Dimensions of Mathematical Explanation

BY

WILLIAM B. D'ALESSANDRO
B.A., Canisius College, 2006

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Philosophy
in the Graduate College of the
University of Illinois at Chicago, 2017

Chicago, Illinois

Defense Committee:

Daniel Sutherland, Chair and Advisor
Mahrad Almotahari
David Hilbert
Kenny Easwaran, Texas A&M University
Marc Lange, University of North Carolina at Chapel Hill

Preface

I first crossed paths with mathematical explanation in 2009, my first year of graduate school. I was browsing at random through the newest issue of *Analysis* and happened to read Marc Lange's "Why Proofs by Induction Are Generally Not Explanatory". For one reason or another, the issue struck me immediately as something I should think more about, and within a few weeks I'd written an earnest reply defending the possibility of explanatory circles in pure math. The paper was naïve and not very good—I knew nothing about the larger literature at the time. But Marc read it and wrote me a very nice email in response. (I'm at a loss to explain why, but he continues to treat my fumbling attempts with similar indulgence even today.)

I'd originally planned to study philosophy of physics in grad school, so I was taking math classes to improve my science chops. But I soon realized that I found pure math plenty fun and interesting by itself and changed focus accordingly. When it came time to settle on a dissertation project, mathematical explanation was an easy choice—I'd already started reading and thinking about it, and it connected up with some other interests of mine in metaphysics, epistemology and philosophy of science. (The main collision of ideas that set the dissertation in motion involved, on the one hand, some thoughts about intertheoretic reduction that I'd picked up from Colin Klein, and on the other, the abundance of interesting intertheory relations in math, and their seemingly complex entanglement with mathematical explanation.) So that was that. In retrospect, it was a better choice than I could've realized at the time. The subject was developed enough that there was plenty of interesting literature to digest, and most of the hard work of justifying the existence of the problem had already been done. But the surface had barely been scratched on all sorts of important questions. So I've had no shortage of opportunities to explore new territory.

My journey through grad school has been long but happy, and I'm glad to have lots of people to thank. Without further ado, let me open wide the floodgates of gratitude.

First, my committee members. Daniel Sutherland is just the sort of advisor I'd have wanted, giving me wise counsel, friendly encouragement, and motivational nudges when I needed them, but also tolerating my independent streak and trusting me to sort things out on my own. I hope this thesis does him proud. Mahrad Almotahari served as a co-advisor in lots of ways. He's been actively involved in the project from the start (which coincided with his arrival at UIC), and his keen mind and extensive knowledge have helped me out of many philosophical tough spots. Although we've yet to meet in person, Marc Lange has been a great inspiration and help. I doubt that many other best-in-their-field philosophers would correspond with an unknown grad student as generously as Marc has. Finally, I'm grateful to Kenny Easwaran for being a math hero and insightful interlocutor, and to Dave Hilbert for teaching me lots of great philosophy and science. It's been a privilege to work with two such preternaturally smart people.

Second, my family. Being the relative of an aspiring academic philosopher is weird in many ways, I'm sure, especially when that person suddenly decides to move halfway across the country in (seemingly endless) pursuit of a Ph.D. But my parents and siblings have been amazingly good sports about it, even though grad school has meant seeing less of each other than we'd like. Dad, Jen, Elijah, Laura, Andrew, J.J., Grandma, Mom and the Woomers, I'm awfully fond of you and I'm glad our family is a thing. Thanks.

Third, my friends. On the Buffalo-diaspora side, I'm grateful for my old and dear ones Anthony, Jane, Jordan, Stacey, Elliott, Mark, Jamie, Chris and Janelle. I don't know what sort of person I'd have been without you, but I'm sure that version of me is worse in every imaginable way. On the Chicago end, thanks to Kelly Lynch and Scafuri Bakery for keeping me well supplied with coffee, pizza and companionship, and for Claire, Eric and for board games, Meat Club and general good times.

My final year at UIC was spent as a Graduate Resident Fellow at the Institute for the Humanities, and I'm very glad for the experience. Linda Vavra, Susan Levine and the Institute made the Institute a wonderful place to work, and I'm thankful to my co-Fellows Tim Soriano, Lindsay Marshall and especially Lisa James for their company inside and outside the office. Finally, Janella Baxter and Robert English made the philosophy department a smarter, kinder and funnier place. Chicago hasn't been the same since they've gone, but I'm glad they've found green pastures elsewhere.

Lots of other people helped this dissertation get done and improved my Chicago philosophy life. These good folks include UIC philosophy department secretary Valerie Brown, current and former UIC faculty Nick Huggett, Sally Sedgwick, John Whipple, Colin Klein and Ramin Takloo-Bighash, current and former UIC grad students Kristin Thornburg, Emily Lacy, Areins Pelayo and Saja Parvizian, the great philosophers Liz Camp, Andy Arana and Chris Pincock, my Virtual Dissertation Group partners Adam Bales and Shay Logan (as well as VDG organizer Joshua Smart), my APA commentator Isaac Wilhelm, and the many excellent students I've had the pleasure of teaching at UIC, to name just a few. Academic life can be hard, but I'm glad it doesn't have to be solitary.

Finally, I want to thank Lauren Woomer, my partner, hero and best friend, who's given me more help, encouragement, inspiration, enlightenment, laughs, long walks and delicious breakfasts than I can count over the past twelve years. It isn't enough to say that I couldn't have done it without you. That I wouldn't have known what's worth doing is closer to the truth.

Table of Contents

Summary

A mathematical explanation occurs when some piece of mathematics—a theorem, a diagram, or a proof, say—explains some other piece. This dissertation consists of three papers on the nature of mathematical explanation (ME) and its relationship to issues in metaphysics, epistemology and philosophy of science.

In "Arithmetic, Set Theory, Reduction and Explanation", I connect ME to the phenomenon of intertheoretic reduction—a relationship much discussed in philosophy of science since Nagel's work in the 1960s, but not yet properly explored in the mathematical setting. My view is that some reductions in pure mathematics, such as the reduction of arithmetic to set theory, have no explanatory benefits. This contrasts with the situation in empirical science, where successful reductions seem always to give rise to explanations regarding the subject matter of the reduced theory.

In "Mathematical Explanation Beyond Explanatory Proof", I try to banish a troublesome ghost that's long haunted the theory of ME—namely the idea that mathematical explanations always involve proofs in some way. In particular, I argue that theorems are often explanatory in their own right, and that this is in fact the right way to understand an important and much-discussed case of ME (namely, Galois's explanation of the solvability of polynomial equations).

Finally, "Viewing-as, Frames and Mathematical Explanation" deals with the practice of viewing one mathematical object as another, and the explanatory insights this practice often yields. I propose that we understand viewing-as as a kind of thought employing cognitive frames. Finally, I argue that "viewing-as explanations" are properly characterized in cognitive rather than metaphysical terms, and hence that the popular ontic conception of explanation may need to be revised.

Introduction

A mathematical explanation occurs when something is explained by some bit of mathematics—a theorem, a diagram, or a proof, say. Many philosophers think there are mathematical explanations in science. These would be cases in which a piece of mathematics explains an empirical fact. Others say that there are mathematical explanations in pure mathematics, where one piece of math explains another. I'm inclined to think both claims are true. In the coming pages, however, I'll have lots to say about the second sort of mathematical explanation and not much to say about the first. So from this point on, "mathematical explanation" (or ME for short) will mean "mathematical explanation in pure mathematics".

This dissertation consists of three papers on various aspects of ME. Since the papers don't comprise a single, unified work, this won't be the sort of introduction that sketches an overall argumentative trajectory. Still, there's a story to be told about why these papers seemed worth writing at this moment, how they fit into a larger conversation, and what I hope they accomplish. I'll try to tell that story here. To do that, I first need to say a few things about the recent history of the theory of ME.

One can find remarks about ME, or something like it, in philosophers as old as Aristotle. Other important historical sources include the Euclid commentators of late antiquity, the mathematician-scientists of the Renaissance and seventeenth century, and the logicians and foundationalists of the nineteenth century. But ME's appearance as a chapter of analytic philosophy starts with Mark Steiner's 1978 paper "Mathematical Explanation" ([Steiner 1978a]).

That paper was, as far as I know, the first-ever scholarly work devoted exclusively to ME. Its main contribution was a theory of explanatory proof based on the idea of "characterizing properties". Although it's now generally agreed that Steiner's theory is unsatisfactory, the paper contained lots of interesting examples and ideas, many of which have been picked up by later authors. But it wasn't greeted with immediate enthusiasm. In fact, aside from a response that appeared some nine years later ([Resnik & Kushner 1987]), philosophers after Steiner were mostly silent about ME for the next two decades. Then, around the turn of the millennium, Paolo Mancosu published an energizing series of articles exploring ME's history, illustrating its role in contemporary mathematical practice, and championing it as a serious issue in philosophy of mathematics ([Mancosu 1999], [Mancosu 2000], [Mancosu 2001], [Hafner & Mancosu 2005], [Mancosu 2008]). Soon other authors were joining in, and the theory of ME was underway.

The period dominated by Mancosu lasted roughly from 2000 to 2010. I think it's fair to say that work on ME during this time had two main goals. The first was to show that ME is a phenomenon worth theorizing about—that is, to prove that mathematicians really do talk about explanation, that they take explanatory concerns seriously, and that philosophers interested in this sort of thing should take note. The second goal was to evaluate existing theories of explanation that might have something to say about ME. Steiner's account was the primary target, but articles also appeared on ME and theories of scientific explanation, including van Fraassen's pragmatic view ([Sandborg 1998]) and Kitcher's unification view ([Hafner & Mancosu 2005]).

This ground-clearing work was valuable, but its focus was narrow, and it rarely made much of an effort to appeal to a wider philosophical audience. Mancosu's papers in particular seemed to be addressed mainly to fellow historians of mathematics and philosophers of mathematical practice. Most of the early literature followed suit—it made occasional forays into philosophy of science (mostly to discuss van Fraassen, Kitcher and the like), but it forged few connections with relevant issues in mainstream metaphysics, epistemology, or even philosophy of mathematics. What's more, the early period saw little in the way of positive theory-building. Mancosu and other authors produced a number of impressive case studies, and they raised some provocative questions. But they rarely ventured detailed proposals about the nature of ME or its relationship to other phenomena.

This has begun to change in the last decade or so. Now that the fight to make ME a respectable subject is over, the field has seen more ambitious and varied projects emerge, and a growing number of papers on ME are appearing in major "generalist" philosophy journals. (2016 also saw the publication of the first book dealing extensively with ME, Marc Lange's *Because Without Cause*.) Works from this second period have dealt, for example, with notions like *coincidence* and *natural property* and their relationship to ME ([Lange 2010], [Lange 2015]), the idea that some mathematical explanations are grounded in metaphysical dependence relations ([Pincock 2015]), the nature of explanatory proof ([Lange 2014]), and the similarities between ME and other kinds of non-causal explanations ([Lange 2016]). (Incidentally, this list also makes it clear to whom the torch has passed as the leading theorist of ME.)

The 2010s have been a fruitful decade, but plenty of work remains to be done. The papers making up this dissertation are meant to carry on the project of broadening and deepening the theory of ME. They do so by clarifying what counts as a mathematical explanation, and by bringing ME into closer contact with traditional issues in philosophy of science, metaphysics, epistemology, cognitive theory, and mathematical history and practice.

In "Arithmetic, Set Theory, Reduction and Explanation", I connect ME to the phenomenon of intertheoretic reduction—a relationship much discussed in philosophy of science since Nagel's work in the 1960s, but not yet properly explored in the mathematical setting. My view is that some reductions in pure mathematics, such as the reduction of arithmetic to set theory, have no explanatory benefits. This contrasts with the situation in empirical science, where successful reductions seem always to give rise to explanations regarding the subject matter of the reduced theory. In "Mathematical Explanation Beyond Explanatory Proof", I try to banish an old Steinerian ghost that's long haunted the theory of ME—namely the idea that mathematical explanations always involve proofs in some way. In particular, I argue that theorems are often explanatory in their own right, and that this is in fact the right way to understand an important and much-discussed case of ME. Finally, "Viewing-as, Frames and Mathematical Explanation" deals with the practice of viewing one mathematical object as another, and the explanatory insights this practice often yields. I propose that we understand viewing-as as a type of frame-theoretic phenomenon. Finally, I argue that "viewing-as explanations" are properly characterized in cognitive rather than metaphysical terms.

The theory of ME has matured over the last couple decades into a serious subject backed by a respectable literature. Still, until recently at least, it's been possible to view ME as a curious footnote to scientific explanation, or a hobby for collectors of arcane mathematical trivia, rather than a topic of broad general interest. I hope this dissertation helps put such views to rest. ME is a phenomenon with the power to upend received wisdom, pose pregnant new questions and reshape traditional debates, and it deserves a place in the philosophical canon.

Chapter 1

Arithmetic, Set Theory, Reduction and Explanation

(Previously published in *Synthese*, "Arithmetic, Set Theory, Reduction and Explanation", DOI: 10.1007/s11229-017-1450-8, William D'Alessandro.

> [I]n the philosophy of science the notions of explanation and reduction have been extensively discussed, even in formal frameworks, but there exist few successful and exact applications of the notions to actual theories, and, furthermore, any two philosophers of science seem to think differently about the question of how the notions should be reconstructed. On the other hand, philosophers of mathematics and mathematicians have been successful in defining and applying various exact notions of reduction (or interpretation), but they have not seriously studied the questions of explanation and understanding. ([Rantala 1992], 47)

Rantala's observation, though now almost a quarter-century old, remains pretty much on target. In spite of the recent surge of interest in various aspects of mathematical explanation[1], philosophers have yet to turn their attention to the explanatory (and more broadly epistemological) dimensions of intertheoretic reduction in mathematics.

I hope the rest of the paper will show that this neglect is unwarranted. As philosophers of science have long known in the empirical context, understanding the link between reduction and explanation is a vital part of understanding theory succession, mathematical progress, and the nature and

[1] See for instance [Steiner 1978a], [Mancosu 2001], [Weber & Verhoeven 2002], [Hafner & Mancosu 2005], [Tappenden 2005], [Avigad 2008], [Lange 2009], [Baker 2010], [Lange 2010], [Lange 2014], [Pincock 2015].

purpose of foundations of mathematics. I aim to illustrate these points by examining a particularly important single case: the reduction of arithmetic to set theory. Although a number of authors have expressed views about the explanatory significance of this reduction, the issues involved haven't yet been weighed as carefully as they deserve to be. This paper sets out to do so, and to draw some considered conclusions.

My thesis is that the reduction of arithmetic to set theory is unexplanatory. After clarifying this claim and giving some positive reasons to believe it, I devote the larger part of the paper to showing that the most serious arguments to the contrary—due to Steinhart, Maddy, Kitcher, and Quine—are unsuccessful. Finally, I discuss some consequences of the view for philosophy of mathematics, philosophy of science and the theory of reduction.

1.1 Intertheoretic reduction and explanation in mathematics

We might start by asking what it means for one theory to be reducible to another (either in general, or in the mathematical context specifically). This is an important question, but I won't say very much about it here. For present purposes I take it as a datum that arithmetic is reducible to set theory; I'm more confident in the correctness of this claim than I am about the details of any particular theory of reducibility.

Nevertheless, I'm inclined to think that the concept of reduction is at least closely related to the model-theoretic notion of (*relative*) *interpretability*. Informally, a theory T_1 is interpretable in a theory T_2 if there's a theoremhood-preserving translation function from the language of T_1 to the language of T_2. (The rather unwieldy formal definition won't be necessary here; see e.g. [Feferman 1960], 49 or [Niebergall 2000], 30-31.)

Identifying reducibility with interpretability has several advantages. For one, interpretability is quite similar to the classic Nagel-style understanding of reduction in philosophy of science[2]—an approach that continues to find plenty of adherents. Note that for Nagel, however, "Reduction....

[2] See [Nagel 1961].

is the explanation of a theory or a set of experimental laws established in one area of inquiry, by a theory usually though not invariably formulated for some other domain" ([Nagel 1961], 338). I don't follow Nagel in taking explanatoriness to be partly definitional of reduction. But it seems likely that Nagel thought this only because he was a deductivist about explanation (in the manner of [Hempel & Oppenheim 1948])[3], and pretty much everyone now agrees that deductivism is false. So the disagreement doesn't trouble me much. (In any case, those who insist on viewing intertheoretic reduction as necessarily involving explanation can simply interpret me as denying that arithmetic is reducible to set theory.)

The interpretability criterion also gives the right results about central cases, including the case of Peano arithmetic and ZF set theory. Finally, one can find in [Niebergall 2000] a sophisticated defense of relative interpretability as superior to other proposed criteria for reduction.

