician was motivated to pursue explanations. Both examples are taken from the work of the mathematician Alfred Pringsheim.[27] Mancosu presents Pringsheim’s approach to the theory of analytic functions.[28] Mancosu claims that Pringsheim’s method emerged from explanatory concerns. Pringsheim sought to effect a reduction to the “elementary” by deriving an “organic merger” of Weierstrass’ definition of an analytic function as a system of interconnected power series with the Cauchy-Riemann theory, presupposing only differentiability. The derivation would be effected from the “four species” (the basic arithmetic operations[29]), or more precisely the concept of the arithmetic mean. This reduction to simpler operations would yield “understanding”.

---

[27] The first example is in [Man01], the second is in [HM05].

[28] An analytic function is a function that is locally given by a convergent power series.

[29] “Species” was a common way of referring to basic arithmetic operations. See [Swe87]: 180.

The derivation was carried out by introducing the concept of the "mean value of a function". Briefly: we can represent every complex number as a vector. Vectors can be added and multiplied using familiar rules. All the unit vectors on the complex plane form a unit circle. Each point on the unit circle is described as a value. It is possible to obtain the average value of some function on the points on the circle. Pringsheim then introduces the concept of the mean value of a function and shows how we can determine the average value of a continuous function f on a given circle Γ with radius r. This is done by inscribing increasingly many varieties of polygons in Γ. We can thus compute the mean values of f at its vertices, and then the mean value of f as a continuous function. This can then be extended to circles with arbitrary centers, then to the average of a sum of functions, and equivalently, the sum of the average of functions. This approach, that begins with the concept of "mean value of a function", is able to serve as an "explanation" of the basic insights of Cauchy's theory which are otherwise the "sensational results of a mysterious mechanism" ([Man01]: 111). This is Mancosu's first example from Pringsheim's work which is supposed to show that some mathematicians are motivated to produce explanatory mathematics.

Hafner and Mancosu (in [HM05]) present a second example from Pringsheim's mathematical work on convergent series. It is as follows. We know that some series converge and some diverge. There are tests to determine convergence or divergence for some series. Consider

$$B = b_1, b_2, b_3, \ldots \tag{3.7}$$

Assume there is a sequence

$$C = c_1, c_2, c_3, \ldots \text{and} \sum c_n \text{converges} \tag{3.8}$$

also that there is a sequence

$$D = d_1, d_2, d_3, \ldots \text{and} \sum d_n \text{diverges.} \tag{3.9}$$

We know that if $b_n \leq c_n$ for all n, then $\sum b_n$ is convergent. And if $b_n \geq d_n$ for all n, then $\sum b_n$ is divergent. These are fairly intuitive tests for convergence. If the series C converges, then a series where each member of B has a smaller corresponding number in the convergent series, then $\sum_{b \in B} b$ should also converge. If a series D diverges, then the corresponding series with larger numbers will also diverge. These are known as comparison tests of the first kind.

Comparison tests of the second kind are based on comparing the quotients of two consecutive terms of a series. So for all n (or all $n \leq m$, for some m),

$$\frac{b_{n+1}}{b_n} \leq \frac{c_{n+1}}{c_n} \Longrightarrow \sum b_n \text{converges} \tag{3.10}$$

Similarly for divergence. Again, this is familiar and intuitive for the same reasons mentioned above. Kummer proposed a far more general test for convergence. Let $B_1, B_2, \ldots B_n$ be *any arbitrary sequence* of positive numbers.

Let the a_n be positive integers, Kummer claimed that

$$\lim_{n\to\infty}(B_n \cdot \frac{a_n}{a_{n+1}} - B_{n+1}) > 0 \Longrightarrow \sum a_n \text{converges} \tag{3.11}$$

Now *this* is surprising. When we have corresponding numbers in a series that are lower than those in a convergent series or higher than those in a divergent series, we can easily see how to get convergence or divergence (respectively) from that. But how do we get a test of convergence out of an arbitrary sequence B_n without a known convergent series?

Hafner and Mancosu echo Pringsheim in stressing that this feature of Kummer's convergence test needs explaining. Pringsheim gives two proofs of equation (3.11). Both Pringsheim and Hafner and Mancosu claim that one of the two proofs explains and the other simpler proof merely proves the correctness of the criterion in a simpler way.

The details of the proof are in [HM05] (227-229), so we do not repeat them here. However, to describe the difference we can say that the first proof, the "non-explanatory one" is less than half the length of the second proof, simpler, and involves only basic algebraic manipulation to show that the convergence test works. There are no clever tricks, and most importantly there is no appeal to the arbitrary sequence B_n - the most "mysterious" part of the test.

But Pringsheim claims that his proof, the second one, gives "the true reason why the C_n which naturally occur [in the original comparison test of the first kind] can eventually be replaced by *completely arbitrary positive* numbers B_n." (229) It proves that equation (3.11) is true by showing how equation (3.11) can be seen as reduced to a previously solved problem, i.e., it can be reduced to equation (3.10) or a comparison test of the first kind. Hafner and Mancosu state that Pringsheim aims at a more "global explanation" by embedding this test for convergence in a "reorganized theory".

Hafner and Mancosu use these examples from Pringsheim's mathematical work to show how Pringsheim was interested in providing explanations in mathematics. The examples we have just seen, from Proclus through those of Pringsheim, are the sum total of the evidence that Mancosu presents supporting claim (1) - that mathematicians were (and are) concerned about explanation in mathematics.

To be sure, as far as the history of mathematics is concerned, Mancosu is very clear regarding the extent to which we can make use of it to understand mathematical explanation. He says that

> it would be ludicrous to think that we can simply repropose the issue as we find it discussed in the early modern period. In particular most of these debates on the causality of mathematical proofs are framed in a context that depends on an Aristotelian account of causality and a logical development which did not go beyond syllogistic logic. ([Man99b]: 120-121)

So we cannot reasonably pick up the debate where the Renaissance and early modern mathematicians left off. However, he continues: "by looking at the historical development, we can be motivated to articulate a problem that can be relevant for our understanding of past and contemporary mathematics" (121). Mancosu seems to argue that as long as history records strong interest in the questions of mathematical causality and mathematical explanation we have sufficient evidence to substantiate the claim that the need for explanations or the intuition that there are explanations drove much of the mathematical practice in certain periods of the history of mathematics. If there are explanations in mathematics then we can hope to come up with a satisfactory account of what they are like. However, even Mancosu acknowledges that the history cannot tell us what a good account of mathematical explanation is, for "the philosophical treatment has to be carried out by employing contemporary conceptual tools."(121)

I want to argue, contra Mancosu, that the history of mathematics does not substantiate the claim that mathematicians looked for explanations in mathematics. I do so not by challenging Mancosu's account of the history of mathematics. Rather I question the lessons that Mancosu draws from it and I offer an alternative reading of the debates and positions summarized above.

In the Forward to Jaakko Hintikka and Unto Remes' *The Method of Analysis* ([HR74]), the historian of science John Murdoch warns of a particular danger in certain kinds of writing about the history of science and mathematics:

> Some historians of science may feel a certain amount of uneasi-

> ness, if not alarm, at the prospect of the 'translation' or analysis of a distinction arising in ancient science in terms that derive ultimately from contemporary logic and analytic philosophy. This will, the complaint reads, not present things as they were; distortion will be all but inevitable. ([HR74]: x)

I believe that Mancosu is presenting us with one of the inevitable "distortions" that Murdoch cautions against. When we take our contemporary concerns and assumptions about explanation and assume that despite the fact that they were not articulated as such, the ancient and Renaissance mathematicians were really addressing our issues or forerunners of our issue, we face the prospect of conflating their concerns with ours.

We agree with Mancosu that in the cases that he presents, mathematicians were often dissatisfied with the standard of proof in mathematics or the standard of proof in some sub-branch of mathematics. Already in Plato's time we find that there had been a considerable dissatisfaction with the traditional methods of Greek geometry. This "is suggested by a certain statement of Archytus of Tarantum, a friend of Plato, to the effect that arithmetic, and not geometry alone gave valid proofs." ([Hux65]: 148) Archytus seems to have written a whole book about it. And Aristotle seems to have agreed with him ([Apo52]: 81). Thus mathematicians often sought other, better, methods of proof. So it is not surprising that we find alternative and new experiments in proof-style, and discussions of rigor, in mathematical contexts throughout the history of the discipline. But there is nothing to indicate that these alternative styles of proof or alternative proofs were necessarily quests for explanations. Rather, all the discussions from the history of mathematics that we have mentioned can be construed as pleas for proofs that are easier, or "foundational", or more in line with some psychological or ideological reality than the proof methods that mathematicians had inherited from their predecessors. For example, Judith V. Grabiner ([Gra81]) considers many social and philosophical issues leading to the rigorization of the calculus including Berkelean attacks on the notion of infintesimals, and the mocking critique that the calculus only worked "by the compensation of errors." The mathematicians' relatively new function as pedagogues also stimulated them toward better means of exposition. The desire to reduce the calculus to algebra - or a simpler familiar branch of mathematics - also motivated the drive to find a more rigorous calculus. Both Rav ([Rav99]) and Dawson ([Daw06]) suggest many reasons why a particular style or a particular new style of proof may be used.

Consider the following thought experiment. Imagine a world where modern logic was first done only with trees. Someone came along later and described the method for derivations. Would Mancosu then claim that these proofs are "explanations" or that these proofs were being used for explanatory purposes? Would proving via derivation be explanatory in a way that trees are not? It is true that a derivation may help some individuals comprehend a theorem better than a tree can. But their mere use would not indicate a program throughout logic whereby it was apparent that logicians were interested in offering non-psychological mathematical explanations.

In a similar light, we can look at each of Mancosu's examples and construct plausible stories to tell about what each mathematician is trying to accomplish that does not appeal to anything we would call explanation. In the case of the debates of the seventeenth century mathematicians, we can interpret the search for "causes" as an approach to winning Aristotelian respectability for mathematics. This is clearly articulated by Mancosu with respect to many Renaissance mathematicians. The mathematicians went about putting Aristotelian constraints on mathematics, despite and because of the metaphysical or epistemological consequences. That is, they argued that their proofs are all "causal", or ought to be.

This question of whether or not mathematicians were genuinely interested in causes and explanation in the Renaissance is ultimately a question about the background theory that the mathematicians in question brought to their investigations. What was in the forefront in their minds? Was their primary concern about mathematics or about epistemology? Were the mathematicians assuming that mathematics provided explanations or were they assuming that epistemology and methodology is uniform across mathematics and science? Their motives are crucial. If the Renaissance-era mathematicians were mainly interested in providing explanations, then Mancosu is correct - explanation was a concern of mathematicians, and it is worth working out at least their account of what explanation is. However, if their primary goal was to secure a uniform epistemology, then the "explanations" are secondary to their programs, dispensable, and most likely not explanations. They are the mathematicians' best attempt at a mathematical epistemology, eagerly discarded for a better epistemology. It seems implausible that the mathematicians started with the intuition that there are causes in mathematics and then hit on the Aristotelian paradigm to account for the intuition. More likely, they were faced with the constraints of a seventeenth century Aristotelian epistemology and needed to somehow subsume mathematics within it. To make his case Mancosu must provide us with evidence of the former. He does not.

The question of the mathematicians' intent is far from trivial here. Mathematicians do many things for many reasons. For example, they hold conferences for *institutional* reasons. They prove theorems for *mathematical* reasons. Perhaps, they seek "explanatory mathematics" for *psychological* or *epistemological* reasons, not mathematical ones.[30] Scientists too hold conferences for institutional reasons. They collect data, make predictions, and offer explanations all for scientific reasons. They often represent this data graphically, as opposed to numerically, for epistemic or psychological reasons. The question of the real motivation behind explanation seeking - epistemological or mathematical - goes directly to the heart of our question: what is the

---

[30] This phrasing comes from [San97]: 2, though our conclusions obviously differ.

nature of these "explanations"? What function did the mathematicians see these "explanations" serving. Were they offering stories about "causes" or explanations in mathematics for mathematical reasons, as Mancosu would have us believe, or were they doing so for epistemological reasons? If the desire for a coherent epistemology motivated the mathematicians' search for causal proofs (and similar mathematical devices), then there is no reason to assert that they had any real interest in explaining mathematical results, but rather an interest in fitting the epistemology of mathematics together with the going epistemic framework of the natural sciences - to get a unified theory of knowledge. Many philosophers of mathematics today share the same philosophical concerns and attempt to reconcile the semantics and epistemology of mathematics with the semantics and epistemology of the natural sciences. If there was a mathematical motivation behind the search for a system of causal proofs and the like, *then* we can say that mathematical practice illustrates the search for explanations. Importantly, we can note that as much of Mancosu's own research on the history of mathematics shows, mathematicians of the seventeenth century were extremely sensitive to philosophical issues of the time. It is also difficult to overlook Bolzano's philosophical sensibilities when reading his work. So we know that it is not implausible to attribute philosophical or psychological motives to the mathematicians in addition to mathematical ones. There is thus a way to look at the mathematical work that Mancosu uses to prove that explanation was a concern for mathematicians without seeing explanation as a mathematical concern, but rather as an outgrowth of philosophical or psychological presuppositions. If so, there was never a mathematical concern for explanations, but rather, for solving Renaissance-era *philosophical* problems.

Against our re-interpretation, Mancosu can argue that all we can use to judge the motives of earlier mathematicians is the evidence we are given. We have to take the motives of the mathematicians at face value - so we must assume that they were interested in explanation. But if we are to take the mathematicians literally at their word, then we would conclude that what really motivated them was their expectation that there were formal causes in mathematics (whatever those are) and not explanations; but Mancosu is not arguing that mathematics provides formal causes, but rather explanations.

The reason we have to think that no mathematician today would think in terms of explanation is that few mathematicians today think (or at least write ([Avi06]: 106)) in terms of explanation. Unlike scientists, mathematicians generally do not describe the point of their work as providing explanations. Or at least, to the extent that serious mathematicians nowadays consider the epistemological status of mathematics, the debates that raged on in the seventeenth century have been replaced with new debates about *epistemology*, not debates about the *explanatory status of mathematics* or about explanatory proofs. For example there are some mathematicians who consider Lakatos to have provided an interesting account of mathematical epistemology (e.g., [Her79] and [Cor03]). The debate that these mathematicians perpetuate is about epistemology, not explanation. Almost no mathematician today says that we ought to be looking for anything resembling a cause of theorems. It was only when causes were the mark of epistemic certainty[31], that there was a need to fit mathematics into this paradigm. We no longer think that causation is *more certain* than mathematics. Consequently causation is no longer part of the mathematician's vernacular. It is only among mathematicians (actually, mostly philosophers of mathematics) who are influenced by strains of empirical mathematics that it is even suggested that mathematics is not more certain than the sciences. And even among those who tout the empirical or quasi-empirical methods in mathematics (e.g., Tymoczko ([Tym79]) and Horgan ([Hor93])), there is still to my knowledge no one who has seriously argued that mathematics makes use of causes. So it seems hard to concede to Mancosu that the *quaestio* (or related issues) in Renaissance-era mathematics was about whether or not mathematics provides explanations. Rather it was about epistemology in general, and the epistemology of mathematics in particular.

Also, praising Guldin's program of removing counter-intuitive double-reductio proofs from classical mathematics ([Man00]: 113) as an explanatory program certainly stretches the concept of explanation. Eliminating some

---

[31] Given an Aristotelian epistemology, causes were the mark of certainty because causes were identified with explanations. All the standard commentators on Aristotle (e.g., Annas, Barnes, Hocutt, McKirihan, and Moravcsik) take the four causes to be the four "becauses". See David-Hillel Ruben's [Rub90] for more on this.

kinds of proofs, like the kind Guldin eliminated, does not indicate Guldin believed that every “acceptable” proof explained their respective conclusions. It means that Guldin felt that some proofs were unsatisfying, difficult, convoluted, counter-intuitive, confusing, inconclusive, or at the extreme, wrong. Mancosu should be able to derive nothing more about explanation from Guldin’s program.

Lastly, when we look at Pringsheim’s proofs a bit more closely we find the same problems that we found when we examined Steiner’s examples. The appearance of explanatoriness in both of the Pringsheim examples can be accounted for as the desire to articulate a result as an outgrowth of simpler concepts, or concepts that are learned earlier in a mathematics education, are more familiar to mathematicians (or even the general public), or were previously established. In the first case, when we looked at the Weierstrass/Cauchy-Riemann synthesis, it is obvious that Pringsheim was attempting to reduce the complicated and unfamiliar concepts of complex analysis to the familiar concepts of arithmetic (i.e, averages). He showed how we can derive the desired results in complex analysis by extending the concept of an average. Arithmetic and averages are merely mathematical operations that are learned earlier than anything else previously used to articulate these facets of complex analysis. Because they are learned earlier, they are thereby more “deeply ingrained”, and more widely known. Being more widely known makes it more “universal”. It is in this sense that they "explain".

In the case of Kummer's convergence test, Pringsheim proved that the "mysterious" test of the convergence of an arbitrary sequence reduced to a test of convergence of the first (or second) kind. Convergence tests of the first and second kinds were well known. Thus Pringsheim was engaged in a program to keep convergence tests familiar and show that Kummer's new tests were reducible to the old ones.