In any case, nothing much about the issues to be discussed below hangs on details about the nature of reducibility. I assume only the uncontroversial premise that arithmetic is reducible to set theory in some interesting sense.

Another point worth addressing: if, as I've claimed, intertheoretic reduction isn't by definition an explanatory relation, why think the two notions are connected at all? To start with, note that the claim in question is quite weak. It leaves open the possibility, for instance, that all cases of reduction are *actually* explanatory, as a matter of empirical fact. (Or, alternatively, that this is true of all sufficiently natural reductions, or of some other relatively large and interesting class of examples.) And indeed, the evidence from philosophy of science strongly suggests the existence of a close connection of some such kind. Consider the (purported) reductions of classical thermodynamics to statistical mechanics, of Mendelian inheritance theory to biochemistry, of Kepler's theory of planetary motion to Newtonian mechanics, and so on.[4] In each such case, the viewpoint associated with the reducing theory yields substantial explanatory insights about the subject matter of the reduced theory. These insights can take the form of clarifications of the nature and properties of the

[3]See [van Riel 2011] for a discussion of Nagel's deductivism and its influence on his account of reduction.

[4]Whether these cases count as bona fide intertheoretic reductions has been the subject of debate. I don't want or need to take a stand on this issue here—letting go of some (or even many) individual examples is no problem for my purposes, as long as there are some genuine instances of reduction that look something like the ones philosophers have discussed.

entities postulated by the reduced theory (e.g., the identification of genes with regions of DNA), of the reasons why the laws of the reduced theory hold and the circumstances under which they do (e.g., the explanation of the Second Law of Thermodynamics in terms of the improbability of entropy-lowering microstate evolutions), of the place of the reduced theory in a larger conceptual or nomic framework, and so on.

Examples from empirical science thus suggest that explanatory reductions are the norm, if not in fact the rule. A priori, there seems to be no obvious reason to expect this pattern to fail in the mathematical setting. It would be an interesting result, then, if a central and uncontroversial case of intertheoretic reduction in mathematics turned out to lack explanatory value. The next section argues that the reduction of arithmetic to set theory is such a case.

1.2 Arithmetic and set theory

1.2.1 The view to be defended

I claim that the reduction of arithmetic to set theory is unexplanatory.[5] I'll have much to say in defense of this claim below, but it's worth noting immediately that the thesis is highly plausible as a claim about other bits of mathematics, too. Perhaps the best example is the set-theoretic reduction of the ordered pair. Recall that, according to the now-standard definition (due to Kuratowski), the ordered pair (x,y) is identified with the set $\{\{x\},\{x,y\}\}$. It's easy to see that the set in question possesses "the characteristic property of ordered pairs", namely the feature that $(x_1,y_1) = (x_2,y_2)$ just in case $x_1 = y_1$ and $x_2 = y_2$. So the reduction of the ordered pair is unobjectionable from a logical point of view. Nevertheless, it seems clear that the reduction gives us no new insight or understanding about the nature of order. Rather than being an explanatory achievement, we want to say, the identification $(x,y) = \{\{x\},\{x,y\}\}$ is a stipulation designed to solve a technical problem—a clever and convenient stipulation that does its job well, granted, but not the sort of

[5]For the sake of variety or readability, I'll sometimes use expressions like "the reduction of numbers to sets" below. This is just loose talk. Unless the context indicates otherwise, such expressions always mean the same as "the reduction of arithmetic to set theory"—I don't mean to invoke some alternative, ontological notion of reduction by speaking of objects instead of theories.

thing that leaves us epistemically better off than before. As Michael Potter's set theory textbook puts it,

> $\{\{x\},\{x,y\}\}$ is a single set that codes the identities of the two objects x and y, and it is for that purpose that we use it; as long as we do not confuse it with the genuine ordered pair (if such there is), no harm is done. In other words, the ordered pair as it is used here is to be thought of only as a technical tool to be used within the theory of sets and not as genuinely explanatory of whatever prior concept of ordered pair we may have had. ([Potter 2004], 65)

Randall Dipert comments similarly that

> the defining quality of an ordered pair is quite clear without a set-theoretic formulation. Namely, two ordered pairs are identical just when their respective members are identical. It *was* quite clear before the advent of sets and set-talk. ...In terms of increased precision and increased clarity, the set-theoretic reduction of ordered pair added nothing, historically and philosophically. ([Dipert 1982], 366-7. Emphasis in original.)

Analogous remarks may apply to other familiar set-theoretic reductions. Take functions, for example. In set theory a function f is identified with the set of ordered pairs $\{(a_1,b_1),(a_2,b_2),\ldots\}$, where a_i and b_i are the elements of the domain and image of f, respectively, and where $f(a_i)=b_i$. Again, nobody disputes the usefulness of this maneuver. But its usefulness arguably lies in its adequacy as a technical device rather than its power to explain anything about functions.[6] The set theorist Yiannis Moschovakis notes in this vein that

> [t]his "identification" of a function [with a set of ordered pairs] has generated some controversy, because we have natural "operational" intuitions about functions and by "function" we often mean a formula or a rule of computation. ...There is no problem with this if we keep clear in our minds that the "definition" [just mentioned] does not replace the intuitive notion of function but only represents it within *set theory*, faithfully for the uses to which we put this notion *within set theory.* ([Moschovakis 2006], 40. Emphasis in original.)

[6]The case of functions is more complicated and controversial than that of the ordered pair. Identifying a function with its set-theoretic graph, for instance, allows one to prove nontrivial theorems that weren't available to earlier mathematicians operating with less clear or rigorous conceptions. (E.g., the theorem that there exist uncomputable functions.) The thesis that the reduction of functions is unexplanatory is probably most plausible when restricted to some class of "classical" functions, for instance differentiable or continuous functions of a real or complex variable.

What's going on in such cases seems to be roughly this. At the beginning of the story, before the question of reduction enters the picture, we have some mathematical concept C with which we're epistemically satisfied in important respects—that is, a concept whose conditions of application aren't in dispute (at least in ordinary cases), and whose basic properties we feel we understand more or less fully.[7] We become convinced at some point of the desirability of set-theoretic foundations (either for mathematics as a whole, or for the particular subject matter to which C belongs). The goal is then to find a set-theoretic construction that can stand in for objects of type C. This is done by identifying the characteristic logical properties of C-type objects and then producing a set (or collection of sets) with the desired properties. If there are many different candidates, then judgments about relative simplicity, convenience, and so on are used to narrow the field down to a single structure S (or at most a small handful of such structures). Given a sufficiently robust consensus, S then comes to be viewed as the canonical representation of C-type objects within set theory. Whatever its virtues *qua* set-theoretic surrogate, though, the purpose of S isn't to explain anything or generate new understanding about C-type objects. Rather, S does its job by hewing faithfully to the defining C-properties. (And it's precisely because we already had an adequate epistemic handle on these properties that we were able to specify them in advance, and to use them to guide our choice of S.)

This, I think, is pretty clearly the correct story about the reduction of ordered pairs and functions to sets. I maintain that it's also the correct story about the reduction of arithmetic. Viewing the natural numbers and arithmetical operations as set-theoretic constructs is a useful expedient for certain purposes, but doing so has no particular explanatory value.

Before saying more, we might as well recall how the reduction is supposed to go. As is well known, there are many set-theoretic structures that model the Peano axioms, and hence many ways to reduce arithmetic to set theory. Historically, the systems of Zermelo and von Neumann have been the most influential. Of the two, the von Neumann approach is usually thought to have a preponderance of nice properties, and hence has become the default choice of most mathematicians.

[7]This isn't to suggest that the correct understanding of the concept in question need have been easy to determine. The modern notion of function, for instance, evolved slowly and rather tortuously from its roots in 17th-century geometry and analysis to its current general form. For a brief overview of this history, see [Kleiner 1989].

A model of Peano arithmetic is a structure $(S, f, e, \prec)$, where S is a set (interpreted as the collection of *natural numbers*), f is a unary function (interpreted as *successor*), e is a constant (interpreted as *zero*), and $\prec$ is a binary relation (interpreted as *less than*). Hence the first step in specifying a reduction of arithmetic to set theory is to choose a collection of sets S to serve as the natural numbers. In the von Neumann system, these sets are the finite von Neumann ordinals, and the correspondence is

$$\begin{aligned} 0 &= \emptyset, \\ 1 &= \{\emptyset\}, \\ 2 &= \{\emptyset, \{\emptyset\}\}, \\ 3 &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \end{aligned}$$

and in general $n+1 = n \cup \{n\}$. This also supplies the definitions of the successor function and the zero constant. For less than, $\prec$ is identified with the proper subset relation $\subset$. It isn't hard to show that the finite von Neumann ordinals, equipped with these definitions, satisfy the Peano axioms. (The Zermelo system uses the correspondence $0 = \emptyset$, $n+1 = \{n\}$, and identifies $\prec$ with the ancestral of set membership. Zermelo's approach is rarely used today, though, so I'll focus on the von Neumann ordinals from now on.)

My view, then, is that correspondences of this sort between arithmetic and set theory are bona fide intertheoretic reductions which are nevertheless unexplanatory. By this I mean that they don't explain (or substantially contribute to explaining) anything about the natural numbers and their arithmetic.[8] Put another way, there's no explanatory benefit to *viewing numbers as sets* or *supposing that numbers are sets*—in the way that there is such a benefit to viewing a sample of radium as an assemblage of Ra atoms, say, or temperature as mean molecular kinetic energy. In a moment

[8] Anything except, perhaps, certain kinds of possible explanations that are too trivial, insubstantial, and *pro forma* to be worth caring about. For instance, if one subscribes to the view that facts of the form $\exists x Fx$ are explained by their instances, then one might think that the reduction of arithmetic to set theory explains why there exists a theory to which arithmetic is reducible. Maybe so. But my interest is in explanations that make a substantive contribution to mathematical understanding, and such cases clearly don't fit the bill.

I'll say why I think so. For now, though, let me make a clarification and fend off some likely misunderstandings about the view.

First the clarification. I've been making claims about the explanatoriness of various bits of science and mathematics, but I haven't yet said exactly what I take such claims to mean. Just what notion of explanation is at stake here? To some extent I'd like to be noncommittal about this issue. Experience suggests that we—that is, mathematicians, philosophers of mathematics and other knowledgeable observers—are usually good at agreeing about what counts as explanatory in specific cases, once all the relevant facts are on the table. So I take considered and informed judgments of this sort, including my own, to be generally reliable. On the other hand, there's very little agreement about which theory of explanation is correct (including, and perhaps especially, in the mathematical setting), and so it seems best not to rely overmuch on any such theory.

Nevertheless, I think one can safely say at least a few things about the properties of (mathematical) explanations. To start with, there are at least two kinds of phenomenon that arguably warrant the name. One familiar type is an *answer to a why-question*.[9] Another type might be termed *rendering intelligible* (or *elucidating* or *clarifying*). Whereas answers to why-questions explain facts, elucidations explain things: they illuminate the nature of an object or concept, where our previous understanding was obscure, incomplete, inconsistent or otherwise problematic.[10] The discussion below will involve claims about both types of explanation. (For instance, section §2.4 considers Maddy's claim that the reductionist viewpoint explains why multiplication is commutative, while §2.6 deals with Quine's remarks about set-theoretic "explications" of arithmetical notions.)

What's more, explanations stand in a characteristic pattern of relations to other epistemic and cognitive goods. Possessing an explanation typically imparts understanding, for example, and often improves learning, inference and problem-solving in the relevant domain.[11] Hence it's reasonable

[9] As philosophers have taken pains to show over the past couple decades, this type of explanation is encountered in pure mathematics no less than in the empirical sciences; working mathematicians often ask why theorems are true, and they often go to lengths in search of answers. (See the references in footnote 2 above.) To take a random recent example, [Tao 2015] proposes a large-scale collaborative effort to find an explanation for certain surprising polynomial identities.

[10] Some philosophers may want to deny that this second phenomenon is a genuine kind of explanation. If they were right, then my main thesis would be that much easier to defend, since there would then be fewer ways for the reduction of arithmetic to set theory to be prospectively explanatory.

[11] See, e.g., [Lombrozo 2006] and [Lombrozo 2012] for references and discussion.

to take the absence of epistemic and cognitive benefits to reliably indicate a lack of explanatoriness. Finally, we know what sorts of things can serve as mathematical explananda and explanantia. Theorems often stand in need of explanation, as do (the natures of) kinds of mathematical entities. The former can be explained by appropriately enlightening proofs, or by other theorems or bodies of theory. The latter can be explained by suitably clear definitions or identifications.

I hope the above gives some reassurance that the question I've posed makes sense, and that we have some resources for trying to answer it. If this isn't yet entirely clear, perhaps the details that emerge in the rest of the paper will help make it so.

I now want to mention and try to dissolve a few possible confusions about what my view entails. First, in denying that the reduction of arithmetic to set theory is explanatory, I don't mean to presuppose that it *ought* to have been explanatory. So I don't claim that the lack of explanatory power means that the reduction lacks value, or that it failed to achieve its intended purpose. Although some people have demanded that set-theoretic foundations be explanatory—as I'll discuss below—my view is that such reductions can be (and are) successful without contributing to explanation or understanding. (See §2.7 for more on these issues.)

Second, it's no part of my view that set theory contributes nothing of explanatory value to *mathematics generally*. This claim is pretty clearly wrong. To take just one example, there were important explanatory issues at stake in the questions about infinite sets of real numbers that originally motivated Cantorian set theory. Consider in this regard Joseph Dauben's comments on Cantor's early work:

> [In 1874 Cantor] had established that $\mathbb{R}$ was nondenumerable. Following this remarkable discovery, certain new avenues of inquiry must have appeared simultaneously. If, in terms of cardinality, there were more real numbers than natural numbers, were there further quantitative distinctions to be made between and beyond them? How did the new discovery contribute to the explanation and understanding of continuity? ...As Cantor raised the question of mapping lines and planes, such questions seem to have been uppermost in his mind. ([Dauben 1979], 58-59)

Not only did Cantor puzzle over the correct explanation of continuity, infinite size, and other notions, his work also yielded illuminating answers to many such questions. The thesis I defend about the reduction of arithmetic isn't intended as a denial of set theory's explanatory achievements in these and other parts of mathematics.

Finally, and perhaps most importantly, my view should also be distinguished from the claim that set theory fails to explain anything about the natural numbers in particular. This last claim sounds much like the one I want to defend. But the two aren't equivalent, and failing to properly distinguish between them is likely to cause confusion and prompt misplaced objections. In fact the view just mentioned is substantially stronger than my own view. Moreover, I think this stronger view is probably false. Arguably the notion of set is needed to properly characterize the natural numbers, since the (second-order) axiom of induction quantifies over sets of numbers.[12] And the fact that the collection of natural numbers has the smallest possible infinite cardinality is an interesting result that's plausibly explained by set theory. But these explanations in no way involve *viewing numbers as sets*. (Rather, they involve taking the collection of natural numbers and its subcollections as sets, which is something else altogether. Numbers need not *be* sets to occur as *elements* of sets.) So it's important to note that cases involving this sort of explanatory relationship pose no threat to my view.[13]

The thesis I'm defending, then, is more modest than the above claims. The view, to reiterate, is just that *the reduction of arithmetic to set theory*—not set theory itself, or the relationship between sets and numbers on the whole—is unexplanatory. There are both positive and negative reasons to

[12]Thanks to both Kenny Easwaran and an anonymous referee for *Synthese* for suggesting this example.

[13]An anonymous referee suggests that results like Goodstein's theorem—which are unprovable in first-order Peano arithmetic, but provable in the second-order setting using set-theoretic methods—might count as set-theoretic explanations of arithmetical facts, and hence might seem problematic for my view. I'm inclined to group this sort of case, together with the others mentioned in this paragraph, under the heading of "ways of applying set theory to arithmetic that may be explanatory, but which don't depend on viewing numbers as sets". So I don't find such cases worrying.

To see why, it's helpful to briefly describe the set-theoretic proof of Goodstein's theorem. One starts by considering an arbitrary Goodstein sequence $G(m)$, which is a certain sequence of natural numbers. One wants to show that $G(m)$ eventually terminates, i.e. that it takes the value 0 at some point. (Goodstein's theorem is the statement that all Goodstein sequences terminate.) To show this, one constructs a sequence of ordinal numbers $O(m)$ with the properties that (1) $G(m)$ can be shown to terminate if $O(m)$ terminates, and (2) $O(m)$ does in fact terminate. This is a sparse sketch of the proof, of course, but hopefully it's clear from the sketch that the proof doesn't exploit the set-theoretic representation of the natural numbers in any way. One could view the numbers as any sort of object whatsoever and the proof would still go through.

In any case, I think it's far from clear that the set-theoretic proof of Goodstein's theorem is explanatory in the first place. The fact that a certain sequence of ordinals terminates, it seems to me, surely isn't *the reason why* the associated Goodstein sequence terminates. The behavior of the two sequences is correlated, but what happens with the sequence $O(m)$ can hardly be said to "ground" or "determine" what happens with the sequence $G(m)$. So I'm doubtful that we're even dealing with a case of mathematical explanation here.

think so. The positive reasons are considerations that directly cast doubt on the explanatoriness of the reduction. The negative reasons consist in the lack of evidence for the contrary claim (that the reduction is explanatory after all). As is often the case when one tries to establish a nonexistence claim, the issues surrounding the negative reasons are arguably more interesting and weighty than those surrounding the positive reasons—after all, once we've shown that the alleged evidence for the existence claim is unconvincing and that the claim isn't needed to account for the phenomena, considerations of parsimony give us good reason to reject it. So most of the work below will focus on answering various arguments purporting to show that the reduction of arithmetic to set theory is explanatory.

1.2.2 Some positive considerations

First, however, the positive considerations. One serious reason for doubt, alluded to above, is the fact that certain other set-theoretic reductions—e.g. that of the ordered pair—are pretty clearly unexplanatory. If it turns out that the reduction of arithmetic is similar to these cases in important respects, then we have good inductive grounds to expect that the reduction of arithmetic isn't explanatory either.

Are the two kinds of case similar in ways that matter? I think so. In particular, the account sketched above about the translation of an antecedently well-understood concept C into set-theoretic language seems to fit the case of arithmetic nearly as well as it fits the case of the ordered pair. In both cases, a specification of the logical properties of the structure preceded the attempt to model the the structure in set theory. Also in both cases, several such surrogates were proposed[14]—each of which satisfied the relevant postulates, but which differed along various other dimensions—before the mathematical community eventually settled on a favorite. And the consensus in each instance

[14]The 1921 Kuratowski definition of the ordered pair was preceded by attempts by Wiener (1914) and Hausdorff (1914), who identified (a,b) with $\{\{\{a\},\emptyset\},\{\{b\}\}\}$ and $\{\{a,1\},\{b,2\}\}$, respectively. (The 1 and 2 in Hausdorff's definition are arbitrarily chosen objects distinct from a, b, and each other.)

was based on practical considerations of simplicity, well-behavedness and so on, rather than epistemic claims about explanatoriness or understanding.

In the case of the ordered pair, this process yielded a construction that pretty obviously has no claim to explanatory value. Since the (von Neumann) reduction of arithmetic to set theory was the result of much the same type of process, it's reasonable to expect much the same outcome.

Another reason for doubt is based on number theorists' own attitudes and behavior. It's evident that mathematicians care a great deal about obtaining understanding, and that this value substantially influences mathematical practice and pedagogy.[15] It's also a commonplace that explanation and understanding are closely related[16], insofar as good explanations typically produce better understanding of their *explananda* (and the associated subject matter). Thus, if viewing numbers as sets had appreciable explanatory value, we should expect number theorists to adopt this vantage point for the sake of exploiting its epistemic and cognitive resources.

But this isn't in fact what happens. Although number theorists draw extensively on other areas of mathematics to gain insight and solve problems—think of Wiles's use of algebraic geometry in the proof of Fermat's Last Theorem, Hrushovski's model-theoretic approach to the Mordell-Lang conjecture, the application of Fourier-type methods in studying prime distributions, and so on—one doesn't find the identification of numbers with sets being deployed for such purposes.[17] Viewing the natural numbers as, say, the von Neumann ordinals has led to no breakthroughs on longstanding problems, insightful new ways of organizing and teaching number theory, or the like. The fact that number-theoretic practice has so little use for the set-theoretic viewpoint suggests that the latter doesn't contribute significantly to understanding. Hence the reduction of arithmetic to set theory is probably not explanatory.

I turn next to negative considerations. In the parts to follow, I'll consider four arguments purporting to show that the reduction of arithmetic to set theory is explanatory after all. The first, due to

[15] See for instance [Avigad 2008], [Carter 2008], [Lipton 2011], [Sierpinska 1994], [Tappenden 2005].

[16] It's sometimes argued (e.g. by [Khalifa 2012]) that understanding is nothing other than having an explanation. At other times (e.g. by [Strevens 2013]) the two states are treated as distinct, but explanation is viewed a necessary component or precondition of understanding.

[17] As I pointed out before, care should be taken not to confuse this point with the claim that set theory *in general* has nothing of explanatory value to contribute to number theory. I take the latter claim to be false, or at least highly dubious.

Zermelo and his followers, centers on the ontological thesis that numbers just *are* sets. The second, due to Maddy, claims that there are explanatory set-theoretic proofs of some arithmetical facts. The third, due to Kitcher, suggests that the reduction of arithmetic to set theory sheds light on some high-level features of the natural number system. The last, due to Quine, maintains that the reductionist viewpoint renders the concept of number more intelligible. These arguments raise many interesting issues, but none, I think, are ultimately successful. Hence we've been given no good reason to think that the reduction of arithmetic to set theory is explanatory.

1.2.3 A reductionist argument

Reductionism about objects of a given type, say A, is the view that there exist objects of some notionally different type B such that the As are in fact identical to some of the Bs. *Reductionism* about As is thus a stronger thesis than the claim that the A-theory is *reducible* to the B-theory, since reducibility (on the Nagel-style view that I endorse) is a formal rather than metaphysical affair.[18] It's often tempting to make the leap from reducibility to reductionism—if the A-theoretic truths correspond systematically to B-theoretic truths (with nothing left over), why not try to keep ontological costs down by doing away with the As entirely? This maneuver has made plenty of appearances in philosophy of science, philosophy of mind and elsewhere. It's not too surprising, then, that some mathematicians and philosophers have embraced *set-theoretic reductionism* of one sort or other.

The version of reductionism that's relevant for present purposes is the claim that the natural numbers in particular are nothing but sets. (This thesis is arguably traceable to, and most famously associated with, the early work of Ernst Zermelo. See [Zermelo 1909a] and [Zermelo 1909b] for statements of Zermelo's reductionism, and [Taylor 1993] and [Hallett 1984] for analysis and commentary.) If reductionism about the natural numbers is true, then the line I'm defending about set theory and arithmetic is in trouble. After all, if numbers are identical to sets, it's hard to see how the reduction of arithmetic to set theory could fail to be explanatory. Ascertaining what numbers really are surely counts as explaining something important about them, to start with. And presumably

[18] See [Klein 2009] for an extended (and convincing) defense of this claim.

this identification will have all sorts of explanatorily valuable further consequences. Compare the identification of water with H_2O, of genes with regions of DNA, and so on. Such identifications are always explanatory achievements in their own right, it seems, and they invariably give rise to explanations involving other properties of the reduced entity—why salt but not sand is soluble in water, or how allelic dominance and recession are implemented at the molecular level, for instance. So it should go with numbers and sets, if reductionism is true.

The most influential argument against reductionism is due, of course, to Paul Benacerraf ([Benacerraf 1965]). Benacerraf's reasoning is roughly as follows. There are several alternative reductions of arithmetic to set theory, each giving a different correspondence between numbers and sets. If numbers are identical to sets, as the reductionist claims, then any particular number must be determinately iden-

[19] Of course, Benacerraf's argument has generated a great deal of discussion over the years, and some post-Benacerrafian philosophers have continued to hold views that ascribe some special metaphysical status to sets vis-à-vis the natural numbers. One might think that a proper examination of the reductionism issue would include some mention of these views. In fact, though, the views in question—or at least the ones I'm aware of—uniformly concede Benacerraf's point that numbers can't be uniquely identified with any particular sets in a principled way. Hence they're not directly relevant to the line of thought I take up here. Some noteworthy examples of what I have in mind are the views of Penelope Maddy, W.V. Quine and Nicholas White.

Maddy's early work (e.g. [Maddy 1990]) argues that numbers are (equinumerosity) *properties* of sets. This is indeed reductionism of a certain sort. But Maddy explicitly says that, for Benacerrafian reasons, there's no hope of identifying numbers with sets themselves. (Whether numbers can or should be identified with properties of sets, and whether such an identification would have explanatory value, is an interesting question. But it's not quite the question this paper is trying to answer.)

Quine (e.g. [Quine 1960]) held a view that can be described as claiming that "numbers are sets", which sounds anti-Benacerrafian. But his version of reductionism basically amounts to the thesis that it's convenient to identify numbers with sets, together with a pragmatic approach to ontology. On Quine's view, we're free to make any identification of numbers with sets that serves our purposes; if there are several equally handy choices, then each identification counts as equally correct, as long as we stick with the one we've chosen. So there's no notion here of getting at a deep truth about what numbers "really were all along". In particular, Quine agrees with Benacerraf that there's nothing metaphysically special about either the Zermelo ordinals or the von Neumann ordinals. (Nevertheless, one might think that Quine's view still imputes a type of explanatory value to set-theoretic reductions, even if it isn't for the metaphysical reasons discussed in this section. See §2.6 below for more on this issue.)

A less well-known but notable reaction to Benacerraf is [White 1974]. White agrees with the anti-uniqueness part of Benacerraf's argument, but from here he takes the unusual line that "the existence of multiple set-theoretic models of arithmetic should prompt us, not to say with Benacerraf that numbers cannot be sets, but rather to suggest that there are multiple full-blown series of natural numbers. Thus, for example, instead of there being only one three, there are after all many threes, and many thirty-sevens, and so on" (112). White's view sounds at first like a version of reductionism, but it later becomes clear that this isn't what he has in mind. His view is rather that objects of *any* kind count as numbers, insofar as they can be placed in an $\mathbb{N}$-like progression. So White doesn't, after all, identify numbers with set-theoretic finite ordinals in particular (although these are *among* the objects that count as numbers for him). As he points out, the view is better described as a sort of Pythagoreanism than a type of set-theoretic reductionism.

Finally, it's worth mentioning that another prominent claim from "What Numbers Could Not Be" is Benacerraf's thesis that numbers aren't objects *of any sort at all*. This is a claim with which many people *have* directly disagreed. I find Benacerraf's argument for this view unconvincing myself, but its truth or falsity doesn't directly bear on anything I say below, so there's no need to canvass responses to it here.

tical to some particular set. So at most one of the alternative reductions gives the right identification of numbers with sets. Suppose that *R* is this reduction. Then *R* must have some special feature in virtue of which it stands out as uniquely correct. But, in fact, none of the alternative reductions has any such special feature. (There seems to be nothing outstanding in the relevant way about the von Neumann ordinals $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\}, \ldots$ as compared to the Zermelo ordinals $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\{\emptyset\}\}, \ldots$, for instance.) Therefore reductionism is false.

The broad consensus in philosophy of mathematics has been that Benacerraf was right, and hence that reductionism is untenable. Even philosophers who are sympathetic to some version of mathematical realism tend to think that numbers are obviously not sets: the platonist Mark Balaguer, for instance, writes that "it is more or less beyond doubt that no sequence of sets stands out as the sequence of natural numbers" ([Balaguer 1998], 64). Still, there are occasional voices of dissent. The most extensive recent defense of reductionism is [Steinhart 2002], which argues that the natural numbers are identical to the finite von Neumann ordinals.[19]

Steinhart's argument has two parts. The first part consists of an enumeration of some nice properties of the finite von Neumann ordinals (FVNOs) as compared to the Zermelo ordinals and other candidates. (The list includes the naturalness of the extension of the FVNOs to the transfinite ordinals, the elegance of identifying the numerical ordering $<$ with set membership, and so on. Other candidate reductions also have some of these features, but the FVNOs arguably boast the most impressive assortment.) As Steinhart points out, mathematicians standardly represent the natural numbers with the FVNOs on account of their having these convenient features. This fact, Steinhart thinks, amounts to "an argument from mathematical practice" that the natural numbers are the FVNOs.

There are at least two ways to understand the argument in question. One possible claim is that mathematicians (or set theorists, at any rate) are themselves committed to the view that the FVNOs are the natural numbers—and so, since we should defer to epistemic authorities about questions within their domain of expertise, we should follow them in being so committed. A second possible claim is that, irrespective of what mathematicians actually happen to believe, their reasons for preferring the FVNOs are in fact compelling grounds to identify the FVNOs with the natural numbers.

Neither version of the argument is very convincing. Starting with the first version, it's doubtful that mathematicians (or set theorists in particular) are generally committed to identifying the natural numbers with the FVNOs. For one, some prominent set theorists are quite clear about rejecting this sort of reductionism. For instance, as we saw above, Moschovakis takes pains to warn his readers against supposing that functions and other familiar objects are identical to the sets used to represent them. Even setting this point aside, however, there's no good reason to think that mathematicians' preference for the FVNOs should be understood as an ontological prescription rather than a practical one. Even if some set theorists have suggested that "the natural numbers are (or ought to be viewed as) the FVNOs", the context in which such claims are made is that of ascertaining which practices are most amenable to doing set theory, not that of answering philosophical questions about identity.

As for the second version, suppose we grant for the sake of argument that the FVNOs edge out other candidates in the possession of nice properties. Does any interesting metaphysical conclusion follow? It's hard to see how it could. It's not clear what sort of plausible principle would let us infer "the Xs are the Ys" from "representing the Xs as the Ys within set theory is moderately more convenient than doing otherwise". (Analogously: there are many possible ways of representing the natural numbers by arrays of dots. Some such representations are more simple, elegant and practical than some others; perhaps there's some reasonable way of assigning each one an overall niceness score, and perhaps if we did so there'd be a convincing overall winner. In itself, though, none of this seems seems to convey any information about whether numbers really are arrays of dots, or which arrays of dots they'd be if they were.) In particular, if this line of reasoning is to be understood as an inference to the best explanation, it's unclear how to make it work: the proffered *explanans* hardly seems to count as any sort of explanation of the *explanandum*, let alone as a uniquely good such explanation.

In any case, Steinhart himself concedes that the "argument from mathematical practice" is inconclusive. But he also offers a second, and supposedly decisive, argument for reductionism. (Indeed, he refers to the second argument as "a precise mathematical demonstration that the natural numbers are the finite von Neumann ordinals" ([Steinhart 2002], 355).[20] The reasoning is as follows. To start with, Steinhart borrows from Benacerraf two sets of conditions that a structure S must satisfy in order to serve as the set of natural numbers. The first set of conditions is the Peano axioms for arithmetic. The second set consists of two "cardinality conditions":

1. For all $n \in S$—i.e., for all putative natural numbers n—there exists a set $n^* = \{m \in S : m < n\}$ of all numbers less than n.

2. The cardinality of any set A is n if and only if there's a bijection between A and n^*.

Steinhart argues that, *pace* Benacerraf and his followers, accepting these two sets of conditions ("the NN-conditions") commits us to identifying the natural numbers with the FVNOs.

Steinhart's argumentative maneuvers aren't always easy to follow, but the general shape of the reasoning seems to be this. First, suppose that reductionism is true, so that the natural numbers and arithmetical operations are uniquely identifiable with a particular set-theoretic structure $(S, f, e, \prec)$, or S for short. (As above, S is a set, f is a one-place function to be interpreted as *successor*, e is an element of S to be interpreted as 0, and $\prec$ is a two-place relation to be interpreted as *less than*.) Then, whatever structure S might be, the cardinality conditions above "[compel] us to form certain definite sets" of the elements of S, namely the sets n^* for each $n \in S$ (351). With appropriately chosen definitions for zero, successor, and less than, however, the sets n^* themselves constitute a structure S^* that satisfies the Peano axioms and the cardinality conditions. This structure S^* is in fact nothing other than the FVNOs. By uniqueness, therefore, $S = S^*$, and so the natural numbers are uniquely identifiable with the FVNOs.

Perhaps the most important problem with Steinhart's argument is that it's not clear why the structure S^* is the only candidate we're permitted to consider in determining the identity of S (and hence that of the natural numbers). That is, it's unclear why the following reasoning is apparently not

[20]Further references to Steinhart are to this paper.

allowed: "Suppose that the natural numbers and arithmetical operations are uniquely identifiable with a particular structure S. As is well known, the finite Zermelo ordinals with the usual operations constitute a structure Z that satisfies the Peano axioms and the cardinality conditions. By uniqueness, therefore, $S = Z$, and so the natural numbers are uniquely identifiable with the finite Zermelo ordinals."