None of Mancosu's cases show that there are and were mathematicians whose program *must* be characterized by a search for explanation. The cases merely suggest that there were factors influencing the mathematicians - an Aristotelian epistemology perhaps, or considerations of ease and familiarity - which were unrelated to demonstrating the theorems in question.[32]

To summarize: as proof for Claim (1) Mancosu provides us with a collection of mathematics that 1) in the case of the Renaissance mathematicians, attempts to force mathematics into an Aristotelian epistemology, 2) in the case of Pringsheim's examples, are merely clever methods of taking simple ideas and deriving more complicated results in ways that make it obvious that the complicated ideas arose from the simpler ones, and 3) in the case of Bolzano, are likely proto-foundational intuitions that comply with an aesthetic intuition of some mathematicians which Bolzano himself was unable to formalize. In most of the cases Mancosu himself states that (or shows

---

[32] We could argue similarly for the case of Pick's Theorem analyzed by Sandborg in [San97].

how) none of the programs he has presented 1) work as theories of explanation (e.g., Bolzano) 2) were consistently applied by the mathematicians themselves (e.g., Guldin), or 3) are unformalizable feelings about mathematics - according to the mathematicians who proposed them (e.g., Cournot and Pringsheim). I thus find it difficult to agree with Mancosu that these are all budding *explanatory* programs trying to emerge. So to the question Mancosu asks ([Man01]: §1.3) "Is explanation a novelty in the philosophy of mathematics?", I would suggest that the answer is yes. If the history is properly interpreted and "explanation" is properly understood, the question about the existence of explanations is a relatively new question in the philosophy of mathematics. Explanation is not obviously a significant concern in the history of mathematics either.

Claim 2: Mancosu's second claim is that the evidence for whether and where there is explanation in mathematics ought to come only from mathematical practice. That is, the only evidence we should accept regarding mathematical explanation should be from the intuitions of mathematicians, the mathematical community, and the history of mathematics. The intuitions of philosophers or anyone else are inconsequential as they cannot provide the "basic evidence" for the existence and nature of explanations in mathematics.

Mancosu first expresses this when criticizing Resnik and Kushner's counterexamples to Steiner's model of explanation. In the previous section we saw Resnik and Kushner use a version of Rudin's proof of the intermediate value theorem. They also used Henkin's proof of the completeness of first-order logic, both as counterexamples to Steiner's model of mathematical explanation. The counterexamples were put forth to show that there are proofs that are intuitively explanatory, but Steiner's model would not count them. (Mancosu contrasts Resnik and Kushner's proofs with Pringsheim's proof of Kummer's convergence test which Pringsheim said is explanatory but Steiner's model cannot accommodate.) Mancosu dismisses Resnik and Kushner's counterexamples as follows:

> . . . what justifications are put forth by Resnik and Kushner for the classification of their examples as indeed (intuitively) explanatory? Besides simply claiming that these proofs "would seem to qualify as explanatory if any do" ([RK87]: 149), it is contended with some - albeit rather vague - reference to mathematical/logical practice that Henkin's proof "is generally regarded as really showing what goes on in the completeness theorem and the proof idea has been used again and again in obtaining results about other logical systems" ([RK87]: 149). . . .
>
> For counterexamples to Steiner's theory to carry real weight they would have to be much more closely related to mathematical practice. . . . mathematicians *often* describe themselves and other mathematicians as explaining. And their judgements concerning explanatory vs. non-explanatory proofs (and other varieties of explanation in mathematics . . .) has to figure as the basic evidence, however subjective or context dependent they may be. Claims to the effect that certain proofs are explanatory come from within mathematics, not from philosophers of mathematics. ([HM05]: 223)

It seems that for Mancosu, the only way for philosophers to contribute to the discussion about mathematical explanation is by amassing a list of mathematical statements that mathematicians have claimed was, or treated as, explanatory and then tease out conclusions from that list.

This is an important claim. It speaks to the nature of the discussion that philosophers can have about mathematical explanation, and the nature of evidence that can be admitted in to the discussion.

This is really only a claim about *the nature* of explanation in mathematics, because it presupposes their existence. Presumably Mancosu would have to accept as "basic evidence" pronouncements from within mathematics denying that there are explanations in mathematics. But without much discussion about mathematical explanation, we are not likely to find many mathematicians claiming that there are none. The lack of debate on this issue, incidentally, is telling. Without a visible debate on explanation it is difficult to understand the claim that the mathematical community takes explanations seriously. It is certainly not so prima facie obvious that there are explanations in mathematics as to render discussion unnecessary. And even if the existence were to be uncontroversial, the nature of explanation is not.

But aside for the fact that this claim presupposes the existence of explanations in mathematics, and regardless of the lack of debate about explanations, there are many intrinsic problems with the claim.

Our first problem is that this leaves little room for an evaluation of the claim. How do we evaluate whether or not 1) mathematical practice is mistaken in its attribution of (some piece of) mathematics as explanatory, or 2) it failed to describe something as explanatory, or 3) it failed to proceed as if something was an explanation? Are only mathematicians entitled to claim that other mathematicians or the mathematical community are mistaken about explanations? Can a "working mathematician" argue with established mathematical practice about the existence of explanations, or about particular cases? If so, why can't anyone else? The deck seems to be stacked against the claim even being evaluated.

Mancosu does provide us with a procedure for excluding some claims as *theories* of explanation. For example, in Nagel's *The Structure of Science* ([Nag61]: 16) his first example of explanation is a mathematical example. Mancosu denies that Nagel's example can represent a conception of mathematical explanation. Without going into the details, Mancosu rejects Nagel's conception on the grounds that given Nagel's approach, we would ultimately be able to derive the conclusion that for some As and Bs, A can explain B and B can explain A. This would be a legitimate reason to reject any theory of explanation ([Man00]: 104). But Nagel was not a practicing mathematician, so it is not clear why Mancosu takes his pronouncement seriously in the first place. However as we noted above, Mancosu rejects Bolzano's and Cournot's concepts of "explanation". Both Bolzano and Cournot were practicing mathematicians ([Man99a]). So Mancosu does allow our prior conception of explanation to impose constraints on to the examples and assessments of mathematical practice, and ultimately offer judgements on them.

Moreover, it is clear that Mancosu argues that philosophers may *interpret* mathematics and the claims of mathematicians. As we saw when we looked at Mancosu's first claim, he certainly deems it acceptable to attribute explanatory motives to mathematicians without their explicit acknowledgement of the fact that they are offering explanations. He further claims that there are a variety of expressions we can interpret to mean "explanation". For example, if a mathematician refers to "a better understanding", a "satisfying reason", "an account of the fact", ... then he is referring to explanation. ([HM05]: 218) I would certainly think that this claim needs proof. Yet Hafner and Mancosu take it for granted that when a mathematician uses one of those phrases, he or she is talking about explanation.

Furthermore, Hafner and Mancosu have little problem taking Steiner's work seriously. The bulk of their paper ([HM05]) analyzes Steiner's "Characterizing Property" model of mathematical explanation. And neither Steiner nor his model are integrated into mathematical practice, certainly no more than Resnik and Kushner and the evidence they provide. Hafner and Mancosu are not entitled to pick and choose their explanatory mathematics. Either Steiner's examples and hence his model is dismissed out of hand, or Resnik and Kushner's examples should be accepted.

Steiner may not be providing “basic evidence”, but nonetheless his essay on mathematical explanation is motivated mostly by examples he found, together with examples that have been only circuitously endorsed in conversation by some mathematicians as being somehow superior. This hardly constitutes an unqualified endorsement from “mathematical practice” about the explanatory status of some mathematical proofs.

But the case of Steiner aside, we have a second problem with claim (2). It is clear what Hafner and Mancosu mean by “practice”: if there is some subset of mathematicians, including apparently subsets with as few as one member, who claim that what he or she is doing in any mathematical context is explaining, then that is evidence for mathematical explanation, and evidence for the nature of mathematical explanation. It is only after amassing such evidence that philosophers may extrapolate a theory (or theories) of mathematical explanation from these statements. This is obvious from Mancosu as he mostly uses individual mathematicians and their examples for illustration.

But this seems extreme. After all, in science we do not expect that only scientists are qualified to judge when they have come across a scientific explanation. Anyone sufficiently versed in the relevant field can evaluate whether Lanchester’s square law explains the course of a battle, or the existence of some gene complex in a species explains some aspect of the species’ phenotype. Any mathematician well enough versed in the meaning of “explanation" should be able to make judgements about whether some mathematical proof is explanatory. And to the extent that an individual needs some analysis of what an explanation is, mathematicians are in no better a position than anyone else, especially someone with somewhat refined ideas about necessary or sufficient conditions for explanations in general.

Let us assume that "explanation" is a philosophical concept, and not a psychological one. Why do we presume that mathematicians have privileged access to mathematical explanations? Perhaps the community of mathematicians as a whole can be relied upon to converge on a definition of explanation, but an individual mathematician can be using the word in a variety of ways that even Mancosu would concede is inappropriate in this context. Mancosu stresses (as did we in §1) that there are many uses of "explanation" that are not the kind philosophers have in mind. The philosophy of science literature contains numerous descriptions of cases that we would not include in a conceptual analysis of explanation (see e.g., [Man01]: 100, [HM05]: 217).

The fact that a mathematician says that something is explanatory is not evidence that it is explanatory in the philosophical sense. It is as much evidence of the fact that it explains in some philosophically interesting sense as it is that it explains to *him*. And as Hempel once pointed out, we would not accept a proof that proved only to some particular mathematician, why accept an explanation that only explained to a particular mathematician? I would suggest that what might count as evidence is a *consensus* in the mathematical literature that there are some things that really explain.

We note that there is little disagreement on the explanatory role that Newtonian mechanics played in physical theory, or that evolution plays in biological theory, or game theory plays in economics. Should we not expect the same level of consensus and acknowledgement about explanations in mathematics? We cannot believe that mathematics is so much more subjective than science that it is harder to come to a consensus on what sort of ways we account for, or explain, mathematical phenomena, or what sort of things count as explanations.

We can examine a recent event in mathematical practice for guidance. The famous computer proof of the 4-Color Theorem inspired much debate in mathematics and the philosophy of mathematics. The 4-Color Theorem is now an accepted part of mathematical practice, i.e., it is a statement accepted by the mathematical community as true. But it is only so because of the acceptance of the practice of computer proofs by the community of mathematicians. Until this method of proof was accepted it was a curiosity. It was a way by which a trio of mathematicians established a theorem. Only given the [immediate subsequent] debate within the mathematical and philosophical communities and ultimate acceptance of the method, do we now consider theorems that are proved by computers to be established.

Do we have the same sort of consensus with explanations in mathematics? I think not. Mathematicians can go their entire careers without using the word "explanation" in a mathematical context. They can formulate proof after proof and never be asked by a colleague if they have explained something, or if or how their proofs are explanatory. *This* lack of interest in explanations seems to represent mathematical practice.

Additionally, when it comes to explanation or other details of mathematical methodology, it is legitimate to ask what philosophers are to make of philosophical and methodological pronouncements by mathematicians.

Perhaps we can get some guidance from Albert Einstein. In the opening words of his 1933 Herbert Spencer lecture, Einstein said with respect to physics

> If you want to find out anything from the theoretical physicists about the methods they use, I advise you to stick closely to one principle: don't listen to their words, fix your attention on their deeds. To him who is a discoverer in this field, the products of his imagination appear so necessary and natural that he regards them, and would like to have them regarded by others, not as creations of thought, but as given realities. ([Ein54]) [33]

John Barrow writes similarly of mathematicians "If you want to know what mathematics *is*, a mathematician is probably the last person to ask." ([Bar04]: 82) Methodological and metaphysical questions are not necessarily best judged by those closest to the field.

---

[33] Two things: First, the story behind the translation of this essay is lengthy. See [Hol81]: n2 for details. Secondly, this passage suggests an obvious absurdity which Einstein addresses in the following two paragraphs of the lecture.

While there have been some mathematicians like Frege (and perhaps Barrow himself) who have had a keen sense of the philosophical and methodological aspects of mathematics, I am inclined to agree with Barrow. Before we admit Pringsheim's claim of the explanatoriness of his proofs into our store of "basic evidence" for explanation in mathematics, recall Pringsheim's pronouncement on the earlier "non-explanatory" proof of Kummer's convergence test: that it only "proves the correctness of the criterion *a posteriori* in a simpler way." (in [HM05]: 226) Are we to value every offhand remark of every working mathematician so highly that we will assume that their pronouncements are the last word on mathematical *a posteriori* as well? Shall we query mathematicians for *a posteriori* proofs because it is part of the informal lingo of some mathematicians? So the fact that Phrigsheim analyzes the two proofs in the way he does, is not evidence that his pronouncements on them reflect mathematical consensus or methodology. The claim of explanatoriness could be equally "offhand".

Mathematical practice is only a valuable guide to methodological and philosophical issues as long as mathematicians agree on an accepted definition or are concerned to articulate specific concepts in the same way that they are concerned about, say, the definition of "limit". Otherwise mathematicians are not speaking *qua* working mathematicians but rather in a non-mathematical vernacular. And the idiosyncratic informal vernacular of some mathematicians should not be a guide to mathematical methodology, or an indication of the methodological or epistemological concerns of the working mathematician.

There have of course been temporary exceptions where mathematical practice allowed for vague definitions for as long as they proved fruitful. "Continuity" was used before a precise definition was given. But in cases of undefined terms, there was always agreement amongst mathematicians that they needed a definition, and finding one was an active area of mathematical research. The definition of "continuity" was eventually given by Cauchy and was considered an impressive achievement.

But there must be constraints on what is considered the relevant vocabulary of the working mathematicians. After all, not every term is intended by the authors for a precise definition. For example, when a mathematician claims that he has a "remarkable proof" or an "interesting proof"[34] or whatever, we do not then insist that the working mathematician has a definition of "remarkable" or "interesting" or whatever. We do not take such pronouncements as basic evidence of anything. But are we, as good analytic philosophers, required to produce a conceptual analysis to see what else fits the pattern of "remarkableness" or "interesting"? Do we then insist that we can segregate proofs by whether or not they are remarkable or interesting? Certainly no such thing has been claimed by practitioners (*any* practitioners, to my knowledge) of mathematics. Nor have philosophers suggested this, not

---

[34] But see [Cor03] for a discussion of "interesting", though even he claims after a brief discussion that it is too vague to consider.

even the “mavericks”. We have no reason to believe that in mathematical contexts the word “explanation” is to be taken in any technical, mathematical, or philosophical sense whatever, certainly not an objective one.

Perhaps a thought experiment: imagine we taught all mathematicians a little bit about the philosophy of science, specifically the debate about explanation and asked them to be precise in their usage of the term “explanation”. Imagine we asked that subjective and clearly psychological uses of the term be avoided. Imagine they agreed to this. Would the term “explanation” still appear in the mathematical literature?

I would argue that “explanation” would disappear from mathematical literature. Here is why. We can easily rewrite the passages in the mathematical literature that contain the word “explanation” by substituting a word or phrase that does not use “accounts for” or some other term that denotes what philosophers really take explanations to be.

I will demonstrate this using one of Hafner and Mancosu’s examples. They claim that the following passage illustrates the drive for mathematical explanation:

> One of the oldest and still one of the most interesting applications of group theory arises in the study of the transformations of an ordinary differential equation. If we know that a given differential equation admits a group of transformations, then we know that the solution set must admit that same group of transformations, and we can deduce the properties of all the solutions from the properties of any one of them. A case in point is offered by the

> celebrated hypergeometric equation whose solutions include many of the most interesting special functions of mathematical physics [...] In 1836 Kummer published a set of six distinct solutions of the hypergeometric equation. [...] A glance at the list of these solutions reveals a rather complicated set of relationships which pleads for some simple explanation. We show here that the Kummer solutions are related by a finite group of transformations which serve to explain their relationships and to exemplify the use of transformation groups in the study of differential equations. (quoted in [HM05]: 219 (ellipses in original))

But far from illustrating an objective concept of explanation, the relevant part of the passage can be rewritten (I think without loss) as follows:

> ...A glance at the list of these solutions reveals a rather complicated set of relationships. We show here that, despite appearances, the Kummer solutions are related because they admit a group of transformations which allow us to deduce the properties of the solution set from the properties of one of its members.

Ironically, eliminating the word "explanation" from the above passage makes it more elucidated and shows exactly what mathematical facts are being sought. We generally cannot rewrite passages in the scientific literature without loss of meaning by eliminating the word "explanation".

So it is not clear why only mathematicians may judge methodology when it seems that others can do so just as well. And further, the evidence suggests that mathematicians do not seem particularly cautious in their nonmathematical discourse to distinguish subjective usages of "explanation" from objective ones.

A further challenge to claim (2) regarding the claim's dependence on the criteria of mathematical practice asks how many mathematicians constitute "mathematical practice"? Above, we attributed to Hafner and Mancosu the view that you can have arbitrarily small subsets of mathematicians who provide evidence on the nature of mathematical explanation. We attributed this view to them as they clearly intended their discussion of Pringsheim's mean-value approach to analysis to be an example of explanation, yet they also clearly state that "Pringsheim's approach did not have a large following. Contemporary textbooks in complex analysis did not follow his approach." ([Man01]: 112) Nor is there an indication that any debate ensued over the nature of explanation following Pringsheim's work. Yet according to Hafner and Mancosu Pringsheim's view expresses an "urge . . . to look for explanations". So Hafner and Mancosu must intend the notion of mathematical explanation to be expressible by individual mathematicians, the "trivial community", whose mathematics may never become a part of the larger mathematical practice. This is a crucial point in an argument that appears so heavily invested in the sociological concerns of the practicing mathematician in the mathematical community.

But we cannot discount the possibility of a rogue mathematician, or as Kitcher and Asprey ([AK88]: 17) might put it, a maverick, who is very much out of step with orthodox mathematics. Clearly we would not want to include his pronouncements as "mathematical practice". Practice can only count if it reflects the attitudes of some non-trivial community of mathematicians, or of some historical thread over time in a mathematical tradition.

If mathematical practice is to only reflect the larger mathematical community, this bodes poorly for Mancosu's claim. The fact is that every explanatory program from the "causal programs" in the Renaissance to Bolzano's notion of grounded proofs to Pringsheim's "organic" derivation of the main results of complex analysis have all been abandoned or ignored by the mathematical community. This suggests that we have exactly zero cases of mathematical explanation. In Bolzano's case, his view of grounded proofs may have impacted his own practice somewhat, but much of Mancosu's claim that emerges from Bolzano's work comes from Bolzano's *Anti-Euklid*, an uninfluential manuscript, first published in 1967 (in Czech) when it was primarily of interest to historians of mathematics, and thus not influential on the mathematical community. This is hardly indicative of something the mathematical community took very seriously.