Steinhart's answer to this challenge seems to involve the idea that "the cardinality [condition] compels us to form certain definite *sets*... Specifically: we have to form, for each [$n \in S$], the set of all numbers less than n. For every number n, we must form the set [$n^* = \{m \in S : m \prec n\}$]. ...We cannot avoid forming all these sets of numbers" (351). Later, he adds that the conditions mentioned above "*do not* allow you to form any other sets of [numbers from S]. They do not assert or imply that the NN-universe contains any other sets built from [such] numbers" (353). ("The NN-universe" is defined to be "the universe of objects to which we may ontologically reduce the natural numbers" (352).)

This is rather puzzling. Steinhart's thought seems to be that we can only consider S and S^* (and not any other structures to which we might reduce the natural numbers) because the existence of S^*, and only that of S^*, is entailed by the existence of S together with the cardinality conditions. But why should this matter? Given that we accept ZFC (or some other system of set theory), we have a universe's worth of sets already at our disposal. It makes little sense to forgo the use of these other sets, or to pretend they don't exist, for the reasons Steinhart gives. Whether or not we can "avoid forming" other sets on the basis of the existence of S and the cardinality conditions seems hardly relevant, and perhaps isn't even a coherent question—the other sets are in some sense already there, and hence not in need of being "formed" by us at all. Thus, as long as we're committed on good independent grounds to the existence of the Zermelo ordinals (say), they ought to be fair game in our deliberations about the identity of the natural numbers. And I take it we are so committed. Hence I fail to see how Steinhart can block the $S = Z$ argument sketched above in a motivated way. (Of course, if he can't do so, then nothing prevents Benacerraf's reductio from arising again: in that case we'll have $S = Z$ and also $S =$the FVNOs, and hence we'll have to reject the assumption that the numbers are uniquely identifiable with a particular series of sets.)

In light of these issues, I conclude that Steinhart's defense of reductionism is unsuccessful. There remains no compelling reason to identify the natural numbers with any particular set-theoretic structure. Hence we needn't think that the reduction of arithmetic to set theory is explanatory in virtue of its revealing the true nature of the numbers.

1.2.4 A Maddian argument

Here's another worry. It might be that general, a priori reflection about set theory and arithmetic is liable to paint an incomplete picture of the relationship between the two. Set-theoretic language and reasoning have a distinctly different flavor from their number-theoretic counterparts, after all, and it's not easy to imagine how the theorems and proofs of arithmetic will look when viewed through the lens of set theory. If we pay close attention to details, the reduction of numbers to sets might yet yield an unexpectedly insightful proof or piece of reasoning.

What might a case of this sort look like? Here's a suggestion from an early paper of Penelope Maddy:

> [L]et me give an elementary example of what I take to be a demonstration of the explanatory power of set theory. Suppose you wonder why multiplication is commutative. You could prove this by induction from the Peano postulates by showing:
>
> (1) $n \cdot 0 = 0 \cdot n$
>
> (2) if $n \cdot m = m \cdot n$, then $(n+1)m = m(n+1)$
>
> ...If you were convinced of the truth of the Peano postulates and of the soundness theorem, this exercise should convince you of the commutativity of multiplication, but you were convinced of this in the first place; your question was why, and the proof did little towards answering it. Now suppose you take a set theoretic perspective and again ask why multiplication is commutative. Here an answer is forthcoming: because if A and B are sets, then there is a one-to-one correspondence between the cartesian products $A \times B$ and $B \times A$. The central idea in the proof of this fact is the old observation that a rectangle of n rows of m dots contains $n \cdot m$ dots, but turned on its side it contains $m \cdot n$ dots. I take it that this explains why multiplication is commutative, and their ability to provide such an explanation provides some theoretical support for the set theoretic axioms involved. ([Maddy 1981], 498-499)

This is an interesting idea, but I don't think it works. The main problem is that the envisaged proof doesn't look like proper set-theoretic reasoning at all, since rectangles made of rows of dots aren't objects of pure set theory. What Maddy is describing is instead a (partly) *diagrammatic* or *geometric* proof—i.e., something like the following.

1. If a and b are natural numbers, then the product $a \cdot b$ is equal to $|A \times B|$ (i.e., the number of elements in the Cartesian product of A and B, where these sets are the ordinals representing a and b respectively). [Fact]

2. $|A \times B|$ is equal to the number of dots in a rectangle consisting of $|A|$ rows of $|B|$ dots. ["Bridge premise", established by a separate argument]

3. The number of dots in a rectangle consisting of $|A|$ rows of $|B|$ dots is equal to the number of dots in a rectangle consisting of $|B|$ rows of $|A|$ dots. [Geometric premise, evident from an appropriately constructed diagram]

4. So $|A \times B| = |B \times A|$. [From 2 and 3]

5. Therefore $a \cdot b = b \cdot a$. Hence multiplication is commutative. [From 1 and 4]

The geometric premise 3 is clearly essential to the argument. I don't think it has any place in a supposedly pure set-theoretic proof, and so I don't think Maddy can claim the proof she describes as an explanatory victory for set theory. But suppose one is convinced for some reason that premise 3 is fair game for the set theorist. Then, presumably, it should also be fair game for the number theorist. In that case, though, nothing prevents the number theorist from constructing a similar proof of her own:

1. The product $a \cdot b$ is equal to the number of dots in a rectangle consisting of a rows of b dots. [Bridge premise]

2. The number of dots in a rectangle consisting of a rows of b dots is equal to the number of dots in a rectangle consisting of b rows of a dots. [Geometric premise]

3. Therefore $a \cdot b = b \cdot a$. Hence multiplication is commutative. [From 1 and 2]

The second proof is simpler, more transparent, and arguably more explanatory than the first, since it avoids the extra detour through set theory. So even if we allow proofs of the commutativity of multiplication that appeal to dot diagrams, the set-theoretic viewpoint still has no explanatory advantage over the "naïve" number-theoretic approach.

In response to this criticism, one might claim that the so-called geometric premise should in fact be viewed as a set-theoretic proposition, so that Maddy's proof is really a purely set-theoretic argument. The idea could be cashed out in a number of different ways. One could hold, for instance, that the notions of dots and rectangles are somehow *dependent* upon, or *grounded* in, or *best understood* in terms of the primitive notions of set theory, e.g. the concept of Cartesian product.

I'm not convinced by this line of thought. The correct response will depend to some extent on the way the details are spelled out, of course. Nevertheless, I see no reason to accept that set theory is epistemically or metaphysically prior to geometry in any interesting sense. As with any mathematical theory, it's certainly possible to model arrangements of dots or whatever with sets. But this doesn't show that such objects are identical to sets, or metaphysically grounded in sets, or rendered more intelligible by sets. And there's plenty of reason to doubt such claims. (Indeed, the idea that geometry depends in some such way on set theory may be more dubious than the corresponding claim about numbers and sets.) Compare these remarks of Moschovakis on the relation between lines, points and sets of numbers:

> What is the precise meaning of this "identification" [of a line with a set of real numbers]? *Certainly not that points are real numbers*. Men have always had direct geometric intuitions about points which have nothing to do with their coordinates... Every Athenian of the classical period understood the meaning of the sentence 'Phaliron is

> between Piraeus and Sounion along the Saronic coast' even though he was (by necessity) ignorant of analytic geometry. In fact, many educated ancient Athenians had an excellent understanding of the Pythagorean theorem, without knowing how to coordinatize the plane. ([Moschovakis 2006], 33, emphasis in original)[21]

If this is true for lines and points, as I think it is, then something analogous is surely also true for arrangements of dots. We can "identify" the content of such a picture with the Cartesian product of two sets if we want. But to suggest that this content depends in some way on products of sets, or that it can only be properly understood in light of set theory, is a different (and much less plausible) story. On the contrary, it's exactly the visuospatial character of the diagram that makes premise 3 of the Maddian argument obvious, and in turn makes the whole proof explanatory. And this visuospatial character is lost when we boil the diagram down to a proposition about sets.

In any case, Maddy's framing of the situation is misleading in a more basic way. To explain why, let's say that a proof of an arithmetical theorem is *reductive* if the proof involves viewing numbers as sets and *non-reductive* otherwise. Maddy's argument compares a reductive proof of the commutativity of multiplication with one particular non-reductive proof, which starts from the Peano axioms and proceeds by induction. Maddy may well be right that this proof is unexplanatory. (Cf. [Lange 2009], which argues that proofs by induction are unexplanatory as a rule.) But the more interesting question here is whether some reductive proof is more explanatory than *any* non-reductive proof. It could be, after all, that Maddy's non-reductive proof is particularly unexplanatory, but that better alternatives exist. Even if Maddy's claims are all true, their force would be less clear in this sort of situation.

So are there better non-reductive proofs of commutativity? I think so. One such proof, for instance, is a simple and compelling argument based on the Principle of Recursive Definition. The idea here is to show that the functions $m \times n$ and $n \times m$ both satisfy the definition of multiplication

$$\forall m, n \in \mathbb{N} : \begin{cases} m \times 0 & = 0 \\ m \times (n+1) & = m \times n + m. \end{cases}$$

[21] The sentence in single quotes is displayed on a separate line in the original.

Since the Principle of Recursive Definition says that there exists exactly one function satisfying such a recursion formula, it follows that $m \times n = n \times m$, so multiplication is commutative. (See Theorem 16.10 and its proof in [Warner 1990].) I don't want to claim too much on behalf of this proof, but it at least doesn't strike me as totally lacking in explanatory value. Note that it allows a straightforward and fairly satisfying answer to the commutativity question: multiplication is commutative because, if it weren't, there would be two distinct functions satisfying the recursive definition of multiplication, which is impossible.

I conclude that Maddy is wrong to claim that set theory explains the commutativity of multiplication. The proof she offers, though plausibly explanatory, is essentially geometric or diagrammatic, and this geometric element can't be harmlessly replaced with more set theory. And the purely set-theoretic, "reductive" version of Maddy's proof is no more explanatory than the non-reductive proof just mentioned.

Obviously Maddy's example represents just one way that set theory could be claimed to explain a fact about numbers. I don't claim to have ruled out all the other possible candidates. On the other hand, few philosophers have thought about set theory as much and as well as Maddy, so the failure of her preferred example is a discouraging sign for this line of thought in general.

1.2.5 A Kitcherian argument

The objection in the last section involved the possibility that, once we've reduced numbers to sets, some number-theoretic facts might turn out to be explained by set-theoretic considerations. The focus there was on individual theorems, like the commutativity of multiplication, and their proofs. Here I'll consider a different type of objection. Instead of (or in addition to) explaining particular results, the thought goes, it might be that set-theoretic reduction makes its explanatory contribution by shedding light on higher-level, relational features of the natural numbers and their arithmetic. Such features might include, for instance, revealing analogies or unexpected logical connections involving arithmetic and other branches of mathematics. A defense of this line of thought comes from Philip Kitcher:

> Mathematical proofs can serve other functions besides that of increasing the certainty of our knowledge. In some cases, as in the case at hand, they can be part of a scheme of ontological reduction in which our justifications for accepting the theorems proved remain unaffected. Moreover, without making us any more certain that a theorem is true, a proof can show us why it is true: proofs may yield explanatory dividends. ...Reducing arithmetic to set theory has explanatory, as well as ontological, value. For, in the light of the reduction, our understanding is advanced through exhibition of the kinship between theorems of arithmetic and theorems in other developments of set theory (in particular, branches of abstract algebra). ([Kitcher 1978], 123)

Kitcher doesn't explain what sort of kinship he has in mind, but I think the general idea is clear enough. When Kitcher speaks of abstract algebra and its influence on our understanding of arithmetic, he's presumably referring to the developments surrounding the birth of "modern algebra" around the beginning of the 20th century. This period—which saw the first general definitions and systematic study of objects like groups, rings and fields, owing to the work of Noether, Hilbert, Artin, Steinitz, van der Waerden and others—had major consequences for almost every area of mathematics, including our understanding of the natural numbers and related number systems.

One type of progress was classificatory. By categorizing number systems according to their abstract algebraic properties, modern algebraists got a better fix on the essential features of particular structures, and a better understanding of genus-species relations between structures. For instance, the natural numbers with addition $(\mathbb{N}, +)$ were seen to be an example of a *commutative monoid*, the integers with addition $(\mathbb{Z}, +)$ an *abelian group*, the integers with addition and multiplication $(\mathbb{Z}, +, \cdot)$ a *commutative ring*, the rationals $(\mathbb{Q}, +, \cdot)$ a *field*, and so on. Along with these taxonomic advances, algebraists figured out how to pass from the natural numbers to other number structures by way of rigorous algebraic constructions. The integers, for example, can be obtained as the Grothendieck group of the natural numbers (as the construction is now called). Similarly, the rationals arise as the fraction field of the integers, and the complex numbers as the algebraic closure of the reals.

Finally, algebraists in the modern era discovered general theorems about algebraic structures that shed new light on the properties of the familiar number systems. For instance, the Fundamental Theorem of Arithmetic—the fact that every natural number greater than 1 has a unique decomposition as a product of primes—had been known at least since Euclid. But only with the help of abstract algebra did the FTA clearly emerge as an instance of a more general phenomenon, involving notions like *commutative ring* and *irreducible element*[22]. The picture is as follows. A commutative ring in which every element factors uniquely as a product of irreducibles (and which lacks zero divisors) is called a "unique factorization domain". The integers $\mathbb{Z}$, of course, are a unique factorization domain, and it turns out that the irreducible elements of $\mathbb{Z}$ are exactly the primes. Taken together, these facts imply the FTA. So abstract algebra gives us a nice story about why the FTA holds for the natural numbers, where before this looked like a more or less brute fact. (It also reveals how FTA-like results hold, or fail to hold, in other structures.)

In light of these kinds of advances, it's reasonable to think that developments in algebra have explained some things about the natural numbers and their place in the mathematical universe. In particular, I take it that the example of the previous paragraph is the sort of thing Kitcher has in mind when he talks about "the kinship between theorems of arithmetic and theorems in... abstract algebra". And I agree that understanding why and to what extent results like the FTA hold in various number systems can improve our understanding of arithmetic, as this line of thought has it.

So where does the reduction of arithmetic to set theory enter this picture? According to Kitcher, both arithmetic and algebra count as "developments of set theory" in some sense, and this is supposed to explain why the reductionist viewpoint deserves the credit for the illuminating parallels between the two subjects. How exactly this argument is supposed to work is a little unclear. But perhaps what Kitcher has in mind is something like this. As a matter of historical fact, it's true that abstract algebra reached maturity around the same time as set theory, and set-theoretic language and tools exerted a notable influence on algebra during this period (as did as the axiomatic method

[22] An element of an integral domain is irreducible if it's neither zero nor a unit, and if it isn't expressible as a product of two non-units. (An integral domain is a commutative ring with no zero divisors, meaning that the product of nonzero elements is always nonzero. A unit is an element with a multiplicative inverse.)

pioneered by Hilbert, along with other trappings of the emerging "modern" style of mathematics). For instance, van der Waerden's *Moderne Algebra*—the first definitive modern algebra textbook, published in 1930—defines a group as a set equipped with a certain type of binary operation. And a great deal of progress during this period was due to Dedekind, who was extremely interested in rigorous set-theoretic foundations, and who used "set-theoretic tools and techniques... to construct new mathematical objects (the natural and real numbers, ideals, modules, etc.) or whole classes of such objects (various algebraic number fields, rings, lattices, etc.)" ([Reck 2016]).

With these facts in mind, one might be tempted to make the following sort of argument:

1. Abstract algebra is fundamentally set-theoretic in nature, and so anything explained by algebra is ultimately explained by set theory.

2. Algebra provides explanations of some arithmetical facts (e.g., the explanation of the FTA mentioned earlier).

3. Therefore, set theory provides explanations of some arithmetical facts.

I don't know whether this is exactly the argument Kitcher had in mind, but it's the most plausible way I can see to make the general idea more precise. Still, I think the argument is unsuccessful for at least two reasons.

For one, the first premise is doubtful, at least under any interpretation that's interesting for present purposes. To start with, it's certainly not the case that algebra developed *from* set theory, or that algebra couldn't possibly be done outside a set-theoretic framework. The roots of modern algebra lie with Lagrange, Ruffini, Galois and others in the 18th and early 19th centuries, well before Cantor was born. Their successes, and those of their pre-Cantorian followers, show that much of abstract algebra could—and surely would—have been worked out without the backing of a formalized set theory. The kinds of algebraic insights about number theory I've mentioned, in particular, mostly work just fine without a rigorous notion of set. There's nothing particularly set-theoretic about the "algebraic explanation" of the Fundamental Theorem of Arithmetic outlined above, for instance.

Actually, though, the truth or falsity of the first premise doesn't matter much in the end. That's because, even if the above argument is sound, it nevertheless misses its mark. The view we wanted to investigate—and the one Kitcher claimed to be defending—is that the reduction of arithmetic to set theory is explanatory. But if the above argument shows anything, it's that *algebra* is reducible to set theory, and that *algebra* explains some things about arithmetic. Nothing about this conclusion depends in any way on viewing *numbers* as sets.

Thus I disagree with Kitcher's assessment of the reduction of number theory to set theory. On the one hand, Kitcher is right that one theory can advance our understanding of another by showing how the latter is situated in the broader mathematical universe—by "exhibiting its kinship" with other areas of mathematics, showing its theorems to be special cases of more general phenomena, and so on. What's more, Kitcher is also right to think that abstract algebra has advanced our understanding of arithmetic in these ways. But he's mistaken to conclude that the explanatory virtues of algebra accrue to set theory. Still less do they accrue to the *reduction* of arithmetic to set theory.

1.2.6 A Quinean argument

I conclude this section with a line of thought due to Quine, tracing to his well-known discussion of set-theoretic reductions in §§53-54 of *Word and Object* ([Quine 1960]).[23] Quine's view is that such reductions constitute major epistemic and theoretical achievements, in that they rendered more intelligible various notions which are extremely useful but which were previously unclear. Indeed, he takes the reduction of the ordered pair to be exemplary of the best way to solve a certain kind of philosophical problem, and he holds the reduction of arithmetic in similarly high esteem. (This discussion is the source of Quine's famous dictum "explication is elimination".)

[23] Further references to Quine are to this work.

Although Quine rarely uses the language of explanation explicitly, his arguments can nevertheless be taken to pose a challenge to the thesis I've been defending. For, as I noted above, rendering a given type of entity more intelligible can be viewed as a sort of objectual explanation. And Quine certainly thinks that set theory plays this clarifying role.

I'll start by focusing on Quine's discussion of the ordered pair in §53, since it's in this section that he lays out his view most clearly. Briefly, Quine's argument is that the ordinary conception of the ordered pair is "defective", "perplexing", "inadequately formulated", and "problematic", whereas the set-theoretic conception is "clear and couched in terms to our liking" (258-260). The argument for this weighty claim, unfortunately, is brief and unconvincing. As evidence for the defectiveness of the pre-set-theoretic conception, Quine cites Peirce's definition of the ordered pair (or "Dyad"), which Quine considers confused and obscure. I won't go into the details of Peirce's account here. Even granting the unsatisfactoriness of the Peircean definition, however, it's not clear how Quine's conclusion is supposed to follow. For we're given no reason to think that Peirce's remarks are representative of the ordinary conception in any way, or that his account is the best sense we can hope to make of the ordered pair without the help of set theory.

On the contrary, understanding the ordered pair from the pre-set-theoretic viewpoint presents no particular problem. The characteristic property of pairs, $(x_1, y_1) = (x_2, y_2) \leftrightarrow x_1 = y_1 \wedge x_2 = y_2$, is perfectly clear, and it makes no mention of sets. (Quine himself seems to agree; he praises the ordered pair postulate as "preternaturally succinct and explicit" (259).) Moreover, the postulate was formulated already by Schröder in 1895 ([Schröder 1895]), some twenty years before the first set-theoretic definition due to Wiener ([Wiener 1914]).[24] It's not at all obvious what's supposed to be so problematic about conceiving of an ordered pair as a *sui generis* object satisfying this axiom. Mathematicians like Schröder apparently did so without suffering any unease or confusion. So Quine's complaints about the epistemic inadequacies of the pre-set-theoretic conception aren't very compelling.

What about the supposed epistemic advantages of identifying pairs with sets? Even if the ordinary conception is clear enough on its own, after all, it's possible that the set-theoretic conception is

[24] See [Dipert 1982], fn. 23 for more on the early history of the ordered pair in modern logic.

better yet in some interesting way. But this is also doubtful. Although the naïve notion of set may have had some claim to special clarity, simplicity, and the like, modern axiomatized set theories are far more complex, and the nature of the sets they describe remains mysterious in many ways. It's hard to see what gains are to be made by trading the *sui generis* ordered pair for a construction that carries so much epistemological baggage. Indeed, I'm inclined to agree with Dipert's opinion that

> the pre-Wiener, primitive notion of ordered pair in Peirce and Schröder is actually clearer than its alleged set-theoretic representations... since the notion of set is itself so cumbersome and obscure. And a New Foundations set is still more obscure. ([Dipert 1982], 367-8)[25]

Michael Hallett notes likewise that

> 'set' (unlike 'aggregate' perhaps) is not an ancient, well-understood concept which can easily be taken as an axiomatic primitive in the knowledge that it can be supported by extra-axiomatic explanation. (Unlike the case of natural number, this is largely *why* set theory is axiomatized, because we do *not* understand the set concept well.) ([Hallett 1984], 300, emphasis in original)

There seems to be little reason, then, to accept either component of Quine's view. The pre-set-theoretic conception of the ordered pair isn't unclear or mysterious, except perhaps to the extent that one finds it hard to wrap one's mind around ordered pairs as *sui generis* objects. But the set-theoretic conception is hardly an improvement on this score. If anything, the sets postulated by ZFC, NF and similar systems are more problematic and harder to understand than ordered pairs. So, contra Quine, there's no obvious epistemic advantage to viewing pairs as sets.

So much for ordered pairs—what about the reduction of arithmetic to set theory? Quine gives essentially the same arguments here as in the previous case, and he draws a similar conclusion. ("But for its greater antiquity and its concern with a more venerable notion," he says, "the philosophical question 'What is a number?' is on a par with the corresponding question about ordered pairs" (262).) Quine's charge is that the reductions of Zermelo and von Neumann brought clarity

[25] New Foundations is Quine's system of set theory, which admits non-well-founded sets and has various other unusual features.

to the pre-set-theoretic conception of the natural numbers, which on its own was muddled and inadequate. In response, we can take essentially the previous line—exchange the ordered-pair postulate for the second-order PA axioms, and as before we have a satisfactory characterization of the relevant domain which doesn't involve viewing numbers as sets, and which set-theoretic reduction does nothing to improve on. Hence there are some plausible epistemic disadvantages, but no apparent advantages, to treating numbers as sets.

Although Quine was wrong to take set-theoretic reductions as his main examples of epistemically salutary explications, I'm inclined to think there are better ones. Consider, for instance, the replacement of the classical notion (or notions) of infinitesimals with the modern construction due to Robinson.[26] Unlike the ordinary conception of ordered pairs or numbers, the old notion of infinitesimals really was defective, perplexing, and inadequately formulated. In certain circumstances, for instance, the classical analysts had to "neglect" infinitesimal quantities by treating them as equal to zero, while at other times infinitesimals were assumed positive in order to function as divisors. These practices were decried by Berkeley and other critics, and only uneasily tolerated by many working mathematicians. Even Newton and Leibniz themselves made various (unsuccessful) attempts to remove infinitesimals from the calculus, or at least to explain them away as mere heuristic devices with no official standing. Later, the 19th-century "arithmetization of analysis" replaced infinitesimals with the now-standard epsilons and deltas. But Abraham Robinson's work on nonstandard analysis in the 1960s[27] led to a sort of redemption. Robinson identified infinitesimals with certain elements of $^*\mathbb{R}$, the non-Archimedean field of hyperreal numbers, and showed how to do analysis rigorously on the basis of this construction.

The history of the infinitesimal concept furnishes what I take to be a more plausible case of epistemic progress through Quinean explication. The story seems to comfortably fit the Quinean mold: it involves the replacement of a useful but problematic concept with a relatively clear and concrete alternative. The alternative—Robinson's nonstandard reals—may not exactly match what the early analysts had in mind, and it may not have any special claim to being the "one true notion

[26] Quine can't be faulted for overlooking this example; *Word and Object* was published in 1960, while Robinson's work on nonstandard analysis didn't appear until later in the decade.

[27] See [Robinson 1974].

of infinitesimal". But this is no objection, for Quine. What matters is just that the Robinsonian infinitesimals can do the useful work done by the original notion without partaking in its defects. ("We do not claim synonymy. We do not expose hidden meanings... we supply lacks" (258).)

Importantly, the Robinsonian explication of the infinitesimal concept is more than just a technical fix. It also represents an epistemic advance over earlier ways of thinking, in that it renders the notion of infinitesimal more intelligible and paves the way for new explanatory proofs in analysis. As F.A. Medvedev says,

> Nonstandard analysis makes it possible to answer a delicate question bound up with earlier approaches to the history of classical analysis. If infinitely small and infinitely large magnitudes are regarded as inconsistent notions, how could they serve as a basis for the construction of so grandiose an edifice of one of the most important mathematical disciplines? ([Medvedev 1998], 664)

Similarly, according to the historian of analysis Henk Bos, "A preliminary explanation of why the calculus could develop on the insecure foundation of [infinitesimals] is provided by the recently developed non-standard analysis, which shows that it is possible to remove the inconsistencies without removing the infinitesimals themselves" ([Bos 1974], 13).

To sum up, the problem with Quine's discussion of explication isn't that it describes a nonexistent phenomenon. In the history of mathematics, at least, epistemic progress does sometimes happen in much the way Quine describes. Only his examples are off the mark. Contrary to Quine's claims, arithmetic and the ordered pair were never in need of set-theoretic explication, for two reasons. First, the relevant concepts aren't unclear or defective in the way Quine suggests. Second, set theory is in no position to offer help at any rate, since the sets of modern axiomatized systems like ZFC are if anything more obscure than numbers and ordered pairs. The contrast with the case of infinitesimals—a genuine example of better understanding achieved through explication—makes these points especially clear.

1.2.7 If not explanation, then what?

A final question is worth addressing before this section ends. If the reduction of arithmetic to set theory has no explanatory value (or other epistemic benefits), as I've argued, then what sort of value does it have? Why bother reducing numbers to sets at all?

One possible answer is that the reduction has no (or very little) value. Someone might think this, perhaps, because they're convinced that a legitimate foundation for mathematics must be a sort of Cartesian-style *epistemic* foundation—i.e., that it should offer greater certainty, clarity, insight, explanation, understanding, or whatever package of epistemic goods one takes to be necessary. This type of view is a recurring theme, for instance, in Hallett's *Cantorian Set Theory and Limitation of Size*. I quote a particularly forceful passage at length:

> Because of the reductionist ambition, we demand that set theory genuinely explain all other mathematical concepts. When I say 'we demand', I mean that is what we should demand. Useful reductionism (philosophically speaking) cannot be just successful theoretical translation, though of course this would suffice in a relative consistency exercise, or as an original part of Hilbert's programme, say. Rather reduction must (in the best instance) be accompanied by a gain in conceptual clarity. Where set theory suffers as a foundation framework is that in general it does *not* bring this conceptual clarity with it. It is no good, philosophically speaking, reducing to set theory something so basic to human thought as the elementary theory of natural number if you cannot also explain why numbers are sets, and why the set concept is even more fundamental. But the set concept is too unclear for any such explanation to be given. ([Hallett 1984], 300-301)

As should be clear by now, I largely agree with Hallett's pessimistic assessment of the epistemology of set theory, and of the reduction of numbers to sets. And Hallett may even be right that an ideal foundation for mathematics would be epistemically illuminating and metaphysically revealing in the ways that set theory often fails to be. If so, we may be justified in continuing to look for an alternative framework—one based on categories or homotopy types[28] rather than sets, maybe. Still, I want to resist the conclusion that set-theoretic foundations are "no good" or not "useful"—either "philosophically speaking", whatever exactly that means, or otherwise. There are

[28] See [Linnebo & Pettigrew 2011] and [Ladyman & Presnell 2014], respectively, for discussion of these frameworks as possible foundations for mathematics.

plenty of ways for a foundational theory to earn its keep, and a failure along the dimensions Hallett mentions needn't be a failure across the board. Hence we don't have to put ourselves in the awkward position of claiming, *pace* long-standing mathematical practice, that set-theoretic reductions are a useless boondoggle.

Many people have written about set theory and its role in foundations, but I find Maddy's recent work on this issue particularly enlightening.[29] [Maddy 2011] and [Maddy 2017] take up more or less the question just raised—given that set theory can't serve as a Hallett-style epistemic and metaphysical basis for the rest of mathematics, in what way can it serve as a legitimate foundational theory?[30] In Maddy's view, set theory succeeds in playing a number of distinctively foundational roles, including the following:

- **Meta-mathematical Corral**: By translating all of mathematics into set theory, it becomes possible, or at least much easier, to obtain general results about consistency and provability for mathematics as a whole. (Cf. Hilbert's program and Gödel's theorems.)

[29] See also [Shapiro 2000], which defends a view similar to Maddy's. For a contrary picture, see [Mayberry 1994].

[30] Maddy's early work, including the 1981 paper discussed above, espoused "set-theoretic realism"—a sort of reductionist view according to which numbers are properties of sets. Maddy has since abandoned set-theoretic realism and its metaphysical commitments, which explains why the recent papers mentioned here strike a quite different tone.

- **Elucidation**: Some set-theoretic realizations of classical mathematical notions introduce a degree of clarity and precision that was lacking in the original concepts (e.g., Dedekind's construction of the real numbers as cuts of the rationals).[31]

- **Risk Assessment**: A proof that a given theory is consistent relative to ZFC (or a lack of such a proof) provides an important measure of the theory's logical "safety" or "riskiness".

- **Shared Standard**: The ZFC axioms function as a shared standard of existence and provability for mathematicians working in diverse specialty areas.

- **Generous Arena**: The set-theoretic universe serves as a single forum in which "all the various structures studied in all the various branches [of mathematics] can co-exist side-by-side, where their interrelations can be studied, shared fundamentals isolated and exploited, effective methods exported and imported from one to another, and so on." ([Maddy 2017], 305)

Clearly these functions are important ones, and each of them plausibly pertains to foundations. Moreover, nothing requires that sets be metaphysically or epistemically fundamental in order to play these roles well. So we have a good start on an answer to the question posed at the beginning of this part. Even if the reduction of arithmetic to set theory has no explanatory benefits, this

[31] Maddy's "elucidation" is apparently much the same sort of thing as Quine's "explication". For the record, I have some reservations about the extent to which Dedekind's construction of the continuum is a good example of this phenomenon. One reason to be careful here is because, as Solomon Feferman has argued, there may in fact be several different, and equally worthwhile, conceptions of the continuum. So it's not even obvious what would count as a successful explanation (or explication or elucidation) of continuity and its properties. Moreover, the device of Dedekind cuts does a poor job capturing at least one such conception, namely the one that's operative in Euclidean geometry. As Feferman writes: "The main thing to be emphasized about the conception of the continuum as it appears in Euclidean geometry is that the general concept of set is not part of the basic picture, and that Dedekind style continuity considerations... are at odds with that picture. It does not make sense, for example, to think of deleting a point from a line, or to remove the end point of a line segment. Given two line segments L and L', we can form a right triangle with legs L_1 and L'_1 congruent to the given segments, respectively; but these share a vertex as a common point, each an end point. Thought of as a set, L is transformed into L' by a rigid motion, and the same for L_1 and L'_1. Thought of in that way, the vertex of the right triangle has displaced one of the end points, but which one? There are many similar thought experiments which dictate that lines, line segments and other figures in Euclidean geometry are not to be identified with their sets of points" ([Feferman 2009], 174-5). Of course, "Dedekind's continuum" was quite useful for the foundations of analysis and for other purposes. In any case, the issue is complex, and I can't pretend to do it full justice here. (Thanks to the two anonymous referees from *Synthese* for prompting me to say more about this example.)

doesn't entail that it and other such reductions are without value. Translating arithmetic and other theories into the language of sets contributes in an important way to a worthwhile end, namely the orderliness, systematicity and communality of mathematics generally.

1.3 Implications and further questions

The contention of the previous section, if correct, raises a number of interesting issues for philosophy of science and philosophy of mathematics. I won't try to discuss these issues in detail here—that's a task for future work– but a few are at least worth pointing out and briefly commenting on.