Certainly only the views of some larger mathematical community ought to count as mathematical practice and not the idiosyncrasies of a few individual mathematicians.[35] If some "explanation" is not widely enough addressed by

---

[35] The *Mathematical Intelligencer*'s regular column "Mathematical communities" comes with the following instructions: "This column is a forum for discussions of mathematical communities throughout the world, and through all time. Our definition of "mathematical community" is the broadest. We include "schools" of mathematics, circles of correspondence, mathematical societies, student organizations, and informal communities of *cardinality greater than one* ..." (emphasis added)

the community of mathematicians, then we ought to take care that the "basic evidence" that Mancosu would have us admit acts as analogous to phlogiston and homunculi - that is, the obsolete or discarded "explanations" - and they are being legitimately ignored by the mathematical community. If this were not the case, and explanations were valuable to mathematicians, would they not have been more widely discussed by the mathematical community in the same way proofs are discussed and even treated as formal objects? Would not the *Abfolge* of Bolzano or the "organic merger" proof of Pringsheim been treated to a more thorough review by the community of mathematicians?

Kitcher requires that we look to the set of accepted statements when we assess mathematical practice. ([Kit83]: 164) Little of what we have seen here would count as "accepted statements". All the proved results are accepted. The insistence on, or perpetuation of, "idiosyncratic" proof-styles is not. To be accepted we would expect to see, at minimum, the proof-style cited and imitated without objection or comment.

To summarize: mathematical practice must constitute more than the pronouncement of some individual mathematician. Thus the "basic evidence" for mathematical explanation must come from the mathematical community over some span of mathematical history. But there is no community or span of history or mathematical program which Mancosu presents us with as examples of mathematical explanation. Also, even if we were to accept the "testimony" of the individual mathematician we have no way of knowing how "offhand" some philosophical remark is meant to be taken, nor do we know how philosophically informed it is. Finally, whence the prejudice against the judgement of non-mathematicians? Others may be equally capable of evaluating claims of explanatoriness just as well.

Claim 3: The third claim Mancosu makes is that there were a few important philosophers and mathematicians who can be classified as hypothetico-inductivists (h-inductivists).

> H-Inductivism is, roughly, a conception of mathematics which asserts that the acceptance of axioms for a mathematical discipline might be motivated not by criteria of evidence and certainty but rather, like hypotheses in physics, by their success in deriving and systematizing a certain number of familiar consequences. ([Man01]: 103)

Mancosu claims that Mill, Russell, and Gödel all defended versions of h-inductivism as did Lakatos and Hersh.

H-inductivism is actually only tangentially related to the argument for mathematical explanation that we are attributing to Mancosu. The connection between explanation and h-inductivism is tenuous, and Mancosu's argument does not seem to depend heavily on it. The discussion of h-inductivism appears to be included in Mancosu's discussion of explanation to lend credence to the notion that some important mathematicians have seen important similarities between mathematics and the physical sciences. Significant similarities may suggest analogies between mathematical and science. If there are analogies to be made from science to mathematics perhaps they include the fact that they both make use of explanations. Here I show that regardless of whether or not mathematicians and philosophers accept h-inductivism it does not significantly contribute to the discussion of explanation in mathematics.

H-inductivism is the same conception of mathematics as Lakatos' quasi-empiricism ([Man01]: 105). In "A renaissance of empiricism in the recent philosophy of mathematics?" Lakatos quotes a host of philosophers and mathematicians who "seem to herald a renaissance of Mill's radical assimilation of mathematics to science." ([Lak78]: 30) The list of people Lakatos quotes is impressive: Russell, Fraenkel, Carnap, Curry, Quine, Church, Gödel, Weyl, von Neumann, Mostowski, and Kalmár. Many of these mathematicians claimed that much of the impetus for the mathematics/science comparison originated in their views of the foundational "crisis" of the early twentieth century. They often alluded to the role Gödel's incompleteness results played in shaking our confidence in the certainty of mathematics.

Lakatos has been interpreted[36] as holding that the function of a proof is to explain the theorem that it is proving. Here (in [Lak78]) Lakatos explicitly makes a connection between mathematics, explanation, and his quasi-empiricism when he states "in a quasi-empirical theory the (true) basic statements are *explained* by the rest of the system." ([Lak78]: 34, emphasis in

---

[36] See e.g., [Aga81] for a note on Peggy Marchi's interpretation of Lakatos along these lines.

original.)

This seems patently false. The statement “$5 + 5 = 10$” does not explain the Peano axioms from which it can be derived. Nor does the set of all statements derivable from the Peano axioms explain them. Perhaps *our knowledge* that $5 + 5 = 10$ together with an idea about the relationship between the axioms and the statement explains why *we believe* that the Peano axioms are true, but surely they do not explain the axioms themselves. And the project of explaining mathematics is not to explain our beliefs in mathematics.

In his attempt to make mathematics a “quasi-empirical science” modeled on a Popperian view of conjectures and refutations, Lakatos frequently invokes *explanation.* Mancosu sees Lakatos raising the notion of explanation in three contexts ([Man01]: 106). They are: (1) global foundational activity, (2) formal theories as explanatory accounts of informal theories, and (3) as a class of explanatory (as opposed to convincing) proofs.

The fact that Lakatos uses “explanation” in so many different ways is quite telling. One cannot help but wonder what Lakatos would answer to the question “what is explanation?”. However, whatever Lakatos might answer, we focus on (1) as that is the only type of explanation relevant to Mancosu’s claim.

When Lakatos speaks of explanation as global foundational activity, he has in mind the following: Mathematics is quasi-empirical, and thus always conjectural. The reason quasi-empirical theories are conjectural is that the hypotheses that organize the system can only be conjectural. Therefore they will at best, provide an explanation for the set of basic facts which we hold as true and an explanation of the whole system. Lakatos finds this view in Russell's writing. Mancosu also recognizes this in Gödel's view of Russell.

Mancosu's claim about h-inductivism is weaker than Lakatos' claim about explanation and quasi-empiricism. Mancosu claims that some philosophers including Mill and Russell draw analogies between science and mathematics such that mathematicians accept axioms of mathematics because the axioms are ultimately successful in deriving and systematizing familiar consequences.

For example, Mancosu takes Russell to be defending h-inductivism in Russell's [Rus07] and other works including *Principa Mathematica*. Mancosu argues as follows: Russell claims that we obtain few epistemic advantages from taking an obvious truism like $1 + 1 = 2$ and defending it using obscure and opaque premises followed by some 250 pages of dense logic. So what Russell and Whitehead's project in *Principia Mathematica* really amounts to is the claim that we assert (believe) the premises of their system (mainly) because the consequences which follow from them are familiar and obvious. Russell claimed that this order resembles induction in the physical sciences.[37]

---

[37] Riordan ([Rio06]) claims that the different philosophical approaches to mathematics (e.g., fictionalism, formalism, logicism, Platonism, and structuralism) are all programs trying to explain mathematical truths. She argues that logicism is not up to the task.

H-inductivism is opposed to the "common notion" that the "real order" (i.e., the epistemological order) that mathematicians (or anyone else for that matter) accept mathematical truths, involves first finding a proper acceptable foundation for mathematics (like Peano's axioms) and then, based on that foundation, accepts conclusions like $1 + 1 = 2$.

We do not take issue with h-inductivism here. However, we do wish to argue that h-inductivism does not speak to the question of mathematical explanation. H-inductivism is a thesis about the fact that some mathematicians and philosophers have treated mathematics as if it were a science, in that they have accepted the truth of certain axioms because of the axioms' fecundity and ability to systematize other known mathematical results. They accept mathematical axioms in the same way that they accept certain formulas in science for their ability to systematize and derive the phenomena (e.g., we accept that Maxwell's equations are true because of their ability to systematize the laws concerning electromagnetism).

To show that h-inductivism is relevant to our question of explanation Mancosu (or Lakatos) would have to show that mathematics is like science in *some relevant way*, not merely by showing that some people have claimed that there are similarities. For example, the belief of some mathematicians that they have similar reasons for giving assent to certain mathematical statements as they would for certain scientific statements, is a minor epistemic similarity. This hardly counts as a warrant to start making other analogies from science to mathematics, especially in the absence of an explicit statement from the mathematicians themselves. We do not know how far the mathematicians would take the analogies between mathematics and science. Also, more importantly, the analogy is not about the mathematics and science, but rather about the mathematicians and their motives.

J. S. Mill was an empiricist, even about mathematics. He argued that there is no difference between the mathematical and the empirical sciences. In order for Mill to show that an analogy between science and mathematics holds, he has to show how mathematical phenomena are the same as scientific phenomena. We can then thereby account for (explain) all the phenomena using a single account of explanation. Separately, we can show that we can come to believe the premises, axioms, or equations in light of what can be derived from them (i.e., h-inductivism).

But a quasi-empirical account of mathematics is not exempt from telling us in what ways mathematics is empirical and in what ways it is not. If quasi-empiricism holds that mathematics is empirical in the same way that Mill holds, then the explanations in mathematics would run the way we described them running for Mill. And just as for Mill, it is then trivial that there are explanations in mathematics. But then one faces the standard problems of Millian mathematical empiricism.

But if quasi-empiricism has a different conception of what is empirical about mathematics (perhaps only some parts of their respective methodologies), then by what rights do they say that mathematics and science share explanatory concerns?

Mathematical empiricism has been under continuous attack at least since Frege's *Grundlagen* and has few defenders today. Quasi-empiricism would have to distinguish itself from standard mathematical empiricism to be immune from the Fregean (and subsequent) counterarguments. H-inductivism is only informative about explanation after we are given a clear analysis of how quasi quasi-empiricism really is. So whether or not we accept quasi-empiricism or h-inductivism, we are no more informed about mathematical explanation now than when we started.

Claim 4: We must approach mathematical explanation from a particularist perspective.

We find a constant tension in many areas of philosophy between approaching a subject from the "top" and approaching it from the "bottom". "Approaching from the top" refers to using a general *theory* when assessing the phenomena that the theory is about so that we take the theory as fixed and then examine various purported instances to see if the theory accommodates the instances. This is contrasted with "approaching from the bottom" which examines clear *cases* of the thing we are analyzing as fixed and then tries to extrapolate the correct theory from the individual cases.

Hume[38] for example, advocated a bottom-up approach to assessing the legitimacy of states. Hume held that there was no single procedure that determines if a government is legitimate. We should look at all the states that people have granted their allegiance to, regardless of how the states came into existence, and then we can try to extract a view of legitimate statehood. John Locke, on the other hand, argued for democracy as the theory of legitimate statehood from a general theory of rights. He held the view that states were legitimate in virtue of their being the right kind of democratic state formed via the proper social contract. Locke was thus a top-down theorist, stressing the social contract theory as granting legitimacy (downward) to states.

Roderick Chisholm[39] pointed out the top-down/bottom-up distinction in epistemology. There are top-down theories that dictate how we get knowledge, and all the knowledge we have has to be accommodated by this theory. Those theories typically take the form of "knowledge is x" where x is some set of criterion. The most famous example, from Plato's *Theatetus*, is that knowledge is justified true belief. Anything that we claim to know must be a justified true belief, and any belief that is true and justified is a bit of knowledge that we have. On the other hand, there are bottom-up accounts in epistemology. Those accounts take individual bits of knowledge and de-

---

[38] See David Hume's "Of the original contract" in *Political Philosophy: The essential texts* Steven Cahn, ed. Oxford, 2005.

[39] See "The problem of the criterion" in [Chi82].

rive an epistemology from them. G. E. Moore's proof of an external world[40] exhibits a bottom-up approach. It is bottom-up because it starts by taking particular cases of knowledge as the "basic evidence" for our theory. Moore states that he knows that he has a hand, and another hand, and thus he knows that he has two hands. If some theory of knowledge cannot accommodate Moore's knowledge of his own two hands, all the worse for that theory of knowledge.

Whether one prefers a top-down approach or a bottom-up approach depends mainly on which of two intuitions are stronger. Do we have stronger intuitions about a) particular instances of what we are looking for or, b) about a theory or the possibility of finding a theory that will subsume all the cases we take the theory to need to include, and exclude all the cases we do not want the theory to include. In our case of mathematical explanation, are we more confident about some particular explanatory pieces of mathematics or are we confident that we know what an explanation should look like? Top-down theories generally seem preferable because we get a clear procedure for determining whether each case is in or out. But the history of philosophy reveals a rather poor record of providing counter-example-proof conceptual analyses of top-down theories.

The merits and problems of top-down and bottom-up approaches in general are beside our point. Chisholm gave the approaches useful names. A

---

[40] G. E. Moore "A Defense of Common Sense" in [Moo59].

"methodist" is one who takes there to be a method for determining whether the cases fit the theory - one who takes a top-down approach. A "particularist" is one who prefers to begin with cases in assessing a given concept. So, Hume and Moore were particularists, Locke and Plato were methodists.

In the topic of mathematical explanation, Steiner is a methodist,[41] so is Kitcher, and so is anyone else who thinks explanation is x, where x is some criterion for distinguishing explanatory proofs, or explanations in mathematics. For Steiner an explanatory proof is one which meets the guidelines he spells out in the characterizing property account discussed in the last section. Regardless of any other merits the proof may have, if a proof is not accommodated by the account, all the worse for the proof, says Steiner - it is just not explanatory. Kitcher's theory is also a methodist theory. Hafner and Mancosu however advocate particularism in the study of mathematical explanation ([HM05] §4).

They offer two reasons for preferring a particularist approach, both of which are at least a little problematic. There is also a bigger problem with this approach which I spell out at the end. First, they argue that models (methods) of explanation are prejudiced against countenancing mathematics. They were designed for the sciences and thus the conceptual resources of the theories will not apply to mathematics. The models may beg the question against mathematical explanation. In their own words:

---

[41] As Chisohlm is careful to point out, and I do so again here: this terminology implies no connection with the religious group.

> As a rule contemporary accounts of explanation have been developed within the philosophy of natural science without addressing the specificity of mathematical explanations. Hence the conceptual resources of those accounts involving, e.g. the notions of causal connections or laws of nature seem inappropriate for capturing explanations in mathematics. ([HM05]: 221)

Although we have spoken about causation and laws in mathematics (§2.1 and §2.2) it certainly is reasonbable to claim that if we seriously consider only theories that have a strong connection to the natural sciences it will beg the question against mathematical explanation.

Of course a proponent of mathematical explanation can reasonably claim that the existence of mathematical explanations demonstrates the futility of a theory of explanation that relies on causation or laws of nature. That is to say that there is no reason to insist that the question-begging only goes one way. Perhaps there are explanations in mathematics and e.g., the causal approach to explanation is thus wrong, or perhaps explanations are causal and there are none in mathematics.

Regardless, it is true that many or even most models of explanation seem to have been designed with the sciences in mind. Curiously, Hafner and Mancosu do not consider that this seems to imply that philosophical intuitions are against the existence of mathematical explanations. No one has bothered to account for mathematical explanation because there is no widespread belief that mathematics provides explanations. (Perhaps to the extent that they do think there are intuitions about mathematical explanations, they take solace in their belief that claim (2) is true - that only mathematicians and not philosophers can be concerned with them.) But Hafner and Mancosu's claim is weakened by the fact that most theories of explanation were not really designed [by philosophers] to be oblivious of mathematics. Aristotle considered a mathematical example when talking about his theory of explanation, suggesting that he deliberately meant to include mathematics in his explanatory schema. Mancosu notes that, like Aristotle, Ernest Nagel's first example in his study of explanation is a mathematical one. Kitcher's work on explanation *first* considered Bolzano's view of explanatory mathematics and only later applied it to science. J. J. C. Smart's holistic account of explanation ([Sma90]) explicitly considers the place of mathematics alongside science. And more recently, like Nagel, Arnold Koslow ([Kos03]: 170) (who is admittedly sensitive to concerns of mathematical explanation) also offers a mathematical example in exploring a novel approach to explanation. So it would seem odd to exclude methodist conceptions of explanation on the grounds that the study of explanation traditionally ignored the case of mathematics. It did not.

As we said earlier on, some conceptions of explanation seem, prima facie, particularly accommodating to mathematics. As long as a conception of explanation does not rely on the physical nature of scientific objects there should be little concern. Also, as we have seen, proponents of some models of explanation that do appear to rely on physical properties have taken functionalist-type attitudes toward causation. They have treated the causal relation as abstracted from the physical or as a kind of relation (perhaps a structural one[42]) that is broad enough to accommodate mathematics.

The second argument for particularism that Hafner and Mancosu give is that methodism

> would mean forcing the evidence from mathematical practice into a predefined mould, thereby narrowing the perspective from the outset and probably leading to distortions.([HM05]: 221)

But this is just the typical particularist retort: "if you find me a framework within which I must fit in all my cases, I lose the freedom to include all the cases I may have." To which the methodist responds: "perhaps those are the cases which ought to be excluded."

More familiarly this is the paradox of analysis. How can we recognize an instance of mathematical explanation if we have no working model of it? And how can we find a model of mathematical explanation without clear instances?

One can normally "bootstrap" oneself out of the paradox by putting forth an exceptionally clear intuition - either of an instance or of a theory. G. E. Moore's "hands example" is certainly a powerful case of objects of knowledge. How does one argue with someone who says that they know that

---

[42] This was suggested by Susan Vineberg in a talk she gave at the 2007 Mid-Western Philosophy of Mathematics Workshop at Notre Dame.

they have a hand? At the very least, the philosopher who attempts to deny someone's knowledge of his own hands faces an uphill battle.