One obvious question to ask is whether or not *all* reductions in mathematics follow the arithmetic-set theory case in being unexplanatory. In my view, the answer is almost certainly no. On the contrary, there are pairs of mathematical theories such that one is plausibly reducible to the other, and also such that the reducing theory sheds significant explanatory light on the reduced theory.[32] To take one example, consider the advances in algebraic geometry made possible by Alexander Grothendieck's theory of schemes. (Roughly speaking, Grothendieck's idea was to view the *varieties* studied by classical algebraic geometry as a species of *scheme*, a superficially quite different-looking type of object consisting of a topological space with a commutative ring paired to each open set.) It's not hard to see that, given suitably formalized versions of both theories, classical algebraic geometry should come out reducible to scheme theory. Moreover, there's broad agreement among mathematicians that the scheme-theoretic viewpoint provides important insights about various classical phenomena. Indeed, according to Vakil, "the wonderful machine of modern alge-

[32] In light of the infinitesimal example from §2.6 and the other cases considered so far, one might wonder, as an anonymous referee did, whether the difference between explanatory and non-explanatory reductions is just the difference between cases where the relevant concepts weren't or were originally "in good working order". It seems to me that the example given below challenges this idea. We seem to have an explanatory reduction of algebraic varieties to schemes, but there was nothing unclear or contradictory about the notion of variety before Grothendieck came along. And I don't think anyone would suggest that the notion of a scheme—which involves a lot of complex technical machinery with no immediately obvious geometric meaning—is simpler, clearer, more intuitive, or otherwise more epistemically or logically adequate than the notion of a variety. What's going on in this case seems to be, rather, that schemes are *richer in structure* and that they *carry more data* than varieties, and this means that the scheme-theoretic viewpoint allows one to see further into the phenomena than the classical approach allows.

braic geometry [i.e., scheme theory and related ideas] was created to understand basic and naive questions about geometry" ([Vakil 2015], 12). (To take a particular example of Perrin's, "In the plane... we can always find lines which are not tangent to a given curve. The notion of a scheme is indispensable for understanding these phenomena" ([Perrin 2008], 213).) Eisenbud and Harris's textbook on schemes makes the broad scope of Grothendieck's achievements clear:

> The theory of schemes... is the basis for a grand unification of number theory and algebraic geometry, dreamt of by number theorists and geometers for over a century. It has strengthened classical algebraic geometry by allowing flexible geometric arguments about infinitesimals and limits in a way that the classic theory could not handle. In both these ways it has made possible astonishing solutions of many concrete problems. ([Eisenbud & Harris 2000], 1)

(This case deserves a much more detailed examination, which I hope to give it in future work.) Other examples of explanatory mathematical reductions include, perhaps, the application of Galois theory to questions about the solvability of polynomial equations, and of matroid theory to the notion of "independence".

Supposing this is right, reducibility relations in mathematics evidently come in two different flavors—those that are substantially explanatory and those that aren't (by virtue, perhaps, of being in some sense merely "conventional"). A second question is then how to account for this difference. Is it, for instance, a matter of metaphysics? On this type of view, explanatory reductions are explanatory because the objects in the domain of the reduced theory are metaphysically related to those in the domain of the reducing theory in some appropriate way. This relation could simply be identity, for example. (One might think that the reduction of arithmetic fails to be explanatory precisely because numbers aren't identical to sets.) Or the relation could be something weaker—parthood, maybe, or even metaphysical dependence of a sort peculiar to mathematical objects.[33] This last idea, incidentally, suggests the need for a closer look at the relationship between mathematical explanation and the currently popular notion of metaphysical grounding.[34]

[33] See [Pincock 2015] for a defense of a dependence-style account.

[34] [Correia & Schnieder 2012] collects some recent work on the subject.

Alternatively, the difference between explanatory and unexplanatory reductions might lie in the logical or epistemological features of the theories themselves, rather than the properties of the objects they describe. On this type of view, explanatory reductions are explanatory because the reducing theory is logically stronger or conceptually richer than the reduced theory, or because it yields new knowledge of an appropriate kind about the reduced theory's subject matter. (I happen to think the second approach has the brighter prospects. But defending that claim is a task for another time.)

At any rate, answering this second question ought to help with a third, related one: why are reductions in empirical science—at least in natural, interesting, and relatively uncontroversial cases—apparently always explanatory, if the same isn't true in mathematics? In addition to the ideas sketched above, one might think that the difference is due to the relatively "empiricist" standards of theory acceptance in force in the natural sciences. Consider string theory, for instance. Contemporary versions of string theory (and its siblings, like M-theory) promise a sweeping reduction-cum-unification of classical particle physics, spacetime theory, and perhaps more besides. As its proponents have pointed out, such a string-theoretic foundation for physics would be formally elegant, conceptually neat, and institutionally convenient—that is, it would provide many of the same benefits that set-theoretic foundations do for mathematics. Unlike set theory, however, string theory has so far failed to win general acceptance. And the proffered reduction of elementary particles to strings is still viewed with skepticism by many physicists.

As is well-known, a major reason for this hesitation is the lack of independent experimental evidence in support of the theory.[35] Without such evidence in its corner, it seems, even the most attractive empirical theory rarely if ever gets itself accepted by the scientific community—and one can't have a proper reduction without a widely accepted reducing theory. On the other hand, if and when we do have good independent evidence of the existence of strings, it seems we'll *ipso facto* have evidence for an explanatory connection between string theory and the theories it reduces. (Since we'll then have good reason to think, for instance, that classical particles are really strings, and that the laws governing the former are just a special case of the laws governing the latter. And

[35] For an extended critique of this sort, see [Woit 2006].

such "interlevel" facts—about cross-theoretic identities, composition relations, and the like—are seemingly always explanatory.) So we seem to have the following situation. If an empirical theory is a candidate for involvement in a "purely conventional" reduction—because it has no suitably direct evidential support, but only considerations of formal niceness and unificatory power in its favor—then the theory and the associated reduction won't command general assent. On the other hand, if an empirical theory amasses enough experimental support to be widely endorsed, then this very evidence will underwrite acceptance of the entities and laws the theory describes. And the "interlevel" information typically derivable on this basis is explanatorily valuable by nature.

As we've seen, things are otherwise in pure mathematics. A mathematical theory, and the reduction relations in which it's involved, can win institutional acceptance on the grounds of elegance, neatness and convenience alone; there's no counterpart to the demand for (relatively) direct empirical evidence. Hence unexplanatory, merely "conventional" reductions are possible in mathematics, but typically not in the sciences. (Or so this line of thought would have it. I think the story has some plausibility, but I'm not certain it's entirely or uniquely the right one.)

This essay has covered a lot of ground; let me briefly summarize where we've been. My main goal has been to show that the reduction of arithmetic to set theory is unexplanatory. I've argued that there are both positive and negative reasons to accept this conclusion. On the positive side, I suggested that the reduction of arithmetic is similar in relevant ways to other paradigmatically unexplanatory set-theoretic reductions, and moreover that number theorists don't regard the reduction as a source of insight or understanding. On the negative side, I considered four arguments purporting to show that the reduction of arithmetic is explanatory after all—the first claiming that numbers are identical to sets as a matter of metaphysics, the second claiming that there are explanatory set-theoretic proofs of arithmetical facts, the third claiming that the reductionist viewpoint explains high-level features of the natural numbers, and the fourth claiming that reductionism renders arithmetical concepts more intelligible. I've tried to show that none of these arguments are successful. This outcome needn't lead to a pessimistic view of set-theoretic foundations, though, since there remain a number of ways in which an unexplanatory reduction to a foundational theory might still be valuable. Finally, as I hope the last section has shown, accepting these conclusions raises a number of provocative questions for philosophers of science and mathematics interested in the epistemology and metaphysics of intertheory relations.

Chapter 2

Mathematical Explanation Beyond Explanatory Proof

Near the beginning of 'Mathematical Explanation'—the founding document of the recent literature on the subject—Mark Steiner writes: 'Mathematical explanation exists. Mathematicians routinely distinguish proofs that merely demonstrate from proofs which explain' ([Steiner 1978a], p. 135). Judging from his treatment of the subject in the rest of the paper, Steiner seems to intend the second claim as something like an elaboration of the first, rather than as an example of one sort of mathematical explanation among others.[1] That is, Steiner appears to endorse approximately the following view:

> Proof Chauvinism: All or most cases of mathematical explanation involve explanatory proofs in an essential way.

Steiner's contemporaneous paper on mathematical explanations in science ([Steiner 1978b]) is perhaps even more explicitly committed to proof chauvinism. In that essay, Steiner claims that we have a mathematical explanation of a physical phenomenon just in case 'when we remove the

[1] Near the end of the paper, Steiner also mentions explanations 'of the "behavior" of a mathematical object' (p. 148), and explanations obtained by regarding one mathematical object as another. It's not clear what Steiner takes to be the *explanantia* in these cases. As best I can tell from Steiner's discussion, the *explanans* in the first case involves a proof or at least some sort of argument. I'm unsure what the *explanans* in the second case is supposed to be. Steiner characterizes these cases as non-examples of 'explanation by proof', but it's unclear how to interpret this remark in light of the surrounding discussion and the commitments he expresses elsewhere.

physics, we remain with a mathematical explanation—of a mathematical truth' (p. 19). A few lines later, Steiner seems to suggest that 'remaining with a mathematical explanation' amounts to obtaining an explanatory proof of the relevant theorem.[2] On this view, then, an explanation of a mathematical fact just is an explanatory proof of that fact.

Much has been written on the subject since Steiner's work appeared, and his theory of mathematical explanation is no longer popular.[3] For one reason or another, however, Steiner's proof-chauvinist assumptions have taken deep root. Subsequent papers on mathematical explanation have almost invariably dealt with some issue in the theory of explanatory proofs—for instance, whether and why a particular proof explains its result ([Sandborg 1998]; [Mancosu 2001]; [Lange 2015]), whether certain general types of proof are explanatory or not ([Cellucci 2008]; [Lange 2009]; [Baker 2010]), or the factors that contribute to a proof's being explanatory ([Mancosu 1999]; [Lange 2014]; [Pincock 2015]). Some authors admit that there might be other kinds of mathematical explanation, but they almost never say what these kinds might be or discuss specific examples. And it's still common to see mathematical explanation in general run together with explanation-by-proof in particular, as in passages like the following.:

> The topic of mathematical explanation seems to me a central one for an accurate understanding of mathematics. Mathematicians, as we have seen, do not simply struggle to obtain rigorous and compelling proofs. They often look for new proofs or consider old proofs unsatisfactory on account of their lack of "explanatoriness". ([Mancosu 2001], p. 102)

The impression given by all this is that mathematical explanations not involving explanatory proof are vanishingly rare, or not very interesting, or both. I aim to show that this view is mistaken,

[2]The comment about explanatory proof is made in the context of Steiner's analysis of a particular example, and he doesn't make it totally clear whether this feature of the example is supposed to generalize. But the whole discussion suggests that Steiner does endorse the generalization. ([Baker 2012]), for instance, certainly reads Steiner this way.

[3]For criticisms of Steiner's view, see for example [Resnik & Kushner 1987], [Hafner & Mancosu 2005], and [Pincock 2015]. [Baker 2012] argues against Steiner's account of mathematical explanations in science. Baker's line is that some mathematical results contribute to scientific explanations even though no explanatory proofs of the results are known. (For instance, the perimeter-minimizing optimality of the hexagonal tiling of the plane partly explains why honeybees build hexagonal cells. But the only available proof of the optimality theorem is not at all explanatory.) As will become clear below, I agree with the idea that theorems without explanatory proofs can nevertheless be explanatory themselves. (Thanks to an anonymous referee for reminding me about [Steiner 1978b] and [Baker 2012] and their relevance to the chauvinism issue.)

and importantly so. Proof chauvinism has already led to missteps in the theory of mathematical explanation, and its continued hegemony threatens to make things worse. (Hence it's not merely, as one might have thought, that philosophers have focused on explanatory proofs just because they're the most distinctively mathematical sort of mathematical explanation. Such an attitude might have led to an imbalanced distribution of research efforts, but it probably wouldn't be so bad on the whole. Rather, I'll argue that proof chauvinism has had more insidious effects—that it's led to theoretical errors and misdiagnoses of important cases, for instance.)[4]

I'll begin by considering the epistemic status of chauvinism. What sorts of reasons could there be for accepting it, as many philosophers apparently do? Neither Steiner nor anyone else, as far as I'm aware, has defended the principle explicitly, but its possible justifications seem to fall into two general categories. The first kind is something like this: We should accept chauvinism because it faithfully reflects mathematical practice. When mathematicians pursue explanations or make claims about what explains what, according to this line of thought, explanatory proofs are generally their targets. Hence, as philosophers striving to ground our theories in the realities of working mathematics, we should be chauvinists. The second sort of possible justification runs as follows: We should accept chauvinism because it has a compelling philosophical rationale. Whatever mathematicians might or might not say, the theory of explanation presents us with good reasons—perhaps emanating from general philosophy of science, or from elsewhere in the philosophy of mathematics—to think that explanatory proofs must be central to mathematical explanation. Hence, as philosophers striving to achieve a unified picture of the theoretical landscape, we should accept chauvinism.

The next section discusses several arguments of the above kinds. Its conclusion is that philosophical considerations offer no support for proof chauvinism, and mathematical practice strongly suggests its falsity. The rest of the paper provides some reasons for caring about the chauvinism issue. I look in some detail at the explanation of the solvability of polynomial equations provided by Galois theory, which has often been thought to revolve around an explanatory proof. I argue

[4] Thanks to an anonymous referee for prompting me to make this point more clear.

that this diagnosis is mistaken. The paper concludes with some general worries about the effects of chauvinism on the theory of mathematical explanation.

2.1 Why I Am Not a Proof Chauvinist

2.1.1 Proof chauvinism and mathematical practice

As suggested above, one reason for embracing proof chauvinism is that it might be thought to reflect mathematical practices of seeking and offering explanations. It's not hard to see why this view might seem attractive. After all, mathematicians certainly do talk often about the explanatory virtues of various proofs (and proof methods)—the recent literature on mathematical explanation has shown at least that much beyond doubt. Generalizing from these cases, one might form the impression that proofs are the only sort of explanation that mathematicians care much about.

But this impression is erroneous. Mathematicians often cite items other than proofs as the *explanantia* of mathematical explanations: for example, theories, diagrams and other sorts of pictures, and particular facts of various descriptions. Examples of each kind of attribution are easy to find. Here are just a few:

- 'The following theorem, which we give here without proof, provides a necessary and sufficient condition for a number to be the sum of two squares: *An integer a is the sum of two squares if and only if any prime congruent to* 3 (mod 4) *in the prime factorization of a appears an even number of times in the factorization.* This theorem explains why 12 is not the sum of two squares.' ([Maor 2007], p. 224; emphasis in original)
- 'The next lemma explains why certain degenerate cases, singled out by the condition $\check{\chi} = \chi$, sometimes simplify [. . .] *Let χ be a character on a group G such that $\check{\chi} = \chi$. If G is generated by its squares, and in particular if G is* 2*-divisible, then* $\chi = 1$.' ([Stetkær 2013], p. 31; emphasis in original)
- 'The fact that $x^m - 1$ is reducible over $\mathbb{Q}$... explains why its roots do not all have the same algebraic properties over $\mathbb{Q}$.' ([Newman 2012], p. 92)
- 'The central limit theorem is the reason why the Gaussian distribution is the limit of the binomial distribution.' ([Keshav 2012], p. 40)

- 'Complex analysis explains why the Taylor series for the function $f(x) = 1/\left(1+x^2\right)$ converges in the finite domain $-1 < x < 1$ even though $f(x)$ is smooth for all real x.' ([Bender et al. 2009], p. 1)

- 'An arithmetic quotient of a bounded symmetric domain has several compactifications, that is, Satake-Baily-Borel's, Mumford's toroidal and Looijenga's compactifications. A general theory explaining the relations between these types of compactifications and those of a geometric nature [. . .] has been developed by Looijenga.' ([Artebani & Kondo 2011], p. 1445)

- 'Let F be the field of rational numbers or the Galois field $\mathrm{GF}(p)$, where p is a prime. Let A be the underlying additive group of F. Pick an abelian group X [. . .] [Then] $\mathrm{Hom}(A,X)$ is A-cellular. Also, the following diagram explains why the "evaluation at one" map is an A-equivalence,

$$\begin{tikzcd} & X \\ A \arrow[ur, "\forall\tau"] \arrow[r, "\check{\tau}:\alpha\to\{1\to\alpha\tau\}"'] & \mathrm{Hom}(A,X) \arrow[u, "ev"'] \end{tikzcd}$$

  where $1 \to \alpha\tau$ is the unique homomorphism taking 1 to $\alpha\tau$.' ([Farjoun et al. 2007], p. 66)

In these and many similar cases, we're confronted with mathematical explanations[5] in which proofs seem to play no obvious role. I think this impression is basically right. Hence I think mathematical practice lends no support to proof chauvinism. The chauvinist, however, has a number of options for trying to resist this conclusion. Let me present what I take to be the most pressing concerns; answering them will help clarify just what is and isn't shown by examples like the above.

An initial worry is whether diagrams can count as proofs, as some have argued that they can (for instance [Brown 1997]). If so, then explanatory diagrams may just be explanatory proofs, so that these cases aren't counterexamples to chauvinism after all.

It may well be true that some diagrams are proofs. But the claim that *every* explanatory diagram is a proof is much stronger, and there's no obvious reason to accept it. For instance, it's doubtful that the commutative diagram above, taken by itself, proves anything about the map *ev*. (Compare the way in which, e.g., an appropriate dot diagram might be thought to prove that the sum of the first n odd numbers is n^2.) At any rate, even if it's true that all explanatory diagrams are proofs, this doesn't put any pressure on the other types of example.

[5]Or apparent mathematical explanations, anyway. I don't mean to take a stand on whether each of the above explanatory claims is actually correct.

Similarly, it's unclear whether claims about explanatory theories might not ultimately depend on facts about explanatory proofs. At the very least, it's hard to deny that true claims about explanation-by-theories are sometimes grounded in facts about explanation-by-components-of-theories. In the empirical sciences, these components are sometimes facts or laws falling under the purview of the theory in question. Hence we might think that biological theory explains why the fox population increased at time t in virtue of the fact that the Lotka-Volterra equations explain why the fox population increased at t. (Compare van Fraassen's proposal that the claim 'theory T explains fact F' is equivalent to the claim 'there are facts which explain F relative to T'; [van Fraassen 1980], p. 101.) In the mathematical setting, couldn't the explanatory components of T just as well be proofs?

I don't deny that some claims about explanation-by-theories in mathematics are true in virtue of facts about explanation-by-proofs. But why think this is the typical case? Assuming that, say, theorems are often explanatory in their own right, we should expect many explanations-by-theory to rest instead on explanations-by-constituent-theorems. And perhaps there are cases in which theories are 'irreducibly' explanatory—that is, explanatory in a way that isn't grounded in the explanatoriness of any of their components.

Maybe so. Still, I admit that it's usually hard to tell exactly what's behind a given claim of the form 'theory T explains why p'. In the above cases, and in most others, there isn't any straightforward way to determine whether T is supposed to be explanatory in virtue of containing an explanatory proof, or an explanatory theorem, or an explanatory something-else, or in its own right.

Since the cases of diagrams and theories involve the complications just discussed, and since they're probably less important anyway, I'll mostly set them aside in what follows. My focus for the rest of the paper will largely be on the explanatoriness of *theorems* and other mathematical facts. (But I continue to maintain that explanatory diagrams and theories can't be generally assimilated to explanatory proofs.)

Another objection: why assume that claims like the above should be taken literally? Mathematicians' talk about explanations and reasons in such contexts, one might think, could just as well be understood as innocent *façons de parler* with no deep philosophical meaning. For instance, claims like 'theorem T explains why p' may just be stylized ways of expressing the thought that T implies p.

As before, there's no point in denying that some authors may sometimes use explanatory language in this way. But neither is there any reason to think this is always or usually the case. More to the point: if this suggestion is supposed to help the proof chauvinist, then she'd have to maintain that mathematicians generally speak literally when they refer to explanatory proofs, but non-literally when they make other sorts of explanatory claims. This idea seems quite ad hoc, and I can't see how anyone would try to defend it.

I've saved the next concern for last because it's weightier than the others, and the response it requires is relatively involved. The worry is this: given only a list of decontextualized claims like those above, it's difficult to tell whether proofs are in fact making some important explanatory contribution which the quoted text doesn't make obvious. For instance, when mathematicians say things like 'theorem T explains why p', this might be a loose or elliptical way of pointing to the explanatoriness of some proof of p from T—one that's offered or indicated elsewhere in the text, perhaps.

Certainly this is possible, and perhaps it's the right way to understand some cases. But typically there's no evidence that anything like this is going on. If the explanations in the above cases involved proofs in some significant way, then one would expect the explanatory features of those proofs to be highlighted at some point during the proof presentations or in the surrounding discussions. Otherwise it's hard to see what the point of the explanatory claims would be: if the intention is to communicate something about a proof, why say something about a theorem instead, and then fail to draw the reader's attention to the features of the proof that one had in mind? This would be a strange and unhelpful expository choice, to say the least.

Thus, if nothing about any proof is flagged as explanatory anywhere in the relevant text, I take this to be good evidence that no claim about explanatory proofs is intended. And this is in fact the situation in cases like the ones I've offered. There's just no suggestion in these cases that an explanatory proof is somehow lying behind an assertion about an explanatory proposition, or diagram, or theory. Indeed, in some cases a proof isn't so much as gestured at. This suggests that the examples in question are just what they appear to be: instances of mathematical explanation that don't involve explanatory proofs.

But perhaps this is too quick. Marc Lange has recently proposed a way of thinking about this issue that's congenial to chauvinism, but which doesn't seem to require that an explanatory proof be flagged in the way I've suggested. Lange's view isn't crudely chauvinist in the sense that it denies explanatory power to anything other than proofs. In fact it deserves credit for being, as far as I know, the only account of mathematical explanation which acknowledges that some theorems are genuinely explanatory. Regarding d'Alembert's theorem on roots of polynomials, for instance, Lange writes:

> Why is it that in all of the cases I have examined of polynomials with real coefficients, their nonreal roots all fall into complex-conjugate pairs? Is it a coincidence, or are they all like that? [. . .] In fact, as I have shown, d'Alembert's theorem is the explanation; any polynomial with exclusively real coefficients has all of its nonreal roots coming in complex-conjugate pairs. Here we have a mathematical explanation that consists not of a proof, but merely of a theorem. ([Lange 2016], pp. 241-2)

So far I'm happy to agree. As it turns out, however, Lange's theory about the nature of explanation-by-theorems is still chauvinist at heart. For Lange, '[a] theorem can explain [one of its instances] only if the theorem is no coincidence and hence only if it has a certain kind of proof' ([Lange 2016], p. 345). The kind of proof in question is what Lange calls a 'common proof' of all the theorem's cases—that is, a proof that exploits some respect in which the cases are alike, and which 'proceed[s] from there to arrive at the result by treating all of the (classes of) cases in exactly the same way' ([Lange 2016], p. 287).[6]

Such proofs, Lange thinks, are necessarily explanatory, at least in the appropriate context.[7] So even though he recognizes that some theorems genuinely explain, his proposal implies that the explanatory power of a theorem derives from the explanatory power of a proof. (If this is the case, why are the relevant features of the relevant proofs not always explicitly pointed out? Perhaps Lange

[6] See [Lange 2010], or Chapter 8 of [Lange 2016], for more on Lange's notions of mathematical coincidences and common proofs.

[7] The type of context in question is one in which we find it noteworthy, or salient, that the theorem identifies some property that all its instances have in common.

would say that his theory is tacitly understood by mathematicians, and assumed to be understood by their readers. So when someone writes 'theorem T explains why p', it's supposed to go without saying that this means 'T has a "common proof" that exploits some respect in which its instances, including p, are all alike'.)

There are a few things to say in response to Lange's view. First of all, his account only applies to theorems which explain their instances by subsuming them under a more general fact. But not all explanatory theorems work this way.[8] For instance, in the third example given above, the fact that the roots of $x^m - 1$ don't have all the same algebraic properties over $\mathbb{Q}$ isn't an instance of the fact that $x^m - 1$ is reducible over $\mathbb{Q}$. And yet the latter fact is supposed to explain the former. So even if Lange is right, his account only applies to a limited class of explanations-by-theorems, and there's no obvious way to generalize the proposal to cover other cases. So the strategy leads only to a partial vindication of chauvinism at best.

Second, it's doubtful whether Lange's view is in fact correct. Consider that mathematicians also make claims about what a theorem would explain if it were true, even though its truth value is actually unknown, and hence even though no proof yet exists. For instance, in his discussion of the current state of evidence for $P \neq NP$, William Gasarch writes:

> If a theory explains much data, then perhaps the theory is true. ...Are there a set of random facts that $P \neq NP$ would help explain? Yes. [. . .] [For example,] Chvatal in 1979 showed that there is an algorithm for Set Cover that returns a cover of size $(\ln n) \times OPT$ where OPT is the best one could do. Moshkovitz in 2011 proved that,

[8]Relatedly, an anonymous referee suggests that it may be useful to distinguish between explanations of specific facts and explanations of general theorems for the purposes of evaluating the plausibility of proof chauvinism. One might think, for instance, that we can often explain a specific fact just by citing a general theorem under which it falls. But when it comes to explaining a general theorem itself, our only recourse is usually to point to some explanatory features of its proof. Hence chauvinism might be false as a thesis about explanations of specific mathematical facts, but at least mostly true as a thesis about explanations of general results.

Maybe something in the neighborhood of this view is right. But it's impossible to say unless we've got some means to tell a specific fact from a general theorem, and I don't know that there's any good way to do this. For, of course, what seems to be a general result in one context often looks like a narrow special case from another viewpoint. Consider, for instance, the fact that all circles in the plane intersect the x-axis at most twice. Is it specific or general? On the one hand, it's about *all* circles, not any one in particular. That seems to indicate generality. On the other hand, the result about circles can be viewed as a very special case of the Fundamental Theorem of Algebra. (And for that matter, the FTA is a particular case of Bézout's theorem, which in turn 'has many strengthenings, generalisations, and variants' ([Tao 2012]).) This sort of case seems very ordinary. So I seriously doubt that enough sense can be made of the specific/general distinction for it to do useful work in clarifying the scope of proof chauvinism.

> assuming $P \neq NP$, this approximation cannot be improved. Why can't we do better than $\ln n$? Perhaps because $P \neq NP$. If this was the only example it would not be compelling. But there are many such pairs where assuming $P \neq NP$ would explain why we have approached these limits. ([Gasarch 2014], p. 258)

This sort of case makes it hard to see how chauvinism, or Lange's account in particular, could be right. We have no proof of $P \neq NP$. Nor does anyone have much idea what a proof of $P \neq NP$ might look like. And yet it seems perfectly natural to say that $P \neq NP$ would explain all sorts of results in computability theory if it turned out to be true.

Of course, this particular example doesn't involve a case of explanation-by-subsumption, but similar remarks apply to cases that do. Consider for instance the famous Hadwiger conjecture in graph theory, which says that, for all $t \geq 0$, a graph is t-colourable if it has no K_{t+1} minor.[9] The conjecture is equivalent to the four-colour theorem when $t = 4$, but the general case is considered extremely difficult, and again nobody has much idea what to expect from a proof. Nevertheless, many mathematicians feel that the Hadwiger conjecture, if true, would explain the four-colour theorem. As Paul Seymour writes:

> The four-colour conjecture (or theorem as it became in 1976) [. . .] was the central open problem in graph theory for a hundred years; and its proof is still not satisfying, requiring as it does the extensive use of a computer. (Let us call it the 4CT.) We would very much like to know the "real" reason the 4CT is true; what exactly is it about planarity that implies that four colours suffice? [. . .] By the Kuratowski-Wagner theorem, planar graphs are precisely the graphs that do not contain K_5 or $K_{3,3}$ as a minor; so the 4CT says that every graph with no K_5 or $K_{3,3}$ minor is 4-colourable. If we are searching for the "real" reason for the four-colour theorem, then it is natural to exclude K_5 here, because it is not four-colourable; but why are we excluding $K_{3,3}$? What if we just exclude K_5, are all graphs with no K_5 minor four-colourable? And does the analogous statement hold if we change K_5 to K_{t+1} and four-colouring to t-colouring? That conjecture was posed by Hadwiger in 1943 and is still open. ([Seymour 2016], pp. 417-8.)

[9] Definitions: A graph G is t-colourable if, given a collection of t distinct colours, G's vertices can be assigned colours in such a way that adjacent vertices never share the same colour. A graph H is a minor of a graph G if H can be obtained from G by deleting edges and vertices and by merging adjacent vertices. The notation K_t denotes the complete graph on t vertices (i.e. the graph where every vertex is adjacent to all of the remaining $t-1$ vertices) $K_{m,n}$ denotes the complete bipartite graph on m and n vertices (i.e. the graph consisting of two vertex sets of sizes m and n, where each of the m vertices is adjacent to all and only the n vertices, and vice versa).

Seymour's discussion strongly suggests that, if the Hadwiger conjecture turned out to be true, then it would explain (i.e. give 'the real reason' for) the four-colour theorem. But neither Seymour nor anyone else is in a position to guess whether the conjecture has a 'common proof' in Lange's sense. And even if there turned out to be no such proof, there's no reason to think this would change anybody's assessment of the situation.

So it seems that the question whether a Langean common proof exists is irrelevant to the question whether the conjecture would be explanatory. Moreover, this is precisely the type of case that Lange's account is designed to handle, since the four-colour theorem is a particular case of the Hadwiger conjecture. Hence Lange's account appears not to work even for cases of explanation by subsumption under a theorem.

I conclude that the verdict reached earlier was the right one. Many theorems are explanatory in a way that doesn't involve or depend on their having explanatory proofs.

So the first strategy for defending proof chauvinism seems to fail. Mathematical practice offers no support for the claim that explanatory proofs are at the centre of most instances of mathematical explanation. On the contrary, mathematicians regularly identify particular facts (as well as things like theories and diagrams) as *explanantia*, and they do so in ways that leave proofs entirely out of the picture.

2.1.2 Proof chauvinism and philosophy

2.1.2.1 An argument from philosophy of science

I turn now to the second possible strategy for defending chauvinism. According to this line of thought, there are compelling philosophical reasons to think that proofs must be central to mathematical explanation—reasons that don't directly involve mathematical explanation-seeking practices, but which hinge on general considerations about the nature of explanation, the role of proofs in mathematics, and the like. Here I'll focus on two fairly natural ways that such an argument might go.

The first argument appeals to the history of philosophy of science. As everyone knows, a view according to which explanations are arguments is defended by [Hempel & Oppenheim 1948] and its sequels. (More specifically—according to the 'deductive-nomological' version of the theory, which is the version that's relevant here—an explanation is a deductive argument whose premises are antecedent conditions and general laws, and whose conclusion is the fact to be explained.) If one wanted to adapt this sort of view to the mathematical setting, the natural thing would be to identify explanations with proofs. And why not give this a shot? The Hempel-Oppenheim theory was enormously popular in philosophy of science for decades, after all, and in some ways it seems even more promising as an account of mathematical explanation.

I'll consider the merits of this idea shortly. But first, why do I say that the explanations-as-arguments view seems better off in the mathematical setting? Well, recall the main objections that eventually led many philosophers of science to abandon the view. Salmon's philosophical history *Four Decades of Scientific Explanation* ([Salmon 1989]) describes three such problems. The first is that irrelevant information is 'harmless to arguments but fatal to explanations' (p. 102)—hence cases like that of 'John Jones' birth control pills' are fine as deductive inferences, but unacceptable as pieces of explanatory reasoning. The second problem is that the explanations-as-arguments paradigm seems unable to account for explanations of low-probability events. The third problem is that legitimate explanations exhibit 'temporal asymmetry'—meaning that the events or conditions appearing in the *explanans* always occur before those appearing in the *explanandum*—whereas arguments are subject to no such constraint. Salmon's preferred solution to these problems, which has since become thoroughly mainstream, is to view scientific explanation as a matter of identifying causes rather than simply giving deductively valid arguments.

In the mathematical context, however, things look quite different. Salmon's second and third problems are no trouble at all for the view that mathematical explanations are arguments (that is, proofs), since the relevant issues about probability and temporality simply don't arise. The first problem is a little more complicated, but perhaps one could say that arguments containing irrelevant information don't qualify as fully satisfactory proofs in the first place. (A good proof, after all, is a piece of reasoning that actually is or would be accepted by the mathematical community, and mathematicians aren't fond of arguments packed with irrelevancies.) In any case, even if we had doubts about the adequacy of the explanations-as-arguments model in mathematics, there's no obvious move available corresponding to Salmon's turn to causation. One could perhaps try to replace causal relations with some species of metaphysical grounding or dependence, but it's not at all clear how to do this in a plausible way. ([Pincock 2015] explores a view of this sort, but I think the proposal can be shown not to work.)

Prima facie, then, the Hempel-Oppenheim approach may look pretty promising as an account of mathematical explanation. And since it identifies explanations with arguments, such an approach would vindicate a type of proof chauvinism. Let me now explain why we shouldn't be convinced by this line of thought.

The first thing to note is that the Hempel-Oppenheim model isn't primarily motivated by the idea that explanations are arguments of some sort. Hempel and Oppenheim didn't start with this conviction, and then proceed to investigate what sort of argument an explanation might be. The real motivation for the view is rather the 'covering law' conception of explanation. This is the thought that

> [an event] is explained by subsuming it under general laws, i.e., by showing that it occurred in accordance with those laws, by virtue of the realization of certain specified antecedent conditions. [. . .] Thus [. . .] the question "*Why* does this phenomenon happen?" is construed as meaning "according to what general laws, and by virtue of what antecedent conditions does the phenomenon occur?" ([Hempel & Oppenheim 1948], p. 136).