The reason for Hafner and Mancosu's particularism is now clear. In the absence of any strong intuitions about a theory of explanation, they still believe that mathematical practice provides strong evidence of particular cases of explanation. This is also what bothers Dale Jacquette. In a comment on [HM05] he states that

> It would have been useful ... for more attention to have been devoted to a nuanced examination of what is meant by and expected of explanation. Explaining a mathematical property or theorem means different things to different mathematicians. It is possible in the limiting case for any proof to be construed as explaining why the final line of an axiomatic derivation is true or why it follows from the axioms. The proof itself on such a conception is the explanation, and none better is forthcoming. If this is not mathematical explanation, or if there is some reason for not interpreting every proof as an explanation at least for the initiated, the authors do not venture to say why. ([Jac06]: 80)

To us the reason why is obvious. There are no intuitions about what explanations are, so Hafner and Mancosu make do with particularism.

But Hafner and Mancosu's case for particularism is unconvincing. First, we have already stated our objections to invoking mathematical practice (when we examined claim (2)). So were we to take our particular cases of mathematical explanation from the practice of mathematics we would have to confront the question of what constitutes mathematical practice and the fact that there is thus far no evidence that mathematical practice, on our account of practice, has ever countenanced explanations.

Secondly, it is not unreasonable for one who believes that there are mathematical explanations to expect that we will someday be able to distinguish some stylistic feature of a proof (perhaps) that enough mathematicians endorse as explanatory. And though this has not happened yet, neither do we have any good specific cases. (To the extent that Mancosu thinks there are good cases, that is, he takes the mathematicians' cases uncritically at face value, he must concede that we have have good theories, as some of the mathematicians he cites (e.g., Guldin, Bolzano, . . .) do give us theories. That is unless he thinks that mathematicians can provide only "basic evidence", i.e., particularist evidence and not methodist evidence.) There is no reason to think that an intuitive theory will not be more quickly forthcoming than a slew of intuitive mathematical cases. After all, if there are explanations, why shouldn't we be able to know what explanations are?

But along these lines, we can voice our final problem with this defense of particularism. The problem is that mathematical explanation has no "Moorean hands". There are no uncontroversial and intuitive cases that will allow us to say, (like Moore) "here is a mathematical explanation. Here is another. I have two mathematical explanations." As Paul Halmos once reportedly said: "A good stock of examples, as large as possible, is indispensable for a thorough understanding of any concept, and when I want to learn something new, I make it my first job to build one." But as both Mancosu and Sandborg have pointed out, there is no stock of examples of mathematical explanations. Neither philosophical practice nor mathematical practice has provided us with a collection of examples that 1) many mathematicians have contended was explanatory, 2) few have disagreed with, and 3) simply appear to the lights of intuition as mathematical explanations. Whether or not there are any mathematical explanations is under dispute; there is no agreement on any particular cases. There are clearly "Moorean" cases of mathematical knowledge (e.g., I know that $5 + 7 = 12$), there seem to be clear "Moorean" cases of scientific explanations[43] (e.g., some version of the theory of evolution explains the proliferation of some species) but there are no Moorean cases of mathematical explanation.

So we conclude that particularism is an unreasonable methodological approach to the study of mathematical explanation. We noted that a methodist approach might be more valuable, as we do have intuitions about what an explanation, at least outside mathematics, needs to accomplish.

Claim 5: The fifth claim we attribute to Mancosu is that mathematical explanation is not homogeneous. Mancosu believes that there are many kinds of explanations in mathematics ([HM05]: 222). When studying mathemat-

---

[43] That there are clear cases of scientific explanation has not gone unchallenged. Jaegwon Kim ([Kim94]: 53) raised this question in a related context. Kim's concern there is whether the cases give us what we want in terms of understanding. As we mentioned in Chapter 1, we don't take understanding to be a necessary condition for explanation. But we do take seriously the concern that candidates for explanation in the sciences be looked at closely.

ical explanation we should not attempt to fit all the explanations into one predetermined model. We should look to the mathematics and the mathematicians to see what it is that is called "explanation" if there appear to be many kinds of explanation.

Hafner and Mancosu explicitly say that they do not argue for this, they merely accept it.[44] (This claim is related to the previous claim (4) on particularism.)

Without clarification it is impossible to know whether Hafner and Mancosu are claiming that they mean that there are *many* models of explanation in mathematics, or there are no models at all - there are just examples of mathematics that explain or are explanatory. The former claim seems particularly unparsimonious. The latter claim implies that our previous particularism was more than a methodological claim (since there is no methodology), it was a claim about the explanations themselves, i.e., there is nothing that any two explanations necessarily have in common.

In the philosophy of science there is some general agreement that scientific explanations all explain by dint of fitting some kind of model, e.g., a syntactic or pragmatic one. Why should explanation in mathematics be so fundamentally different? After all, is mathematics so much more heterogeneous than science that it should have so much more variation on what is

---

[44] Almost as an afterthought, Steiner too asserts without elaboration that beside his own characterizing property model, there may be other methods of explanation in mathematics. ([Ste78a]: 147) David Sandborg also argues that there are a plurality of ways mathematics explains. ([San97])

explanatory? Or, is the philosophy of science similarly misguided in taking there to be a model of explanation.

If there is nothing that any two explanations necessarily have in common then in virtue of what do we call each an explanation? There must be some standard. But giving a standard commits you to some level of homogeneity in your explanations. Do we claim that there are family resemblances between our various explanations? If so that still implies some homogeneity.

What seems more likely is that to the extent that accounts of mathematical explanation must be heterogeneous, they are psychological accounts of explanation very unlike the accounts of explanation that philosophers of science have come to consider. After all, it would be odd to have such an unparsimonious and messy account of explanation in mathematics with a wide array of models or cases and yet still have so much agreement in the philosophy of science about the fact that such heterogeneous phenomena can be subsumed under one account. Again, there may not be agreement on what the correct model of explanation is, but few have argued that Kitcher's model works for some sciences, Hempel's for others, and van Fraassen's for the remainder. However, given the fact that different mathematicians have different skills, educations, cultural particularities, psychological proclivities, and mathematical *weltenschaaung*, wouldn't they be likely to have different views about what is *appealing*, or easy, or special in mathematics? This claim is simply more plausible as a psychological claim and not a philosophical claim about mathematical methodology.

If so, then Mancosu's program would be more accurately described as part of a program involving understanding mathematical psychology - a reasonable question asked often enough by practicing mathematicians. As such it would be a program of discovering which mathematicians find which kind of mathematics appealing, and why.

However, psychological claims have little to do with accounting for mathematical "phenomena" themselves. Claim (5) is simply more plausible as a psychological claim - i.e., that there are a variety of approaches to mathematical problems, and different mathematicians have different methodological and stylistic preferences.

Also, on a methodological note, Mancosu's claim of the heterogeneity of mathematical explanations cannot be the final word on this topic. We need some account of what would act as a constraint on a mathematical explanation. Is there some goal that a mathematical explanation is trying to reach? Is there some clear way it must account for phenomena? Do we know what it means to *account* for something in mathematics? What is an explanation supposed to do? So at the very least we would expect an argument for the homogeneity of mathematical explanations, that takes into account some of the goals of explanation.

Finally, it would certainly be odd if it turned out that there is only one account of scientific explanation and a plethora of accounts of mathematical explanation. The philosophical literature is full of *competing* models of scientific explanation, and there are few pleas for the acceptance of a variety of models. The consensus is (weakly) that if there is a model of scientific explanation the we must find the correct one. This claim can be narrowed: even if there is little consensus that there ought to be one model of explanation across science (construed broadly), we can plausibly argue that *within some particular field* there ought to be some uniformity in the explanations. So say we accept one type of explanation for geology, another for physical anthropology, and a third for mathematics, we can still claim that mathematical explanations, should they exist, should be heterogeneous in structure. It would require an argument to justify such a significant methodological disparity between mathematics and science and a stronger one to argue for such disparity within mathematics. And if this methodological disparity is justified, we may then question what they each mean when they speak of the general goal of explanation.

Because of the above considerations, I think that the claim that mathematical explanations are heterogeneous, is false. If there are mathematical explanations - which I doubt - then it is more likely that there is only one model which will subsume them all.

Given the above problems with the five claims I attribute to Mancosu, I contend that his case for explanations in mathematics remains wanting.

Chapter 4

Against mathematical explanation

> Mathematics itself is never the explanation
> of anything – it is only the means by which
> we use one set of facts to explain another,
> and the language in which we express
> our explanations.[1]

This chapter concludes our argument against mathematical explanation. Up until now we have been challenging the existing accounts of mathematical explanation. Here, we begin by reviewing the scant literature that also challenges the existence of mathematical explanations; we then offer a not unrelated set of positive arguments against mathematical explanation which in part consolidates and expands on arguments offered earlier which are scattered throughout Chapters 2 and 3. We argue that there are good reasons to expect that no theory of explanation will ever work for mathematics for reasons that are inherent in the concept of explanation and in the nature of

---

[1] Steven Weinberg, *Dreams of a Final Theory*: 56.

mathematics.

4.1 From mathematical, philosophical, and scientific practice

Though few thinkers have (at least in print) opposed the view that there are explanations in mathematics,[2] we begin by looking at those philosophers, scientists, and mathematicians who did. Perhaps the earliest opposition to mathematical explanation can be found referred to in Aristotle's *Metaphysics*. Aristotle reports that Aristippus and other sophists who, out of their contempt for mathematics, seized on the fact that there are no *reasons* given for mathematical truths. For example, the fact that the interior angles of a triangle equal two right angles, is brute.; there is no reason for it. This made them regard mathematics inferior even to carpentry and cobbling ([Ari52]: 996a32). Rivaltus argued that mathematics merely appeals to any argument that gets results. Several geometers in the seventeenth century seem to have agreed (see [Man00]: 111). There was thus no sense for them in which some proofs are to be regarded as explanatory or in any way special.

In the modern period, David Hume alludes to the fact that mathematical necessity precludes the need for any mathematical explanations. In his *Dialogues Concerning Natural Religion* ([Hum80]) the character Philo (who is usually taken to speak for Hume) offers a final counterargument to an *a*

---

[2] Though in conversation many philosophers have said to me that it is "obvious" that there are no explanations in mathematics.

priori version of the cosmological argument for the existence of God. Philo is arguing that certain facts of the world may need no explanation, contrary to the Principle of Sufficient Reason. Those facts that need no explanation are necessary and indeed are brute, like the facts of mathematics. Hume writes that:

> It is observed by arithmeticians, that the products of 9, compose always either 9, or some lesser product of 9, if you add together all the characters of which any of the former products is composed. Thus, of 18, 27, 36, which are products of 9, you make 9 by adding 1 to 8, 2 to 7, 3 to 6. Thus, 369 is a product also of 9; and if you add 3, 6, and 9, you make 18, a lesser product of 9. To a superficial observer, so wonderful a regularity may be admired as the effect either of chance or design: but a skillful algebraist immediately concludes it to be the work of necessity, and demonstrates, that it must forever result from the nature of these numbers. ([Hum80] §IX)

Consistent with Hume's scepticism about explanations in general, here Hume is arguing that nothing more needs to be said about mathematical truths, presumably including explanations. There is no reason they need accounting for, and it hardly makes sense to seek out explanations. The facts of mathematics are there of necessity; they are brute and unexplained.

In a 1963 work Carl Hempel, who in an important sense was an intellectual heir of Hume, claims[3] that the covering law model (which together with his I-S model were the original canonical models) of explanation, does not

---

[3] See above Chapter 2.2, and [Hem63]: 126.

countenance mathematical cases. Rudolf Carnap in 1966 essentially denied that there are mathematical explanations as well, and claims that the function of laws in science (in the context of D-N and I-S models of explanation in science) is to take on the dual role of explanation and prediction. He explicitly tells us that the laws he is talking about are those that tell us about the world, i.e., scientific laws and he excludes the "laws" of mathematics, which he claims "merely state relations that hold between certain concepts, ... only because those concepts are defined in certain ways" ([Car66]: 8, 9).[4]

A few years later, in a 1981 interview, the scientist Richard Feynman, who gave us the Feynman diagram - a method of doing relatively difficult calculations in quantum field theory using relatively intuitive diagrams - tells the following story:

> I hear [my cousin] talking about x. He says, "What do *you* know – $2x + 7 = 15$," he says "and you're trying to find out what x is. I says (sic) "you mean 4." He says "Yeah, but you did it with arithmetic, you have to do it by algebra," and that is why my cousin was never able to do algebra, because he didn't understand how he was supposed to do it. There was no way. ... the whole idea was to find out what x was and it didn't make any difference how you did it - there is no such thing as ... you do it by arithmetic, you do it by algebra - that was a false thing that they had invented in school ... ([Fey99]: 6)

There is no reason to believe that Feynman would have a different attitude

---

[4] Other works on philosophy of science of that period mention mathematical explanation as well e.g., [LJ70].

than the one he expressed about any other area of mathematics. As Rivaltus held, one method of solving mathematical problems is as good as another, and proof style is presumably a matter of convenience. There is no better way of finding x, which means that there is no method that "explains" how to get it either, otherwise that way would be better. Feynman's sentiment goes completely counter to the attitude of some proponents of mathematical explanation who see proofs as more and less explanatory, or some mathematics as providing explanations and others failing to.

Despite her keen interest in the unification of diverse branches of mathematics, Emily Grosholz ([Gro00]: 81) remarks explicitly that "mathematical truths are constructed or demonstrated, not explained." In a similar spirit we have an exchange by mathematicians prompted by a review of James Gleick's *Chaos: Making a New Science* ([Gle88]) in the *Mathematical Intelligencer*. In reviewing the book John Franks stressed the cast of characters, the perception of chaos theory by mathematicians, as well as the book's apparent lack of interest in mathematical rigor. Morris W. Hirsch, in a comment on the review, laments that there was not enough space in Gleick's book given to the rigorous mathematical proofs that accompanied the emergence of chaos mathematics. He writes that he wishes there was more proof and less publicity to the "nonexistent science of chaos".

James Gleick responded that mathematics was akin to science in the sociological sense as there are problems and personalities in both, and that chaos mathematics emerged less from the proofs as from the whole mathematical enterprise surrounding it. He writes:

> There are times when mathematical proof (essential though it is!) comes, historically, as an afterthought. …Lanford's proof of Feigenbaum [sic] was ingenious and admirable, but it did little, really, to validate Feigenbaum's breakthrough - experiments accomplished that ([Hir89]: 9).

The mathematician David Gale responded with a letter that displays an acute sensitivity to the issues of explanation in the philosophy of science and mathematics. The letter is worth quoting at length:

> …The point is that Lanford and other mathematicians were not trying to *validate* Feigenbaum's results any more than, say, Newton was trying to validate the discoveries of Kepler on the planetary orbits. In both cases the validity of the results was never in question. What was missing was the *explanation*. Why were the orbits ellipses? Why did they satisfy these particular algebraic relations? In his celebrated "afterthought" Newton provided the explanation while, incidentally, adhering meticulously to the "theorem-proof methodology" which, according to Gleick and Keith Devlin, "most scientists find peculiar."
>
> I believe many mathematicians share Gleick's excitement over the fascinating experimental discoveries of Feigenbaum as well as the beautiful self-symmetries of the Julia sets, the omnipresence of the Mandelbrot set, and so on. Indeed, …providing a whole new range of mathematical phenomena, phenomena which, to a mathematician fairly cry out for explanation. Thanks to the work of Lanford and others one can now account for a great many of these experimental results …

> But . . . there's a world of difference between validating and explaining. Physics wouldn't be much of a science if physicists simply measured and recorded the light spectra of the various elements and let it go at that. The main goal of all science is first to observe then to explain phenomena. In mathematics the explanation is the proof. It's as simple as that, and I doubt that any scientist who understands this can find our methodology peculiar. . . . there is nothing peculiar about using the theorem-proof methodology, because in fact it's the only methodology there is. ([Gal90]) (Emphasis in original.)

This is a clear statement from a prominent mathematician articulating an understanding of what scientists and philosophers are looking for in their respective disciplines, and argues that there is no mathematical counterpart to scientific explanation. Undoubtedly this reflects a view of mathematical practice, and it is perhaps the clearest statement about mathematical explanation on the topic by a mathematician.

The philosophers Jeremy Avigad [Avi06] rejects the use of "explanation" in descriptions of mathematical practice. He claims that (1) 'explanation' "is not so very often used in mathematical practice" and (2) "it is certainly not the only term which is used to voice positive judgments about proofs." Further, he "would prefer to remain agnostic as to whether there is a single overarching concept that accounts for all such positive judgments, or rather a constellation of related notions; and also as to whether the particular virtues here are best labeled 'explanatory.' "

Avigad realizes that it is becoming common practice in the philosophical literature to label many kinds of positive judgments about a proof that go beyond establishing their theorems as "explanatory". Avigad rejects this practice, and so do we. One of Avigad's main goals in [Avi06] is to expose those virtues which are valued in mathematical proofs,[5] e.g., generalizability, brevity, "pointing" to the way other problems can be solved, establishing stronger statements than the theorem they purport to prove, introducing definitions and methods useful elsewhere, etc. He does say that some of these can be classified as explanatory, though he offers no further analysis.

Another philosopher, Richard L. Epstein ([Eps02]), while not explicitly opposing the existence of mathematical explanation, does seem to express a degree of hesitancy about the kind of mathematical explanations we are looking for. He gives the following dialogue as an example of an explanation:

> Student: Why is the square of the length of the hypotenuse in a right triangle equal to the sum of the squares of the lengths of the two sides of the triangle?
>
> Professor: Because (gives a proof of the Pytahgorean theorem).
>
> *Analysis*: The student seems to be asking for an explanation, the reason why the mathematical claim is true. Certainly the student is asking for an inference of some sort. But that inference isn't an explanation, it is an argument or proof. It is meant to convince the student that the conclusion is true; the student doesn't already believe the conclusion. (269)

---

[5] This has numerous affinities with Sandborg's [San97], who would, however, certainly disagree on the specific case of explanation.

But for Epstein, in a case like this, "explanation" means something like a surveyable proof" (138). A surveyable proof tells us why the conclusion is true. Surveyability is a condition for an inference (or set of inferences) to be a good explanation. So "*a proof in mathematics is a kind of explanation, and a better proof is a better explanation*" (*emphasis* in original). So a more surveyable proof (which comprises a vast majority of the proofs already proved) would presumably produce a better explanation. "But," Epstein concedes, "just as for explanations, there is no agreed-on standard of what counts as a proof or a good proof, though there is often strong agreement on particular cases." And certainly we can add that there is no agreed upon standards of what (if anything) would count as a *better* explanation, or for that matter a *better* proof. Nor are there agreed-upon instances of particular cases of explanation in mathematics (as we saw in §3.2).

Epstein does seem ready to concede that a proof may actually be (provide) an explanation, though his final remark about particular cases suggests that he thinks that there are some proofs that explain and some that do not. In the end however, he qualifies all remarks about mathematical explanation by saying that his discussion of explanation and proofs is ultimately subject to an analysis of "explanation".

To the best of my knowledge, there is no other significant writing opposing mathematical explanation. Numerous philosophers have scrutinized scientific explanation at least since Aristotle's time. Mathematical explanation however, has garnered comparably little attention. David Sandborg ([San97]: 8) claims that the fact itself that the topic was ignored calls for some explanation, though given our conclusions in this work, the reason for the lack of attention can be due to the fact that in the sense that we use it, "explanation" has always been designated to range over scientific phenomena and is ruled out in the mathematical cases.

Going back to Aristotle's four causes in his *Physics* and his rejection of the Forms in the *Metaphysics*, *prima facie* explanatory considerations involved a four-fold account that addressed matter, form, process, and purpose,[6] only one of which looks like it can possibly be taken seriously to apply to mathematics.[7] So tradition and the nature of the traditional explanatory program as generally viewed by philosophers, barred discussion of mathematical explanation almost by definition.

Another reason for the lack of attention to mathematical explanation is the apparent attitude of the mathematics community. There is a general background assumption that mathematicians are supposed to justify their theorems by proving them. Mathematics texts at all levels teach their readers how to do proofs. At the very least they offer clear examples of proofs and

---

[6] Mathematics and Aristotle's Prime Mover are exceptions. The Prime Mover is purely final cause as it is immaterial, formless, and eternal. Mathematics is easily construed as being purely formal cause. But this kind of exception does not detract from the *prima facie* case that mathematics is not causal. Anyone who claims otherwise must justify this counterintuitive position. Thus it is still to have been expected that there is little discussion of explanation in mathematics.

[7] See our discussion of this in §3.2.

point out that that is what they are doing. Mathematics texts do nothing comparable for explanations. Proofs are valid or invalid, sound or unsound, but mathematics texts rarely treat proofs as explanatory or unexplanatory, there are certainly few texts (if any) which explicitly teach how to explain mathematics or offer mathematical explanations.

There are many considerations that have always factored in to mathematicians' judgement of proofs, such as their aesthetic properties[8] or length[9]. But for the most part none of these other considerations have had their relevance to an evaluation of a proof clearly or carefully articulated by mathematicians or philosophers. It seems likely that while these considerations are of concern to some mathematicians, they do not further any mathematical goals. Great pains are taken by mathematicians to further the goals of mathematics. Among their goals are proof, and perhaps precision, and the development of new methods. Explaining is not a stated goal.

Now that we have seen that there are some mathematicians and philosophers who have offered assertions or arguments against mathematical explanation, we offer a number of our own.

---

[8] There are surprisingly very few studies of the aesthetic properties of mathematics, given that mathematicians put so much stock in the notion. See however, Reviel Netz's [Net05] for a recent discussion.

[9] Until the twentieth century it was not at all clear how to even phrase questions about the length of proofs, and there are still open questions. But we need not delve into those issues here. However the following anecdote related to me by Nickolas Pappas may prove interesting. W. V. Quine once insisted to Warren Goldfarb, despite Goldfarb's strong resistance, that some proof was superior to some other proof of the same theorem, a priori, because it had one line less. This was despite Goldfarb's argument that his longer proof had other virtues.

We take various approaches to the counterarguments. Largely, we take the criteria which have been generally accepted as either necessary, or at least important, features of explanations and show how each of these criteria cannot accommodate the mathematical cases. If the criteria associated with explanation cannot be applied to mathematics then any future theory of mathematical explanation that is in line with these criteria for scientific explanation will be precluded. There is no accepted list of necessary or sufficient conditions for scientific explanation. Assembling such a list is still a serious goal of the philosophy of science. Nonetheless, we can identify a number of restrictions on explanation whose absence would be glaring in a fully articulated theory of explanation. We inquire into what mathematical explanation could have in common with the traditional explanatory project, and find significant disanalogies.

Some of the arguments that follow will be familiar from earlier sections. Here we consolidate them and take them out of the context of responding to some particular philosopher or model of explanation, and use the argument straight-forwardly against any model of mathematical explanation.[10]

4.2 The absence of explanation

Our first argument is an argument from absence. As we just saw, Avigad points out that there are relatively few references to explanation in main-

---

[10] Thanks to Michael Levin for suggestions which formed the core of two of the following arguments.

stream mathematical literature. Of those references to mathematical explanation that do appear in the literature of pure mathematics,[11] most are pedagogical-type explanations that offer clarification or something else that does not resemble the kind of mathematical explanation philosophers are looking for. We have seen a number of examples of this in the preceding chapters. Though philosophers have lately been pointing out exceptions (see e.g. in [Man08b]), the examples tend to be rather exotic or obscure. There has been no systematic analysis of standard and well-discussed mathematics texts illustrating any pattern of mathematical explanations.

Moreover, the desire to pursue explanations of mathematical facts did not motivate new mathematics in the way the need to explain the phenomena of the natural world motivated science. The desire to prove theorems (or model the physical world, develop new mathematical methods, unify mathematical theories, ...) is what generally motivated mathematicians. We addressed many of the claims to the contrary above.[12]

Further, in stark contrast with the vocabulary of science, mathematicians' professional vocabulary rarely include descriptions of explanatory projects. (If universities insisted on moratoria on explanations, mathematicians would not see their projects curtailed, scientists would.) Descriptions of mathematical projects can dispense with explanation talk with no loss of clarity or

---

[11] See the Appendix to [San97] for one method of searching for "explanation" in mathematical literature.

[12] See above Chapter 3.2.

meaning, unless the program is expository or pedagogical.

Arguments from absence however do not imply the absence of arguments. So the fact that we do not hear about explanations from many mathematicians or philosophers is not a decisive argument against their existence. But it does speak to the nature of mathematical practice. We note that if mathematicians failed to notice such a fundamental part of their discipline - and proponents of mathematical explanations construe explanations as an important part of mathematics - then we have reason to be concerned. A large number of mathematicians thoughtfully contribute to the philosophical and methodological development of their field as they have always done from Euclid and Proclus to Russell, Frege, Gödel, and von Neumann, yet rarely do we encounter any sustained discussion about explanations and their nature. One would expect mathematicians with a philosophical bent to notice this phenomenon in mathematics and treat it with the seriousness that such a methodologically valuable feature deserves. Since we do not find this we have reason to believe that we should not expect to find, on analysis, explanation as part of the methodology of mathematics. If mathematics did countenance explanations then we would expect to find more discussion by mathematicians. But we find no such discussion, so at least our scepticism about the existence of mathematical explanation seems *prima facie* reasonable.

4.3 The Symmetry Thesis

Our next argument against mathematical explanation takes a well-discussed property of explanation - the Symmetry Thesis - as a starting point. In the context of his major exposition of the D-N model of scientific explanation, Hempel ([Hem65a]: 367) argued in favor of the thesis of structural identity, or the Symmetry Thesis.[13] The thesis is that (i) every adequate explanation is potentially a prediction and (ii) every adequate prediction is potentially an explanation. So there is a symmetry between prediction and explanation such that they function as two sides of the same proverbial coin. Where there are predictions there will be explanations, and where there are explanations there could have been predictions. The rationale that Hempel provides for the Symmetry Thesis is that is that any rationally acceptable answer to the question "why is it the case that X?" must offer information which shows that X was to be expected. Thus, the explanatory information should provide adequate grounds for believing X, otherwise we would have no ground to say that we have explained X. So an explanatory account satisfying this condition is a potential prediction that could have predicted X had we known the information in the explanans early enough (367-368). Briefly, predictions are made prior to an experiment or observation. A prediction is justified with certain facts. It is these same facts that are invoked after the experiment or

---

[13] The Symmetry Thesis appears earlier in Mill's [Mil41]: Bk III, ch XII §6, p 310, and also in Popper's [Pop59].

observation as an explanation of the phenomena. For example, the return of Haley's comet in 1758 was predicted using Newtonian physics. Once the comet was observed the same Newtonian calculations that we used to predict the return, could be used to explain why it returned when it did. The information that is necessary and sufficient for an explanation should also be necessary and sufficient for prediction.

Now, since Hempel's time[14], (ii) is usually denied though philosophers still tend to accept (i).[15] (ii) faces serious counter-examples which generally involve explanans that are undetectable for practical reasons prior to the explanandum.[16] So though it is not necessarily the case that the existence of predictions implies the existence of explanations, it is still taken to be the case that if there are explanations, then there are predictions. Thus if there are mathematical explanations in the same way that there are scientific explanations, then there ought to be mathematical predictions as well. But, as we argued in §2.3.1, there is no such thing as a mathematical prediction. Mathematicians certainly make guesses, but not predictions. Any information that would justify a prediction would be sufficient as a proof. Just as in science, theories make predictions, scientists make guesses or hypotheses. So

---

[14] The Symmetry Thesis was widely discussed in the 1960s, and continues to be of interest. See [Sal89].

[15] See [Sal89]: 49.

[16] Two famous early counterexamples are 1) a barometer reading being a predictor of a storm but certainly not an explanation of that storm, and the pre-Newtonian ability to predict features of the tides given the position and phase of the moon without being able to explain them for lack of an adequate theory.

given the Symmetry Thesis and the absence of prediction in mathematics, we claim that there is no explanation either.

4.4 Argument from nonubiquity

Another argument against mathematical explanation is based on an assumption made in scientific practice and not mathematics. Scientists assume that everything in nature is, in principle, explainable. But the philosophers of mathematics who have addressed mathematical explanation make no comparable claim for all of mathematics. Moreover, on some views of mathematical explanation this claim is impossible for mathematics, though it is quite commonplace in science. This disanalogy is yet another way in which mathematics cannot share in the traditional project of explanation.

In Dale Jacquette's review of Mancosu's [HM05] he asks:

> should we . . . require that deductive inferences must carry at least some burden of explanatory content in order to qualify as proofs? None of the authors in this volume are willing to go that far, and the specter of psychologism looms above recommendations in the philosophy of mathematics concerning the definitions of metamathematical concepts and categories. ([Jac06]: 80)

In addition to pointing out that explanations in mathematics may very well be all psychological, Jacquette notices that even the most ardent exponents of explanation in mathematics see large portions of mathematics unexplained or non-explanatory.[17] None of the authors he addresses goes so far as suggesting that all proofs are required to have some explanatory content to count as valid proofs. So some proofs can be completely non-explanatory and still perfectly acceptable as proofs. The reason for this seems to be that there are at least some clear cases of proofs that everyone can agree on that have no explanatory content whatsoever, and there seem to be many mathematical facts that clearly cannot act as explanations that are accepted by the mathematical community without comment.

On the other hand, each scientific fact about the world is presumably subsumed under some law or is somehow accounted for within some explanatory framework. Certainly all the ordinary facts about the universe are taken to have explanations. And when science discovered facts that could not be subsumed under ordinary accounts of explanation, philosophers invented new models of explanation that subsumed the new facts with the old. Statistical laws and their corresponding models of explanation come readily to mind. It is never claimed for mathematics that everything falls under some explanation, nor is it clear what it would mean to explain mathematical laws in the same way that we may want to explain scientific laws. So mathematics exhibits a clear disanalogy with science with respect to explanation. Within the domain of some science, in whatever sense we mean "explanation" science proceeds under the assumption that everything falls under one. If we

---

[17] Note also Salmon's [Sal89]: 104 where he remarks negatively about all deductive arguments being considered explanatory.

went to a doctor with some particular pain, we would expect that the doctor would tell us that the pain could somehow be accounted for. The doctor may not know what accounts for your pain, he may not be able to offer an explanation, but we would assume he was a quack if he claimed that there simply was no account.[18]

The converse holds as well. In practice, unlike in science, the search for explanations certainly has not motivated all mathematical research. Even proponents of mathematical explanation must seek out rather exotic examples to justify their claims that *some* mathematicians are interested in explanations.[19] In science, on the other hand, the search for explanations has motivated quite a lot. Everything beyond specimen collecting in science is driven by a need for explanations. Even specimen collecting is often in service of explaining some larger phenomenon or regularity. The things we take to be the major accomplishments of science were explanations. Consider especially those frameworks provided by Galileo, Kepler, Newton, Darwin, Dalton, Faraday, and Einstein. All of them provided scientific explanations of various phenomena. In mathematics, on the other hand, the major significant achievements are methods and proofs: the introduction of zero, the calculus, negative and imaginary numbers, Hilbert's tenth problem, Fermat's last theorem, the cardinality of the continuum, or the Pythagorean Theorem. Sometimes mathematicians sought proofs for things that most mathemati-

---

[18] This example is modified from one given by Steven Cahn.

[19] See especially [Ste78a], [Man00], [HM05], and [San97].

cians already believed were true (e.g., the Four Color Theorem), sometimes they discovered proofs for things that were surprising (e.g., Gödelian incompleteness).

So there is no expectation that explanations are ubiquitous in mathematics or that mathematical discoveries are typically motivated by a need for explanation. We do not assume that all mathematics falls under some explanatory scheme. So our outlook regarding mathematical explanation is different from the one we have toward scientific explanation for yet another reason.

One may respond to the above objection by noting that the objection is still consistent with the idea that in science, explanations play a large role, while in mathematics explanations play a small, but non-zero role. And this would be accurate. But the disparity between the role of explanation in science and the role it would play in mathematics would make any analogy between the two - perhaps not meaningless - but certainly stretched. A proposed role for explanation that was so small would also hardly satisfy proponents of explanation who say that it is such an important part of mathematics.

A related disanalogy between science and mathematics consists in the expectation that in principle there is an underlying order to the universe to be explained. Scientists work under the assumption that there is a fundamental theory that will explain everything and it is waiting to be discovered, and the pursuit of it is a fundamental active research question. Especially given the seeming failure of foundational programs and contemporary scepticism about the need for axioms, a comparable assumption is rarely made in mathematics. To the extent that there is interest in in foundations for mathematics, we now speak of set theory, category theory, and higher-order algebra for example, as all serving as foundations for mathematics. The idea of *a* foundation for mathematics has less currency. Therefore it is both unsurprising that there is a dearth of discussion of mathematical explanation and there is reason to claim that explanations do not pertain to mathematics.

4.5 Reducing surprise

A view of explanation attributable to Ernst Mach is that one of the primary functions of an explanation is to reduce surprise ([Mac93]: 7). This view holds that the goal of an explanation is to show how the explanandum was to be expected. In scientific matters an explanation manifests itself as showing, for example, that given an understanding of some basic principles about the universe, some particular facts can be expected to be the way they are. In other words, given some information, i.e., the explanans, the rest of the information, i.e., the explanandum, is unsurprising. We see this manifested in most of the major candidates for a theory of explanation. The reduction of surprise itself is not a candidate for a theory of explanation, but rather it is taken as a criterion of explanation, thus explanations should reduce surprise. Here we argue that one reason why there cannot be explanations in mathematics is because there is nothing a mathematician can offer that will make a mathematical result objectively *less surprising* [in the way we outline below] in a way that is relevant to our program of mathematical explanation.

We find the concept of the reduction of surprise in mathematics in Bolzano's ([Bol10] §36) distinction between *problems* and their *solutions*[20] on the one hand, and theorems on the other. He claims that the distinction is "not *objectively scientific*" and only concerns the mere *manner of presentation.* He then goes on to say that "For those propositions whose statement *does not cause surprise*, and whose truth one can grasp immediately, even though dimly, from what went before, the form of the *theorem* is suitable. *On the other hand, for those which would never have entered one's mind*, the form of a *problem* is more suitable." The straight-forward psychological nature of Bolzano's notion of the reduction of surprise is even clearer when we look at his example. The proposition: "factors taken in a different order give the same product" is a proposition that as a "theorem", is not very surprising. Whereas the proposition about "*how the highest common factor of two numbers is to be found*" is always more appropriately presented as a "problem". Presumably this is because the latter is less obvious.

So for Bolzano the distinction between that which causes surprise and that which does not is merely a question about what is more and less obvious – a

---

[20] E.g., of the kind found in Euclid's *Elements* Book II Prop. 14, or Book III prop. 1.

distinction that Bolzano does not take to be an objective one. A fortiori, if the distinction is not objective, neither is the transition from something that causes more surprise to something that causes less.

David Sandborg's discussion [San97] of Polya's [Pol49] attempts to show that there are different kinds of proofs: the kind which exhibits "motivation" for the solution and the kind that does not. The kind that exhibits "motivation" shows how the result follows from the nature of the proof. The kind that does not exhibit motivation has the result come out of nowhere, so to speak; it takes the reader of the proof by such surprise that he would not have known how it could have been solved.

George Boas, in a discussion of Emile Meyerson's understanding of scientific explanation says that "mathematical equations are no pure tautologies; . . . When one says, for instance, $(a + b)(a - b) = a^2 - b^2$, the equality is a discovery and reveals something hitherto unknown to the mathematician. . . . the equation asserts an identity where there is apparent difference; the beginner in mathematics is *surprised that the identity should exist.*" ([Boa30]: 98, *emphasis added.*) But clearly mathematical explanations ought not to be said to exist on account of having to reduce *this* kind of surprise in mathematics. This merely points out something some people didn't know, or perhaps didn't expect. It is hardly an objective feature of mathematics.

I argue that though the distinction between surprising and not surprising can sometimes be pedagogically or psychologically useful, it will not yield a principled way to distinguish mathematical theorems which have been explained from those which have not. This is because it does not make sense to speak of reducing surprise in mathematical contexts. Why not? Because (as per the general assumption) in mathematics everything true is necessarily so, and from a mathematical perspective (as opposed to the subjective perspectives of the practitioners) the necessary features of mathematics should not be surprising. Things are often taken to be surprising to the extent that they could have been otherwise. It should not be possible to clearly and distinctly conceive of some answer other than the correct one for some given mathematical problem, so the correct solution should not be surprising.

Below, we outline the specific sense in which we use "surprise". There is a psychological sense of surprise by which we may say that many mathematicians can be made less surprised by a particular result. We have no qualms with that . But surprise, to the extent that it will be useful in an analysis of explanation, must be somehow made in to an objective concept. It is in that objective concept which we describe below which we show how it cannot be reduced in mathematics. One may not know an answer to a particular problem, or one may be mistaken about an answer, or one may not expect a specific answer, but surprise is not reduced when you are informed of your mistakes or told the answer. Surprise is reduced when you can consider the answers that can be true given all the possible worlds, and then discover which world is the actual one. One can be baffled by the array of possible worlds and then be surprised that some particular world is the actual world. However, the fewer possible worlds that exist the less initial bafflement there is. The less initial bafflement, the less surprise that there is when you discover which is the actual world. The less surprise that there is in discovering the actual world, the less room there is for an explanation that reduces surprise. In mathematics, we shall see that there is no room for reducing surprise due to a lack of possible worlds in which the answer is different.

In order to use "surprise" without any psychological connotations or possible world apparatus at all, we employ the information-theoretic definition of "surprise" by defining the "surprise factor" involved in an explanation, and consider an argument in terms of the contrast class involved in explanations. Contrastive explanations, or explanations that use contrast classes, explain why something is true relative to alternative possibilities. We generally want to explain the occurrence of P relative to a class of contrasting possibilities. Thus *one way* to reduce surprise would be by shrinking the contrast class. Decreasing the size of the contrast class makes it less surprising that the answer is some given member.

Another way of saying this is that an explanation shows why P, given some probability distribution on a set of possibilities. Surprise is reduced when we show that P is more probable than the other possibilities.

So, if for some situation there are only two (more or less equally likely) possibilities, A and B, and you discover that B obtains, it is mildly surprising because it had to be one of the two. If you discover that C obtains, where C is one in a series of 100 possibilities, then it is more surprising because you have discovered 99 options that do not obtain as opposed to just one. That is: it could have been any one of those 100 and thus much more surprise was reduced. You are surprised it was not A. You are surprised it was not B, you are surprised it was not D ... Imagine being surprised that *your* name was drawn from a hat. If there were only two names in the hat, you would not be that surprised. If there were 100 names, you would be pretty surprised that it was not A's name, B's name, C's name..., but rather your name.

Information theory provides us with a definition of "surprise" that we can use to measure this. Surprise is measured (in bits) as the information content associated with the outcome of a random variable. By definition the "surprisal" contained in a probabilistic event depends only on the probability of that event. So, the lower the probability of some event occurring, the higher the surprise when it occurs. A highly probable event is hardly surprising. The surprisal (in bits) of an event with $Pr(x)$ is $-\log_2 \Pr(x)$. In other words, the amount of information gained in discovering x and eliminating alternatives to x is $-\log_2 \Pr(x)$. So, for example, if for some situation we have a number n of possible outcomes, the probability of one of them being the actual outcome is $\frac{1}{n}$. So, the surprisal, or "surprise factor" is $-\log_2 n$. When n is a real number of alternatives, say 8, then the probability of each alternative is $\frac{1}{8}$. Thus $-\log_2 \frac{1}{8}$ is the surprise factor. $-\log_2 \frac{1}{8} = 3$. Thus can we quantify the amount of surprise as 3. So when the possibilities are $A_1, A_2 \ldots A_8$, discovering that A_7 is the actual one choice is surprising to degree 3.

However, in the case of mathematics, where all of our possibilities are necessary truths, the probability that an outcome is true is 1/1 or 1. $-\log_{2^1} = 0$. There are no alternatives to consider. At least there is no finite set of alternatives that we can consider as the contrast class for the actual answer. So we should not be "surprised" at all to discover that some theorem is true. Surprise comes as we eliminate wrong, but possible answers. In the mathematical cases there are certainly wrong answers, but not answers that are possible alternatives to the actual answer. (We note that since the probabilities are defined information theoretically we do not face the problems associated with a subjective account of probability.)

We do not make the case here that the reduction of surprise ought to be central to scientific explanation, though there have been suggestions that a related notion - the elimination of possibilities - ought to be (see [Kos03]). Even in the case of scientific explanation there are a few potential difficulties for this view that come readily to mind.

First, measuring the surprise in the scientific cases where we have numerous scientific alternatives is not easy. In science not every alternative is equiprobable. However, this fact changes nothing significantly except that the probability distribution we described will not be a fraction, but rather a (complicated) function.

Also, in most scientific cases n will likely be infinite. For example, when we ask why the planets follow an inverse square law, we are contrasting the inverse square law with an inverse cube law, an inverse π law, . . . , and an infinite number of other possibilities.

We can nonetheless treat this method of looking at surprise as relevant to scientific explanation. We would have to offer though that (1) our method is just a useful idealization, or (2) in scientific cases there is an infinite amount of surprise being eliminated, or (3) as humans, we partition probabilities in various humanly conceivable ways, such that there is a humanly conceivable relation between each possibility and the whole of the probability space. So this way of looking at explanation likely has merit in the sciences.

However, given the fact that there is no way to design a model of explanation for mathematics that will reduce the only kind of surprise we can make objective and that will hold for all perspectives, we have yet another reason to be skeptical about the prospects for explanation for mathematics.

4.6 Argument from unemployment

Our next argument against mathematical explanation is that explanations in mathematics add nothing relevant to what we already know about mathematics and are thus methodologically superfluous. Whereas in cases of scientific explanation we justify the existence of explanations on the grounds that they contribute something new to what we already know, we do not have the same justification in mathematics. Explanations in mathematics would be "unemployed". Without any use, considerations of parsimony dictate that we do not countenance them.

In §1.2 we briefly mentioned the views of some proto-positivist thinkers who did not even countenance scientific explanations. They believed that explanations were part of the bloated metaphysics of an antiquated philosophy of science and should be purged together with bad explanatory programs such as those which countenanced such entities as phlogiston and entelechies. Their arguments varied but essentially they argued that there is no need to postulate explanations in order to do science, and as such explanations are superfluous. These science-minded philosophers were concerned that the answers to questions about explanations would lead us astray of proper scientific methodology. They were concerned to dispense with the question "what can explanations be used for?" and its answers. They believed that to the extent that explanations were used to inform us of the ultimate *purpose* of the universe or the *essence* of things in it, then the explanations are mistaken, or at the very least irrelevant to science. Many philosophers thus altogether abandoned the idea that explanations played any role in the scientific enterprise ([Sal89]: 126-127).

Though the positivists agreed that explanations had the potential to add an unnecessary metaphysical burden on science, but scientists nonetheless insisted that explanations were an integral part of the scientific enterprise, and scientific practice it seems, continued on as if explanation-talk was indispensable. To accommodate both their own ontological scruples and the scientists' description of their enterprise, the Positivists adopted Hempel's D-N model of scientific explanation; thus beginning the modern discussion of scientific explanation.

The consensus today is that there are indeed scientific explanations. However, regardless of how explanations fare as a part of the scientific enterprise we should demand the proponents of mathematical explanation should face up to the same question that the proponents of scientific explanation must face: i.e., what do explanations [in mathematics] add to our knowledge? Scientific explanations play an integral role in scientific theories and thus contribute to the scientific program, as we outline below. Mathematics on the other hand, has no "theories", certainly not in the same sense that science has.[21] There is thus no explanatory role for our mathematical "explanations" to play. We can see this by exploring the roles that scientific explanations play in science according to various theories of scientific explanation.[22]

Let us begin by considering a physically omniscient (LaPlacian) being P

---

[21] As we noted in §2.4, unlike in the case of scientific theories, number theory for example, is not a theoretical framework under which we subsume the data, inferences, predictions, and explanations of arithmetical "phenomena".

[22] The following argument partially repeats and refers back to some of the arguments put forth in Chapter 2.

who knows all and only true facts about all the objects in the physical world but not the explanations involving them. P only has descriptive knowledge.[23] Consider also a logically omniscient being L who knows all and only the logical and mathematical facts. Assuming that P and L have some reason for seeking explanations about the world, is there something that an explanation can add to their knowledge? In what way is P or L's knowledge incomplete?

Wesley Salmon ([Sal89]: 128-135) argues for each theory of explanation that P gains knowledge when provided with an acceptable explanation of that theory's type. The following shows that none of the answers that can be provided that shows that P gains additional knowledge will apply in a non-trivial way to show that L gains additional knowledge.

(a) Let us start with the modal conception of explanation. As we outlined in our introduction to Chapter 2, the modal conception of explanation shows that, for some event E, it *necessarily had to happen.* Suppose that E occurs, and for all we know, it might not have happened. We can explain E by showing that, given any other circumstances, it had to happen. Many philosophers have advocated modal conceptions of explanations. For those who claim that explanations add a modal dimension to our descriptive and predictive knowledge, explanatory knowledge is knowledge of what is necessary and what is contingent.

In mathematics there is no need to add information about what is neces-

---

[23] P may also need knowledge of concepts, but these have no explanatory powers.

sary because everything is; everything must be as it is. Every mathematical truth has to be true.[24] Thus there is no need to show that whatever mathematical fact is being "explained" had to be that way and could not be any other way. Knowledge of what is necessary and what is impossible in mathematics is useless (to L) because whatever answers questions about necessity will just be standard mathematical proofs.

Mary Leng expressed this problem similarly. In the case of contingent facts, presumably where empirical phenomena are concerned, when we ask why some event E occurred, it is often because we realize that E might not have occurred. So when we ask for an explanation of E we are really seeking other contingent facts and laws that made it the case that E and not, say, F. Therefore, we cannot explain E by deriving it from background assumptions, but rather we need to understand how the assumptions contributed to E. But mathematical truths are necessary. So a proposition p follows from mathematical theory T necessarily. So there is no way to consider how p could have been different ([Len05]).

Thus, on a modal conception of explanation we have no answer to the question "what good are mathematical explanations?" On this conception, P gains knowledge of which facts are necessary and which are contingent. L

---

[24] D. H. Mellor, I believe is expressing similar concerns in his reply to Koslow in [Kos03] (223). He contrasts statements about chance with those of seemingly non-contingent statements. He writes: "... because 'P is contingent' and hence '∼ P is possible' are necessary if true, they need no truthmakers. Similarly for the sense in which truth or falsity are the possible truth values of any 'P' and '∼ P'. Similarly again ... of possible cases invoked in mathematical proofs."

gains nothing because all the facts L dealt with are necessary. (Trivially, he may learn that some mathematical fact is necessary, but then proponents of mathematical explanation must explain why it is not a problem that every fact that L knows has the same explanation.) If explanations add no knowledge, then in this context, postulating their existence violates Ockham's razor. Since violating Ockham's razor is unwarranted, there is no mathematical explanation on the modal conception of explanation so construed.[25]

(b) If you take, with Salmon, an ontic conception of explanation, then the knowledge P would get from an explanation is knowledge of causes. P gets "recognition of the difference between causal and noncausal laws, the difference between causal processes and pseudoprocesses, and the difference between causal interaction and mere spatio-temporal coincidences."(128)

As we argued in Chapter 2.2, there is no way to get causes out of mathematics, nor is there any need to talk about causation in mathematics. Regardless of one's version of causality, there is little to be gained from assuming that mathematicians literally engage in the investigation of causes. "Prove" is not a causal verb. (Even mathematical constructivists do not take it literally that mathematicians cause theorems to be true.) We also argued[26] that those early modern mathematicians who spoke of causal proofs cannot be

---

[25] A similar version of this argument is attributed to Thomas Forester. The argument he gives is as follows: There is no explanation without counterfactuals. There are no counterfactuals without contingency. All of mathematics is necessary. Therefore there are no explanations in mathematics.

[26] See §3.2 above.

taken to have been advocating a program of causation in mathematics along the lines that we think of causation today. And contemporary causation is the version that gives us causal explanations. So the ontic conception of explanation will not be of any use to mathematics either. L thus gains nothing because the model is only offering causal knowledge, which is inapplicable in mathematics.

(c1) The third approach to explanation is the epistemic version of which there are three varieties. The first version of the epistemic view is the inferential version, or the D-N model. By subsuming the explananda under some kind of lawful regularity, the D-N approach provides *nomic expectability.*

It is difficult to see what P gains from this kind of explanation. It is actually a matter of some dispute. But that is not our concern. We are interested in knowing what L gains. Presumably L has no use for nomic expectability. If L is already logically omniscient, then L can learn nothing new because the only thing that we could expect in mathematics is to discover what follows from applying various logical tools to what is already known. But L already has this knowledge, since being logically omniscient *is* having knowledge of everything that follows from applying logical devices to logical truths.

Furthermore, as we argued in Chapter 2.1, L has no laws of the sort that the D-N model could use from which to start explaining anything. So L gains nothing on the first epistemic view.

(c2) The second version of the epistemic view is the information-theoretic version. On this view “the explanatory value of a law or theory is measured in terms of its efficacy in improving our predictive power.” (131) Salmon suggests that this efficacy can be understood in terms of information-theoretic systemization. Thus when we are able to take a great deal of information and encode it in a short “message” we have explained the information by increasing our understanding.

Regardless of our misgivings about the role of understanding in explanations, we saw how this view fails in the case of mathematics when we examined Kitcher’s view in §2.3.1. We saw Hung’s and Humphreys’ examples where mathematical axioms and sequences were compressed, and we saw the category theory proof encompassing many proofs in one - all of which did not generate explanations. The theory of explanation is plausible in the scientific cases, but not the mathematical ones. Any more compact expression of a mathematical theory is also known to L whereas a compact expression of all the information that P has can perhaps be useful. So P can benefit from the systematizing of scientific information by adding organizational rules to the knowledge of the physical facts. L cannot so benefit.

One possible reason for this is that when we can show that many scientific argument patterns can be compressed under one pattern it allows us to show that many sets of phenomena are suddenly - parsimoniously - mathematically tractable. That is, this kind of abstraction may be useful in [our understanding of] science. Whereas in mathematics, adding yet another layer of abstraction to some mathematics rarely makes that bit of mathematics more tractable, more comprehensible, more understood, and only sometimes makes it more useful.

(c3) Finally, on the erotetic approach to explanation, explanations are answers to why-questions. Thus one way of looking at the model is that the only the only difference between explanatory knowledge and descriptive knowledge is pragmatic. Prior to the emergence of van Fraassen's model it was generally taken for granted that explanations must be something other than descriptions of the phenomena they are explaining. By putting the description in a pragmatic context van Fraassen claims to be able to have descriptions as explanations. All explanations then are descriptions, and it is merely a matter of pragmatic concern which descriptions we happen to be interested in at any particular time. On such an account, the why-question approach adds no knowledge to either P or L. And that is the very criticism we raised in Chapter 2.4. However, in the case of mathematics the criticism is especially acute. The answer to any particular mathematical why-question will end up being a proof. Different why-questions will provide different proofs of the same theorem. But L already knows all the proofs of all the theorems - the "explanatory" and the non-explanatory ones. So L does not need explanations of this sort. Presumably an ideal agent who is logically omniscient has no pragmatic contexts. But if we assume that L asked "why

R?", any proof of R would answer the question for L. So getting explanations for R is the same thing as finding a proof. And any proof of R will do. This seems to trivialize the notion of explanation, and reduce "explanation" to proof. This will not do, as explanation in mathematics - by all accounts - cannot merely be proof.

If another being L' asked "why R?", where L' does not know why R is the case because L' is logically omniscient *except for some bits of knowledge*, then explanation would simply consist of filling in a bit of mathematical information. We might have to inform L' that some sequence x is convergent, or that the number 13 is prime, or whatever. But pointing out a property of a sequence (or some other banal mathematical fact) that may or may not be well known can hardly be considered explanatory, except on the most trivial account of explanation.

Thus, each type of explanation finds itself superfluous when it is applied to mathematics.

One reason that mathematical explanations find themselves superfluous is that they do not seem to play the right kind of role in theories. A model of explanation must work in concert with a theory of theories. Explanations do not exist in methodological vacuums. They are parts within the structure of the larger theory of the science in which the explanation is to reside.

If there is explanatory mathematics, they will not play the correct role in our mathematical theories. Explanations in science, at least, generally serve to offer more confirmation of the theory, or under-gird it in some other way. What makes a theory valuable, in part, is that it contains explanations. Theories subsume scientific explanations which in turn subsume the phenomena of the science and include some organizing principle.

Mathematical explanations would offer nothing to encourage us to accept some particular mathematical theory. Even the most obstinate believer in mathematical explanation accepts that proofs are good enough to establish theorems. It is the proven theorems that play a role in a given mathematical theory. This is all for the reasons we stressed above, that theories in mathematics are not the same kinds of things that theories in science are.

We accept a theory in mathematics, or at any rate, we find a theory in mathematics useful to the extent that we can prove theorems in it. The role of explanation is bypassed.

At the minimum, the burden is thus on the proponent of mathematical explanation to show that the mathematical theory they are working with would not be accepted without the explanations. This however is makes no sense for mathematics, nor is there reason to believe that the history of mathematics will bear this out.

4.7 Conclusion

To justify the conclusion that there are mathematical explanations one would need a model of explanation that makes mathematical results “accounted for” in a way that ordinary proofs fail to do. Whatever this extra mathematical information or method is, it must be independent of proof or consist of a special kind of proof. But because of the nature of proof, explanation, and mathematics, for the reasons we have elaborated throughout this work, it seems unlikely that there is an account that meets our explanatory needs. In a proof, once we say correctly something of the form p follows from B, we have just given a proof of p. There is nothing of mathematical importance beyond that.

Some models of scientific explanation have been criticized for offering nothing but a way to repeat the descriptions of scientific phenomenon. Explanations must go beyond such repetition, and that is why philosophers have inserted for example, deduction, causation, or compression in their models of explanation. Mathematical proofs are different from both descriptions and explanations. Descriptions tell us what happened or what obtains. Explanations account for how and/or why. Proofs in mathematics, however interesting or sophisticated, are epistemic devices that show how theorems fit together and function within a mathematical system. They do not otherwise account for the necessary fact of their existence or truth.[27]

Given the above arguments against the applicability of the existing models of explanation to mathematics and against the efficacy of the existing models of mathematical explanation, combined with the difficulty of the phi-

---

[27] This is not intended to suggest any metaphysical commitments.

losophy of science to settle on a model of explanation, there is not likely to be any good account of mathematical explanation in the future. Any account of explanation for mathematics would have to overcome the difficulties we presented above. It would have to show that (1) either there is indeed a history of explanations in mathematics or that there is a good reason that they were never noticed, and (2) either there is a reasonable account of prediction in mathematics or prediction is not a serious desideratum in explanation, and (3) it can explain away the methodological disparity between what and how much we expect to fall under a mathematical as opposed to a scientific explanation, and (4) the reduction of surprise is not a desideratum in mathematics explanation, and (5) there is some function that mathematical explanation serves for mathematics. Alternatively, to make mathematical explanation plausible, mathematicians, philosophers of science and philosophers of mathematics would have to completely reconceptualize their understanding of explanations in terms that are radically different from those we currently use. But as the Bourbaki slogan goes: "*Depuis les Grecs qui dit 'Mathematique' dit 'démonstration'.*" For the most part we agree with this austere view of mathematics and believe that there is no good reason to expect that we will also find explanations.

Bibliography

[Abe76] Reuben Abel. *Man is the Measure: A Cordial Invitation to the Central Problems of Philosophy.* The Free Press, New York, 1976.

[Ach83] Peter Achinstein. *The Nature of Explanation.* Oxford University Press, New York, 1983.

[Aga81] Joseph Agassi. Lakatos on proof and mathematics. *Logique et Analyse*, 95-96:437–439, Septembre-Décembre 1981.

[AK88] William Asprey and Philip Kitcher, editors. *History and Philosophy of Modern Mathematics.* Minnesota Studies in the Philosophy of Science. University of Minnesota Press, Minneapolis, 1988.

[Ans93] G. E. M Anscombe. Causality and determination. In Ernest Sosa and Michael Tooley, editors, *Causation*, pages 88–104. Oxford University Press, New York, 1993.

[Apo52] Hippocrates George Apostle. *Aristotle's Philosophy of Mathematics.* The University of Chicago Press, Chicago, 1952.

[Apo61] Leo Apostel. Towards the formal study of models in the non-formal sciences. In Hans Freudenthal, editor, *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*, pages 1–37. D. Reidel Publishing Company, Holland, 1961.

[Ari52] Aristotle. *Metaphysics*. Columbia University Press, New York, 1952. Translated by Richard Hope.

[Ari75] Aristotle. *Posterior Analytics*. Oxford, London, 1975. Translated by Jonathan Barnes.

[Avi06] Jeremy Avigad. Mathematical method and proof. *Synthese*, 153:105–159, 2006.

[Aye36] Alfred Jules Ayer. *Language, Truth and Logic*. Gollancz, London, 2nd edition, 1936. reprinted, New York: Dover, 1946.

[Azz00] Jody Azzouni. *Knowledge and Reference in Empirical Science*. Routledge, New York, 2000.

[Bac80] Bachye ben Asher. *Encyclopedia of Torah Thoughts*. Shilo Publishing House, New York, 1980.

[Bar34] Isaac Barrow. *The Usefulness of Mathematical Learning Explained and Demonstrated: Being Mathematical Lectures Read in the Publick Schools at the University of Cambridge*. Stephen Austin, London, 1734. Translated by John Kirkby. Reprinted by Frank Cass & Co. Ltd. (London, 1970).

[Bar04] John D. Barrow. Mathematical explanation. In John Cornwell, editor, *Explanations: Styles of Explanation in Science*, chapter 5, pages 81–109. Oxford, New York, 2004.

[Bec71] Petr Beckmann. *A History of π (Pi)*. The Golem Press, New York, 1971.

[Ben99] Donald C. Benson. *The Moment of Proof*. Oxford, New York, 1999.

[BGT91] Richard Boyd, Philip Gasper, and J. D. Trout, editors. *The Philosophy of Science*. The MIT Press, Cambridge, MA, 1991.

[Bla97] D. Blatner. *The Joy of Pi*. Walker Publishing, New York, 1997.

[Boa30] George Boas. *A Critical Analysis of the Philosophy of Emile Meyerson*. The Johns Hopkins Press, Baltimore, 1930.

[Boh52] Harald Bohr. Address of professor Harald Bohr. In International Congress of Mathematicians, editor, *Proceedings of the International Congress of Mathematicians, August 30 - September 6 1950*, pages 127–134, Cambridge, MA, 1952. American Mathematical Society.

[Bol04] Bernard Bolzano. *Betrachtungen über Einige Gegenstände der Elementargeometrie*. Karl Barth, Prague, 1804. English translation by Steven Russ in [Ewa96] vol I, pp 172-4 as "Preface to Considerations on Some Objects of Elementary Geometry".

[Bol10] Bernard Bolzano. *Beiträge Zu Einer Begründeteren Darstellung der Mathematik*. Caspar Widtmann, Prague, 1810. English translation by Steven Russ in [Ewa96] vol. 1, pp. 174-224, as "Contributions to a Better-Grounded presentation of Mathematics".

[Bol17] Bernard Bolzano. *Rein Analytischer Beweis Des Lehrsatzes, Dass Zwischen Je Zwei Werten, Die ein Entgegengesetztes Resultat Gewähren, Wenigstens eine Reelle Wurzel der Gleichung Liege*. Gottlieb Haase, Prague, 1817. English translation by Steven Russ in [Ewa96], vol. 1, pp.225-48 as "Purely Analytic Proof of the Theorem that Between Any Two Values Which Give Results of Opposite Sign There Lies at Least One Real Root of the Equation".

[Bol72] Bernard Bolzano. *Theory of Science: Attempt at a Detailed and in the Main Novel Exposition of Logic With Constant Attention to Earlier Authors*. Oxford, Great Britain, 1972. Edited and Translated by Rolf George.

[Bou50] Nicolas Bourbaki. The architecture of mathematics. *American Mathemaical Monthly*, 57:221–232, 1950. Originally published in French in 1948. Translation by Arnold Dresdin. Reprinted in [Ewa96], vol. 2, p1265-76.

[BP83] Paul Benacerraf and Hilary Putnam, editors. *Philosophy of Mathematics: Selected Readings.* Cambridge, Cambridge, MA, second edition, 1983.

[Bra53] Richard Braithwaite. *Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science.* Cambridge University Press, London, 1953.

[Bro56] T. A. A. Broadbent. Review of *Mathematics and Plausible Reasoning* by G. Polya. *The Mathematical Gazette*, 40(333):233–234, October 1956.

[Bro97] James Robert Brown. Proofs and pictures. *British Journal for the Philosophy of Science*, 48:161–180, 1997.

[Bro99] James Robert Brown. *Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures.* Routledge, New York, 1999.

[Bus56] Robert R. Bush. Review of Polya's *Mathematics and Plausibele Reasoning.* *The American Journal of Psychology*, 69(1):166–167, March 1956.

[But01] Samuel John Butchart. *Evidence and Explanation in Mathematics.* PhD thesis, Monash University, 2001.

[Cam57] Norman Robert Campbell. *Foundations of Science.* Dover Publications, New York, 1957.

[Car66] Rudolf Carnap. *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science.* Basic Books, New York, 1966. Republished in 1995 by Dover Publications, Inc. New York.

[Car06] Nancy Cartwright. Where is the theory in our "theories" of causality? *The Journal of Philosophy*, CIII(2):55–66, February 2006.

[Cas88] D. Castellanos. The ubiquitous pi. *Mathematics Magazine*, 61, 1988.

[Cha75] Gregory J. Chaitin. Randomness and mathematical proof. *Scientific American*, pages 47–52, May 1975.

[Chi82] Roderick M. Chisholm. *The Foundations of Knowing*. University of Minnesota Press, Minneapolis, 1982.

[Chu89] P. Churchland. *A Neurocomputational Perspective*. MIT Press, Cambridge, MA, 1989.

[Col01] Mark Colyvan. *The Indispensability of Mathematics*. Oxford University Press, New York, 2001.

[Cor03] David Corfield. *Towards a Philosophy of Real Mathematics*. Cambridge University Press, Cambridge, 2003.

[Cro97] Alfred W. Crosby. *The Measure of Reality: Quantification and Western Society 1250-1600*. Cambridge University Press, New York, 1997.

[Dav67] Donald Davidson. Causal relations. In Donald Davidson, editor, *Essays on Actions and Events*, pages 149–162. Oxford, New York, 1967.

[Daw06] John W. Dawson, Jr. Why do mathematicians re-prove theorems? *Philosophia Mathematica*, 14(3):269–286, October 2006.

[Des73] Jean-Toussaint Desanti. L'explication en mathématiques. In L. Apostel *et al.*, editor, *L'explication Dans Les Sciences*. Flammarion, Paris, 1973.

[Det88] Michael Detlefsen. Fregean heirarchies and mathematical explanation. *International Studies in the Philosophy of Science*, 3:97–116, 1988.

[DH81] Philip J. Davis and Reuben Hersh. *The Mathematical Experience*. Houghton Mifflin, Boston, 1981.

[Dho93] Jean Dhombres. Is one proof enough? Travels with a mathematician of the baroque period. *Educational Studies in Mathematics*, 24:401–419, 1993.

[Duh54] Pierre Duhem. *The Aim and Structure of Physical Theory*. Princeton University Press, Princeton, NJ, 1954.

[EAFD97] Dawn E. Mccormick Ellice A. Forman and Richard Donato. Learning what counts as a mathematical explanation. *Linguistics and Education*, 9(4):313–339, 1997.

[Ein54] Albert Einstein. *Ideas and Opinions*. Crown Publishers, New York, 1954.

[Eps02] Richard L. Epstein. *Five Ways of Saying “Therefore”: Arguments, Proofs, Conditionals, Cause and Effect, Explanations*. Wadsworth, Belmont, CA, 2002.

[Ewa96] William Ewald, editor. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Oxford, New York, 1996.

[Ezr95] Abraham Ibn Ezra. *The Secret of the Torah: A Translation of Abraham Ibn Ezra’s Sefer Yesod Morah Ve-Yesod Ha-Torah*. Aronson, Northvale, NJ, 1995. Translated by H. Norman Strickman.

[Fey99] Richard P. Feynman. *The Pleasure of Finding Things Out*. Perseus Publishing, Cambridge, MA, 1999.

[Fra80] Bas C. Van Fraassen. *The Scientific Image*. Oxford, New York, 1980.

[Fri74] Michael Friedman. Explanation and scientific understanding. *The Journal of Philosophy*, 71(1):5–19, January 1974. Reprinted in [Pit88] p 188-198.

[Gal90] David Gale. Proof as explanation. *The Mathematical Intelligencer*, 12(1):4, 1990.

[Gar81] Alan Garfinkel. *Forms of Explanation: Rethinking the Questions in Social Theory.* Yale University Press, Hew Haven, CT, 1981.

[Ger99] Gersonides (Levi Ben Gershom). *The Wars of the Lord*, volume 3. The Jewish Publication Society, Philadelphia, 1999. Translated by Seymour Feldman.

[Gia05] M. Giaquinto. Mathematical activity. In Paolo Mancosu, Klaus Frovin Jorgensen, and Stig Andur Pedersen, editor, *Visualization, Explanation and Reasoning Styles in Mathematics*, pages 75–86. Springer, The Netherlands, 2005.

[Gle88] James Gleick. *Chaos: The Making of a New Science.* Penguin Books, New York, 1988.

[Goo65] R. L. Goodstein. *Proof by Reductio Ad Absurdum*, chapter 1, pages 1–11. Leicester University Press, 1965.

[Gra81] Judith V. Grabiner. Changing attitudes toward mathematical rigor: Lagrange and analysis in the eighteenth and nineteenth centuries. In H. N. Jahnke and M. Otte, editors, *Epistemological and Social Problems of the Sciences in the Early Nineteenth Century*, pages 311–330. D. Reidel Publishing Company, 1981.

[Gro80] Emily R. Grosholz. Descartes' unification of algebra and geometry. In Stephen Gaukroger, editor, *Descartes: Philosophy, Mathematics and Physics*, chapter 6, pages 156–168. The Harvester Press, Sussex, 1980.

[Gro81] Emily Grosholz. Wittgenstein and the correlation of logic and arithmetic. *Ratio*, pages 31–42, 1981.

[Gro85] Emily R. Grosholz. Two episodes in the unification of logic and topology. *The British Journal for the Philosophy of Science*, 36(2):147–157, June 1985.

[Gro92] Emily Grosholz. Was Leibniz a mathematical revolutionary? In Donald Gillies, editor, *Revolutions in Mathematics*, chapter 7, pages 117–133. Oxford University Press, New York, 1992.

[Gro00] Emily R. Grosholz. The partial unification of domains, hybrids, and the growth of mathematical knowledge. In Emily Grosholz and Herbert Breger, editors, *The Growth of Mathematical Knowledge*, pages 81–91, Dordrecht, 2000. Kluwer Academic Publishers.

[Had45] Jacques Hadamard. *An Essay on the Psychology of Invention on the Mathematical Field*. Princeton University Press, Princeton, NJ, 1945.

[Haf99] Johannes Hafner. Bolzano's criticism of indirect proofs. *Revue d'histoire des sciences*, 52(3-4):385–398, Juillet-Décembre 1999.

[Han06] Bengt Hansson. Why explanations? fundamental, and less fundamental ways of understanding the world. *Theoria*, LXXII(1):23–59, 2006.

[Har67] G. H. Hardy. *A Mathematician's Apology*. Cambridge University Press, Cambridge, England, 1967.

[Hei77] Martin Heidegger. *Basic Writings*. Harper Collins Publishers, San Francisco, revised and expanded edition, 1977.

[Hem63] Carl G. Hempel. Explanation and prediction by covering laws. In Bernard Baumrin, editor, *Philosophy of Science: The*

Delaware Seminar, pages 107–133, New York, 1963. John Wiley and Sons.

[Hem65a] Carl G. Hempel. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* Free Press, New York, 1965.

[Hem65b] Carl G. Hempel. The function of general laws in history. In *Aspects of Scientific Explanation*, pages 231–243. Free Press, New York, 1965.

[Hem66] Carl G. Hempel. *Philosophy of Natural Science.* Prentice Hall, Englewood Cliffs, NJ, 1966.

[Hen05] Richard Conn Henry. The mental universe. *Nature*, 436(7):29, July 2005.

[Her79] Ruben Hersh. Some proposals for reviving the philosophy of mathematics. *Advances in Mathematics*, 31:31–50, 1979. Reprinted in [Tym98].

[Her93] Ruben Hersh. Proving is convincing and explaining. *Educational Studies in Mathematics*, 24:389–399, 1993.

[Her97] Reuben Hersh. *What is Mathematics, Really?* Jonathan Cape, London, 1997.

[Hir89] Morris W. Hirsch. Chaos, rigor, and hype. *The Mathematical Intelligencer*, 11(3):6–9, Summer 1989. With a reply by James Gleick.

[HM05] J. Hafner and Paolo Mancosu. The varieties of mathematical explanation. In Klaus Frovin Jorgensen Paolo Mancosu and Stig Andur Pedersen, editors, *Visualization, Explanation and Reasoning Styles in Mathematics*, pages 1–32. Springer, 2005.

[HO48] Carl G. Hempel and Paul Oppenheim. Studies in the logic of explanation. *Philosophy of Science*, 15:135–175, 1948. Reprinted in [Pit88] and with a postscript in [Hem65a].

[Hol81] Gerald Holton. Thematic presupposition and the direction of scientific advance. In A. F. Heath, editor, *Scientific Explanation*, pages 1–27. Oxford University Press, New York, 1981.

[Hor93] John Horgan. The death of proof. *Scientific American*, 269(4):92–103, 1993.

[HR74] Jaakko Hintikka and Unto Remes. *The Method of Analysis: Its Geometrical Origin and Its General Significance*. D. Reidel Publishing Company, Boston, 1974.

[Hum77] David Hume. *Enquiry Concerning Human Understanding*. Oxford; The Clarendon Press, 1777. Reprinted in 1975.

[Hum80] David Hume. *Dialogues Concerning Natural Religion*. Hackett, Indianapolis, 1980. Contains the Posthumous Essays. Edited by Richard H. Popkin.

[Hum89] Paul Humphreys. *The Chances of Explanation: Causal Explanation in the Social, Medical, and Physical Sciences*. Princeton University Press, Princeton, NJ, 1989.

[Hum93] Paul Humphreys. Greater unification equals greater understanding? *Analysis*, 53(3):183–188, July 1993.

[Hun97] Edwin Hung. *The Nature of Science: Problems and Perspectives*. Wadsworth Publishing Co, Belmont, CA, 1997.

[Hux65] George Huxley. Greek mathematics and astronomy. In Hugh Lloyd-Jones, editor, *The Greek World*, pages 142–167. Penguin Books, Baltimore, MD, 1965.

[Jac06] Dale Jacquette. Review of [MJP05]. *The Mathematical Intelligencer*, 28(2):79–81, 2006.

[JBM99] Paul Cobb Janet Bowers and Kay McClain. The evolution of mathematical practices: A case study. *Cognition and Instruction*, 17:25–64, 1999.

[KA88] Philip Kitcher and William Aspray. An opinionated introduction. In William Aspray and Philip Kitcher, editors, *History and Philosophy of Modern Mathematics*, pages 3–57. The University of Minnesota Press, Minneapolis, 1988.

[Kah63] Russell Kahl, editor. *Studies in Explanation*. Prentice Hall, Englewood Cliffs, NJ, 1963.

[Kim94] Jaegwon Kim. Explanatory knowledge and metaphysical dependence. *Philosophical Issues*, 5:51–69, 1994.

[Kit75] Philip Kitcher. Bolzano's ideal of algebraic analysis. *Studies in the History and Philosophy of Science*, 6:229–269, 1975.

[Kit76] Philip Kitcher. Explanation, conjunction, and unification. *Journal of Philosophy*, 73:207–212, 1976.

[Kit81a] Philip Kitcher. Explanatory unification. *Philosophy of Science*, 48:507–531, 1981. Reprinted in [BGT91] 329-347, and [Pit88] p167-187.

[Kit81b] Philip Kitcher. Mathematical rigor - who needs it? *Noûs*, pages 469–493, 1981.

[Kit83] Philip Kitcher. *The Nature of Mathematical Knowledge*. Oxford, New York, 1983.

[Kle11] Felix Klein. *The Evanston Colloquium: Lectures on Mathematics*. American Mathematical Society, New York, 1911.

[Kle97] Robert Klee. *Introduction to the Philosophy of Science*. Oxford University Press, New York, 1997.

[Kli67] Morris Kline. *Mathematics for the Nonmathematician*. Addison-Wesley Publishing Company, Reading, MA, 1967.

[Koe92] Noretta Koertge. Explanation and its problems. *British Journal for the philosophy of science*, 43:85–98, 1992.

[Kos03] Arnold Koslow. Laws, explanations and the reduction of possibilities. In G. Rodriguez-Pereyra and H. Lillehammer, editors, *Real Metaphysics: Essays in Honour of Hugh Mellor*, chapter 11, pages 169–183. Routledge Press, 2003. With a reply by D. H. Mellor, p 232-234.

[KS89] Philip Kitcher and Wesley Salmon, editors. *Scientific Explanation.* Minnesota Studies in the Philosophy of Science. University of Minnesota Press, Minneapolis, 1989.

[Lak76] Imre Lakatos. *Proofs and Refutations.* Cambridge University Press, New York, 1976.

[Lak78] Imre Lakatos. A renaissance of empiricism in the recent philosophy of mathematics? In J. Worall and G. Curry, editors, *Philosophical Papers*, volume 2. Cambridge University PRess, Cambridge, 1978. All references are to the reprint in [Tym98].

[Lam95] Leslie Lamport. How to write a proof. *American Mathematical Monthly*, 102:600–608, August-September 1995.

[Law69] F. William Lawvere. Diagonal arguments and cartesian closed categories. *Springer Lecture Notes in Mathematics*, (92):134–145, 1969.

[LBB04] J. M. Borwein L. Berggren and P. B. Borwein. *Pi: A Sourcebook.* Springer, New York, 2004.

[LC08] Aidan Lyon and Mark Colyvan. The explanatory power of phase spaces. *Philosophia Mathematica (III)*, 16(2):227–243, June 2008.

[Len05] Mary Leng. Mathematical explanation. In C. Cellucci and D. Gillies, editors, *Mathematical Reasoning and Heuristics*, pages 167–189. King's College Publications, London, 2005.

[Lev77] Michael Levin. Explanation and prediction in grammar (and semantics). In *Midwest Studies in Philosophy*, pages 128–137. 1977.

[Lev01] Joseph Levine. *Purple Haze: The Puzzle of Consciousness*. Oxford, New York, 2001.

[Lew86] David Lewis. Causal explanation. In *Philosophical Papers: Vol II*, pages 214–240. Oxford University Press, New York, 1986.

[LJ70] Karel Lambert and Gordon G. Brittan Jr. *An Introduction to the Philosophy of Science*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1970.

[LSL02] Maureen Loomes, Alex Shafarenko, and Martin Loomes. Teaching mathematical explanation through audiographic technology. *Computer Education*, 38(1-3):137–149, 2002.

[Mac93] Ernst Mach. *The Science of Mechanics: A Critical and Historical Account of Its Development*. Open Court, La Salle, 1893. Illinois.

[Mac74] J. L. Mackie. *The Cement of the Universe: A Study of Causation*. Oxford, London, 1974.

[Mad90] Penelope Maddy. *Realism in Mathematics*. Oxford University Press, New York, 1990.

[Mah90] Michael S. Mahoney. Infinitesimals and transcendent relations: The mathematics of motion in the late seventeenth century. In David C. Lindberg and Robert S. Westman, editors, *Reappraisals of the Scientific Revolution*, pages 461–492. Cambridge University Press, Cambridge, 1990.

[Man91] Paolo Mancosu. On the status of proofs by contradiction in the seventeenth century. *Synthese*, 88:15–41, 1991.

[Man92] Paolo Mancosu. Descartes's *Géométrie* and revolutions in mathematics. In Donald Gillies, editor, *Revolutions in Mathematics*, pages 83–116. Oxford University Press, New York, 1992.

[Man96] Paolo Mancosu. *Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century*. Oxford University Press, New York, 1996.

[Man99a] Paolo Mancosu. Bolzano and Cournot on mathematical explanation. *Revue d'histiore des sciences*, 52(3-4):429–455, Juillet-Décembre 1999.

[Man99b] Paolo Mancosu. Literature survey: Recent publications in the history and philosophy of mathematics from the Renaissance to Berkeley. *Metascience*, 8(2):102–124, 1999.

[Man00] Paolo Mancosu. On mathematical explanation. In E. Grosholz and Herbert Breger, editors, *The Growth of Mathematical Knowledge*, pages 103–119, Dordrecht, 2000. Kluwer Academic Publishers.

[Man01] Paolo Mancosu. Mathematical explanation: Problems and prospects. *Topoi*, 20:97–117, 2001.

[Man08a] Paolo Mancosu. Mathematical explanation: Why it matters. In Paolo Mancosu, editor, *The Philosophy of Mathematical Practice*, pages 135–150. Oxford University Press, New York, 2008.

[Man08b] Paolo Mancosu. Explanation in mathematics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2008.

[Maz98] Massimo Mazzotti. The geometers of God: Mathematics and reaction in the kingdom of Naples. *Isis*, 89(4):674–701, December 1998.

[McL90] Colin McLarty. The uses and abuses of the history of topos theory. *British Journal for the Philosophy of Science*, 41:351–375, 1990.

[Men97] Elliot Mendelson. *Introduction to Mathematical Logic.* Chapman & Hall, London, fourth edition, 1997.

[Mic90] Ronald C. Mickens, editor. *Mathematics and Science.* World Scientific, Teaneck, NJ, 1990.

[Mil41] John Stuart Mill. *A System of Logic.* Longmans, Green and Co., London, eighth edition, 1941.

[MJP05] Paolo Mancosu, Klaus Frovin Jorgensen, and Stig Andur Pedersen, editors. *Visualization, Explanation and Reasoning Styles in Mathematics*, Dordrecht, 2005. Springer.

[Moo59] G. E. Moore. *Philosophical Papers.* George Allen & Unwin, London, 1959.

[MV90] Paolo Mancosu and Ezio Vailati. Detleff Clüver: An early opponent of the Leibnizian differential calculus. *Cenataurus*, 33:325–344, 1990.

[MV91] Paolo Mancosu and Ezio Vailati. Torricelli's infinitely long solid and its philosophical reception in the seventeenth century. *Isis*, 82:50–70, 1991.

[Nag61] Ernest Nagel. *The Structure of Science: Problems in the Logic of Scientific Explanation.* Routledge & Kegan Paul, London, 1961.

[Ner79] Grahm Nerlich. What can geometry explain? *The British Journal for the Philosophy of Science*, 30(1):69–83, March 1979.

[Net70] Gwynn Nettler. *Explanations.* McGraw-Hill, New York, 1970.

[Net05] R Netz. The aesthetics of mathematics: A case study. In Klaus Frovin Jorgensen Paolo Mancosu and Stig Andur Pedersen, editors, *Visualization, Explanation and Reasoning Styles in Mathematics*, pages 251–293. Springer, Dordrecht, 2005.

[NS00] W. H. Newton-Smith. Explanation. In W. H. Newton-Smith, editor, *A Companion to the Philosophy of Science*, pages 127–133. Blackwell, New york, 2000.

[Oka02] Samir Okasha. *Philosophy of Science: A Very Short Introduction.* Oxford University Press, New York, 2002.

[Osb04] Catherine Osborne. *Presocratic Philosophy: A Very Short Introduction.* Oxford University Press, New York, 2004.

[Par87] Rohit Parikh. Knowledge and the problem of logical omniscience. In Z. W. Ras and M. Zemankova, editors, *Methodologies for Intelligent Systems*, pages 432–439. Elsevier Science Publishing Co., Inc., 1987.

[Pit88] Joseph C. Pitt, editor. *Theories of Explanation.* Oxford University Press, New York, 1988.

[Pol49] G. Polya. With, or without, motivation? *American Mathematical Monthly*, 56:684–691, 1949.

[Pol54] G. Polya. *Mathematics and Plausible Reasoning.* Princeton University Press, Princeton, NJ., 1954. 2 volumes. Volume I - Induction and Analogy in Mathematics. Volume II - Patterns of Plausible Inference.

[Pol57] G. Polya. *How To Solve It.* Princeton University Press, Princeton, New Jersey, second edition, 1957.

[Pol87] Stephen Pollard. What is abstraction? *Noûs*, XXI(2):233–240, June 1987.

[Pop59] Karl R. Popper. *The Logic of Scientific Discovery.* Harper & Row, New York, 1959. Republished by Routledge in 2002. References are to this edition.

[Pri02] Graham Priest. *Beyond the Limits of Thought.* Oxford University Press, New York, 2002.

[Psi02] Stathis Psillos. *Causation and Explanation.* McGill-Quee's University Press, Ithaca, 2002.

[Qui78] W. V. O. Quine. *Set Theory and Its Logic.* Harvard University Press, Cambridge, 1978.

[Rav99] Yehuda Rav. Why do we prove theorems? *Philosophia Mathematica*, 7:5–41, February 1999.

[Res70] Nicholas Rescher. *Scientific Explanation.* Free Press, New York, 1970.

[Rio06] Sally Riordan. Explaining mathematical truths with a switch of structure. *Rerum Causae*, 1(1):22–29, 2006.

[RK87] Michael D. Resnik and David Kushner. Explanation, independence and realism in mathematics. *British Journal for the Philosophy of Science*, 38:141–158, 1987.

[Ros00] Alex Rosenberg. *Philosophy of Science: A Contemporary Introduction.* Routledge, New york, 2000.

[Row86] David E. Rowe. "Jewish mathematics" at Göttingen in the era of Felix Klein. *Isis*, 77:422–449, 1986.

[Rub90] David-Hillel Ruben. *Explaining Explanation.* Routledge, New York, 1990.

[Rus07] Bertrand Russell. The regressive method of discovering the premises of mathematics. In Douglas Lackey, editor, *Essays in Analysis*, pages 272–283. George Allen & Unwin, London, 1907.

[Sal84] Wesley C. Salmon. *Scientific Explanation and the Causal Structure of the World.* Princeton University Press, Princeton, 1984.

[Sal89] Wesley C. Salmon. Four decades of scientific explanation. In Philip Kitcher and Wesley C. Salmon, editors, *Scientific Explanation*, Minnesota Studies in the Philosophy of Science, pages 3–219. University of Minnesota Press, Minneapolis, 1989. Reprinted separately in 1990. References are to the 1990 edition.

[Sal92] Wesley C. Salmon. Explanation. In Jonathan Dancy and Ernest Sosa, editors, *A Companion to Epistemology*, pages 129–132. Blackwell Publishing Co., New York, 1992.

[San97] David A. Sandborg. *Explanation in Mathematical Practice.* PhD thesis, University of Pittsburgh, 1997.

[San98] David Sandborg. Mathematical explanations and the theory of why-questions. *British Journal for the Philosophy of Science*, 49:603–624, 1998.

[Saw55] W. W. Sawyer. *Prelude to Mathematics.* Penguin Books, London, 1955.

[Scr75] Michael Scriven. Causation as explanation. *Noûs*, 9:3–16, 1975.

[SEG$^+$92] Merrilee H. Salmon, John Earman, Clark Glymour, James G. Lennox, Peter Machamer, J. E. McGuire, John D. Norton, Wesley C. Salmon, and Kenneth F. Schaffner. *Introduction to the Philosophy of Science.* Prentice Hall, New Jersey, 1992. Republished by Hackett Publishing, Indianapolis, IN, 1999.

[Seg03] Sanford L. Segal. *Mathematicians under the Nazis.* Princeton University Press, Princeton, NJ, 2003.

[Sma90] J. J. C. Smart. Explanation–opening address. In Dudley Knowles, editor, *Explanation and its Limits*, pages 1–19. Cambridge University Press, Cambridge, 1990.

[ST93] Ernest Sosa and Michael Tooley, editors. *Causation*. Oxford University Press, New York, 1993.

[Ste78a] Mark Steiner. Mathematical explanation. *Philosophical Studies*, 34:135–151, 1978.

[Ste78b] Mark Steiner. Mathematics, explanation, and scientific knowledge. *Noûs*, 12:17–28, 1978.

[Ste79] Mark Steiner. Quine and mathematical reduction. In Robert W. Shahan and Chris Swoyer, editors, *Essays on the Philosophy of W. V. Quine*, chapter 9, pages 133–144. Harvester Press, Sussex, 1979.

[Ste83a] Mark Steiner. Mathematical realism. *Noûs*, pages 363–385, 1983.

[Ste83b] Mark Steiner. The philosophy of mathematics of Imre Lakatos. *The Journal of Philosophy*, LXXX(9):502–521, September 1983.

[Ste89] Mark Steiner. The application of mathematics to natural science. *The Journal of Philosophy*, LXXXVI(9):449–480, September 1989. There is a Heberw version of this article in *Iyyun: A Jerusalem Philosophical Quarterly*.

[Ste95] Mark Steiner. The applicibilities of mathematics. *Philosophia Mathematica (3)*, 3:129–156, 1995.

[Ste98] Mark Steiner. *The Applicability of Mathematics as a Philosophical Problem*. Harvard, Cambridge, MA, 1998.

[Ste00] Mark Steiner. Penrose and platonism. In Emily Grosholz and Herbert Breger, editors, *The Growth of Knowledge*, pages 133–141, Dordrecht, 2000. Kluwer Academic Publishers.

[Ste07] Ian Stewart. *Why Beauty is Truth: A History of Symmetry*. Perseus, New York, 2007.

[Stu97] David J. Stump. Reconstructing the unity of mathematics circa 1900. *Perspectives on Science*, 5(3):383–417, Fall 1997.

[Swe87] Frank J. Swetz. *Capitalism and Arithmetic: The New Math of the 15th Century.* Open Court, La Salle, Illinois, 1987.

[Swo99] Chris Swoyer. How ontology might be possible: Explanation and inference in metaphysics. In Peter French and Howard Wettstein, editors, *Midwest Studies in Philosophy. Vol XXIII: New Directions in Philosophy*, pages 100–130. Blackwell, London, 1999.

[Syl96] James Joseph Sylvester. Presidential address to section 'A' of the British Association. In William Ewald, editor, *From Kant to Hilbert: A Sourcebook in the Foundations of Mathematics*, volume I, chapter 12, pages 511–522. Oxford, New York, 1996.

[Tap05] Jamie Tappenden. Proof style and understanding in mathematics I: Visualization, unification and axiom choice. In Klaus Frovin Jorgensen Paolo Mancosu and Stig Andur Pedersen, editors, *Visualization, Explanation, and Reasoning Styles in Mathematics*, pages 147–214. Springer, Dordrecht, The Netherlands, 2005.

[Tha99] Paul Thagard. *How Scientists Explain Disease.* Princeton University Press, Princeton, NJ, 1999.

[Tro02] J. D. Trout. Scientific explanation and the sense of understanding. *Philosophy of Science*, 69:212–233, June 2002.

[Tro07] J. D. Trout. The psychology of scientific explanation. *Philosophy Compass*, 2(3):564–591, 2007.

[Tym79] Thomas Tymoczko. The four-color theorem and its philosophical significance. *The Journal of Philosophy*, lxxvi:57–83, 1979. Reprinted in [Tym98].

[Tym98] Thomas Tymoczko, editor. *New Directions in the Philosophy of Mathematics.* Princeton University Press, Princeton, New Jersey, 2nd edition, 1998.

[WA96] Rega Wood and Robert Andrews. Causality and demonstration: An early scholastic *Posterior Analytics* commentary. *The Monist*, 79(3):325–356, 1996.

[Wai82] Fredrich Waismann. *Lectures on the Philosophy of Mathematics.* Rodopi, Amsterdam, 1982.

[Whe60] Philip Wheelwright, editor. *The Presocratics.* Bobbs-Merrill, Indianapolis, 1960.

[Wig60] Eugene P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics*, 13(1), February 1960. Reprinted in [Wig67].

[Wig67] Eugene P. Wigner. *Symmetries and Reflections.* Indiana University Press, Bloomington, 1967.

[Wil66] L. Pearce Williams. *The Origins of Field Theory.* Random House, New York, 1966.

[Wit83] Ludwig Wittgenstein. *Remarks on the Foundations of Mathematics.* MIT Press, Cambridge, MA, 1983.

[Woo93] James Woodward. A theory of singular causal explanation. In David-Hillel Ruben, editor, *Explanation.* Oxford University Press, New York, 1993.

[Woo03] James Woodward. Scientific explanation. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Summer 2003.

[WV02] Erik Weber and Liza Verhoeven. Explanatory proofs in mathematics. *Logique et Analyse*, 179-180:299–307, 2002.