The identification of explanations with (deductive-nomological) arguments is offered, then, as a way of making the covering law picture more precise. But the latter is really the heart of the view. Why is this point relevant to the chauvinism issue? In the first place, because the covering law conception has no plausibility whatsoever as a theory of mathematical explanation. Even if we limit our attention to proofs—and even if we assume we can make sense of 'antecedent conditions' and 'general laws' in mathematics—the covering law picture gives plainly wrong conditions for explanatoriness. Subsumption under a law isn't necessary, because many uncontroversially explanatory proofs just don't have this form. Consider for instance the proof of Desargues' theorem discussed in [Lange 2015], whose explanatoriness seems rather to lie in its 'exploit[ing] [a] feature of the given case that is similar to the remarkable feature of the theorem' (p. 438). (One could also consider the many explanatory proofs that succeed by transferring a problem to a geometric setting, or by making it otherwise visualizable.) And subsumption isn't sufficient, either. If it were, then countless proofs like the following would presumably be explanatory: 'All odd numbers are divisible by some prime number; 9 is odd; hence 9 is divisible by some prime.'[10] I take it we'd like to avoid this result.

Perhaps contrary to first appearances, then, the motivation for the Hempel-Oppenheim view fails to transfer in any meaningful way to the mathematical setting. The 'covering law' conception of explanation, which the Hempel-Oppenheim theory is supposed to formalize, is a non-starter as an account of explanatory proof. And the fault doesn't lie just with the deductive-nomological style of argument in particular. In fact, there seems to be no interesting formal structure of any kind that's common to all explanatory proofs. So nothing even in the spirit of the Hempel-Oppenheim view looks promising.

For these reasons, even if such a view was wildly successful as a theory of scientific explanation, this would provide no obvious support for the notion that mathematical explanations are proofs. But of course the Hempel-Oppenheim view is far from wildly successful. The current consensus is rather that it 'jams an alien picture of explanation onto the scientific phenomena [. . .] fails to fit swathes of explanations in the sciences, and has been consequently widely rejected in the philosophy of science as the truth about scientific explanation' ([Gillett 2016], p. 48). Indeed, it's

[10] I don't claim to know what a 'general law' of mathematics is, or what it would be if there were such things. One plausible idea is something like this: a general law is a true universal statement in which only sufficiently natural mathematical predicates appear. Since oddness and divisibility by a prime seem quite natural, the example I give would seem to involve a general law by this criterion.

Another approach, suggested by an anonymous referee, is to identify general laws with axioms. This is an appealing idea, since axioms, like laws, are a special class of (typically general) truths that play a distinguished foundational and organizational role in the theories to which they belong. But this approach has problems too. One issue, noted by the referee, is that *all* proofs might be thought to rest ultimately on axioms, and hence to have the form of covering-law arguments. If this were right, then a covering-law theory of mathematical explanation would (wrongly) predict that all proofs are explanatory. I'm not so sure myself that all proofs should be viewed as proofs from axioms, but I do doubt that there's any good principled way to distinguish the axiom-based proofs from the rest. Another obvious worry is that the distinction between axioms and non-axioms is relative to a choice of axiomatization. Since lots of mathematical theories can be axiomatized in more than one way, whether or not a given proof is explanatory would often turn out to depend on which statements we happened to be taking as axioms. This seems quite counterintuitive.

worth remarking that—along with the objections due to Salmon described above—a longstanding complaint is that the deductive-nomological picture wrongly precludes singular facts from serving as *explanantia*. As Michael Scriven wrote more than fifty years ago:

> If you reach for a cigarette and in doing so knock over an ink bottle which then spills onto the floor, you are in an excellent position to explain to your wife [. . .] why the carpet is stained (if you cannot clean it off fast enough). You knocked the ink bottle over. This is the explanation of the state of affairs in question, and there is no nonsense about it being in doubt because you cannot quote the laws that are involved [. . .] ([Scriven 1962], p. 68)

I think this is right, and I think the same point is well taken in the mathematical setting. Not all legitimate explanations, whether mathematical or scientific, involve the deduction of a special case from a general law. In light of these issues, it's very hard to see how the Hempelian paradigm could provide the proof chauvinist with any encouragement.

2.1.2.2 An argument from philosophy of mathematics

I now turn to the second philosophical argument for chauvinism. The argument comes from philosophy of mathematics, and it appeals to the supposed epistemological primacy of proofs (as compared to axioms, theorems and the like). This type of proof-centred viewpoint is defended with vigor by [Rav 1999]. Rav writes, for instance, that

> [*P*]*roofs rather than the statement-form of theorems are the bearers of mathematical knowledge*. Theorems are in a sense just tags, labels for proofs, summaries of information, headlines of news, editorial devices. The whole arsenal of mathematical methodologies, concepts, strategies and techniques for solving problems, the establishment of interconnections between theories, the systematisation of results—the entire mathematical know-how is embedded in proofs. (p. 20; emphasis in original)

On this sort of picture, chauvinism looks almost inevitable, since a mere theorem is presumably too slender a reed to bear the weight of being explanatory. Only proofs seem epistemologically robust enough to handle the job.

In defense of his view, Rav points out that some theorems (or conjectures) have generated proofs (or proof-attempts) which are much more interesting, useful and deep than the original statements themselves, such as Fermat's last theorem and the Goldbach conjecture. Similarly, there are statements like the Continuum Hypothesis that turned out to be independent of the relevant axioms (and hence neither provable nor disprovable), but whose would-be proofs led to important kinds of progress.

True enough, but such examples hardly vindicate Rav's view. All that these cases show by themselves is that *some* theorems are less epistemologically valuable than their proofs. It's quite another thing to claim that theorems in general occupy a lower rung on the epistemological ladder than do proofs in general. And indeed there's plenty of reason to deny this. Mathematicians routinely identify theorems as profound, beautiful, powerful, deep, essential, and the like. High praise for mere tags! And just as some run-of-the-mill theorems have interesting proofs, the reverse is also true. The Fundamental Theorem of Arithmetic, for instance, is a result worthy of its name: the unique prime factorization property of the natural numbers is a highly nontrivial feature[11] that provides a foundation for much of what goes on in number theory. But the proof of FTA doesn't seem noteworthy in any way. At least, it isn't especially deep or revelatory, and its development didn't lead to the invention of any groundbreaking new machinery: FTA could have been (and practically was) proved by Euclid, using just the modest number-theoretic resources at his disposal. (Euclid's lemma, which is *Elements* VII.30, is the key ingredient of the proof.) Plenty of other results fit the same pattern. One could mention, for instance, Ramsey's theorem in combinatorics, the intermediate value theorem in analysis, the binomial theorem in algebra, Brouwer's fixed-point theorem in topology, and the five lemma in category theory as results which are more important and interesting than their proofs. (See also the discussion of Morley's triangle theorem in [Lange 2016], pp. 252-3; this result has been called 'spectacular', 'amazing', and 'the marvel of marvels', but its known proofs are all notably unenlightening.)

So I don't think we can get to proof chauvinism by way of a Rav-style view about the epistemological primacy of proofs. Rav is right to insist on proof's distinctive value, and to remind us that

[11] For an interesting discussion of the non-triviality and non-obviousness of FTA, see ([Gowers 2011]).

mathematics wouldn't get very far without the creative stimulus provided by the search for new problem-solving methods. But it doesn't follow that proofs are the only things worth knowing. Far from being mere tags or labels, many theorems also have great epistemological significance in their own right. Indeed, it's puzzling why anyone would go to the trouble of devising proofs in the first place if the statements to be proved had no intrinsic intellectual worth. This thought is corroborated by the section on the goals of mathematical research in the recent *Princeton Companion to Mathematics*:

> The object of a typical paper is to establish *mathematical statements*. Sometimes this is an end in itself: for example, the justification for the paper may be that it proves a conjecture that has been open for twenty years. Sometimes the mathematical statements are established in the service of a wider aim, such as helping to explain a mathematical phenomenon that is poorly understood. But either way, mathematical statements are the main currency of mathematics. ([Gowers et al. 2008], p. 73; emphasis in original)

Indeed. Theorems are epistemologically valuable; more specifically, they have explanatory value, as I've been arguing.

2.2 An Example: Galois Theory and Explanatory Proof

I've now finished making the case that proof chauvinism is false. But it's not yet clear that its falsity should be any great concern. It would be useful to know how, if at all, chauvinist assumptions have actually done harm in the theory of mathematical explanation, and how they might do more in the future.

In order to get clearer about the perils of chauvinism, then, let me now consider a particular case in somewhat more detail. One of the standard examples in the literature has long been Évariste Galois's explanation of the solvability of polynomial equations. The case deserves the attention it gets; Galois's work was a major explanatory achievement, and a watershed moment in the history of mathematics. I'll argue, however, that explanatory proof plays no role in this case. What's more, the prejudice in favor of identifying mathematical explanation with explanatory proof has led to some confusions about Galois's work.

First, a brief review of what Galois theory accomplished. The oldest and most famous problem in the theory of equations involves finding the set of solutions to a given equation 'in radicals', that is, as a function of the equation's coefficients involving only basic arithmetical operations and nth roots. (The quadratic equation $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$ is a familiar example of such a solution formula.) Solving a linear equation $ax+b=0$ in radicals is easy; the single root is given by $x = -\frac{b}{a}$. The quadratic formula, or something essentially equivalent, was known already in antiquity. But the cubic and quartic cases are much harder. General solution formulas for these equations were found only in the sixteenth century, by a laborious trial-and-error process involving complicated substitutions.

Following these discoveries, mathematicians sought to understand whether and why an arbitrary polynomial (of any degree) is solvable in radicals. After centuries of limited progress, Niels Henrik Abel proved the unsolvability of the quintic in 1824, and Galois put the general problem to rest shortly thereafter. The approach he developed, now known as Galois theory, involves the study of a certain algebraic object associated with a given polynomial (namely its 'Galois group', a group of permutations of the roots). Galois's major breakthrough was the discovery that, if the Galois group of a polynomial has a certain property (also known as 'solvability'[12]), then the polynomial is solvable in radicals; if not, then no solution formula exists. This result is 'Galois's criterion for solvability'.

To obtain a complete picture of which equations are solvable and which aren't, Galois combined this criterion with two other results. The first is that the Galois group of an nth-degree polynomial is a subgroup of the 'symmetric group' S_n, i.e. the group of all permutations of an n-element set. (Moreover, S_n itself is the Galois group of the 'general' polynomial of degree n.) The second fact is that S_n and all its subgroups are solvable for $n \leq 4$, but not for $n \geq 5$. Together, these facts imply

[12]A group G is solvable if it has subgroups $\{1\} = G_0 < G_1 < \cdots < G_k = G$ such that G_{j-1} is normal in G_j, and G_j/G_{j-1} is a commutative group, for $j = 1, \ldots, k$.

that an arbitrary equation of degree ≤ 4 is always solvable in radicals, but not so for equations of higher degrees.

Many mathematicians and philosophers view Galois's accomplishment as an important case of mathematical explanation. And so it is. Unfortunately, the chauvinist preoccupation with explanatory proof has led to some mistakes in the philosophical understanding of the case. For instance, [Pincock 2015] presents a theory of what he calls 'abstract mathematical explanation' motivated largely by Galois theory. I won't go into the details of Pincock's account here. For now I only want to note that, although little is said to justify either of these assumptions, Pincock takes it for granted that the case involves an explanatory proof, and he fixes on Galois's proof of the unsolvability of the quintic as the relevant item. (Pincock then suggests that this proof is a paradigmatic example of an abstract mathematical explanation, that is, a proof that explains 'by appeal to an entity that is more abstract than the subject-matter or topic of the theorem' ([Pincock 2015], p. 2).)

I think this is a mistake. Galois theory isn't explanatorily successful by virtue of containing an explanatory proof. Instead, I claim that the *explanantia* in this case are theorems, notably Galois's criterion itself and his results about the solvability of symmetric groups.

There are a couple of reasons to think so. For one, my proposal reflects how mathematicians actually talk about the case—they simply don't ever, as far as I can tell, mention any proof (or any aspect of any proof) in their discussions of Galois's explanatory achievements. Instead, they identify Galois's theorems, or sometimes Galois theory as a whole, as the bearers of explanatory value. The following sorts of remarks are typical:

- '[Galois] showed that the group of the general nth degree equations is the symmetric group S_n and that solvability of an equation by radicals is reflected in a group-theoretic "solvability" property, possessed by S_3 and S_4 but *not* by S_5. This explains why the general cubic and quartic are solvable by radicals and the general quintic is not' ([Stillwell 1995], p. 58).
- 'The group of the general equation of the nth degree is the symmetric group [S_n] [. . .] whose factors of composition... [for $n = 3$] are 2, 3; for $n = 4$ they are 2, 3, 2, 2... all of them prime; this is the reason why the general equations of the third and fourth degree are solvable by radicals' ([Bolza 1891], pp. 138-139).[13]

[13] The condition that a group G is solvable is equivalent to the condition that G's composition factors all have prime order. Earlier presentations of Galois theory, including e.g. Jordan's influential *Traité des Substitutions*, tended to use the composition factor condition.

- 'In Sect. 2.1, we show that if the group G is solvable, then the elements of the field P are representable by radicals through the elements of the invariant subfield K of G [. . .] When P is the field of rational functions of n variables, G is the symmetric group acting by permutations of the variables, and K is the subfield of symmetric functions of n variables, this result provides an explanation for the fact that algebraic equations of degrees 2, 3 and 4 in one variable are solvable by radicals' ([Khovanskii 2014], p. 48).

- 'An algebraic equation of the form $[a_n x^n + \cdots + a_1 x + a_0 = 0]$ is solvable by radicals if and only if its Galois group is solvable [. . .] In view of [this theorem], the nonsolvability by radicals of the general polynomial equation of degree ≥ 5 is then explained by the following result [. . .] *The symmetric group S_n is solvable if and only if $n < 5$*' ([Pantea et al. 2014], p. 425; emphasis in original)

- 'The quadratic formula for polynomials of second degree [. . .] holds that an equation of the form $ax^2 + bx + c = 0$ has two solutions [. . .] Galois theory explains why there is no comparable formula for equations of degree 5 and higher' ([Nadis & Yau 2013], p. 170).

So the mathematical literature provides no support for a proof-chauvinist interpretation of the Galois theory case. But it provides ample support for the idea that Galois's theorems are explanatory. (Pincock claims that there's 'considerable evidence that practitioners consider [Galois's unsolvability proof] to be more explanatory than its predecessors' ([Pincock 2015], p. 1), but none of the sources he cites mentions anything about the proof; all of them talk only about Galois's ideas or Galois theory in general.)

In any case, the suggestion that a proof was Galois's main explanatory contribution is problematic by itself. For one, as Leo Corry observes, 'Galois's writings were highly obscure and difficult, and his proofs contained many gaps that needed to be filled' ([Corry 2004], p. 26). Presumably, an argument as problematic and incomplete as the one Galois actually gave is a poor candidate for an explanatory proof. Presumably, also, this is why Pincock presents a modern field-theoretic version of the proof of Galois's criterion, even though the reasoning is substantially different from what Galois himself could have had in mind.

But this line of thought leads in some troubling directions. If we accept chauvinism, and if we also admit that Galois's original proof was unexplanatory, then we apparently have to say that Galois's work contained little or nothing of explanatory value. This is an unappealing view, and I doubt that Pincock would want to accept it. On the other hand, if we're willing to bite the bullet and deny that Galois explained anything, it follows that a genuinely explanatory proof of Galois's criterion appeared at some point after the 1830s and before the present day. Which version of the proof was this, exactly? (Or, if we think of explanatoriness as a matter of degree, what was the first reasonably explanatory version of the proof?) Although perhaps not unanswerable, these questions are at least uncomfortable, and Pincock says little to suggest how he might either respond to or avoid them.

By contrast, the idea that Galois's achievement was to prove a number of explanatory theorems suffers from none of these difficulties. For one, as we've seen, this proposal is squarely in line with what mathematicians actually say about the case. Those who identify specific *explanantia* almost invariably point to Galois's results, rather than his proofs—in particular, they identify Galois's criterion, and the theorems about the solvability of symmetric groups, as the bearers of explanatory value. Additionally, the view vindicates the consensus opinion that Galois achieved something of explanatory significance. After all, Galois was certainly the first to conceive of and state the theorems in question, but it's much less clear that he discovered adequate (let alone explanatory) proofs for these theorems.

Let me conclude by considering another suggestion in the chauvinist spirit, due to Lange. Like Pincock, Lange thinks that Galois theory explains by virtue of containing an explanatory proof. He writes:

> I suggest that the distinction between an explanation and a mere proof of the quintic's unsolvability is grounded in this result's most striking feature: that in being unsolvable, the quintic differs from all lower-order polynomial equations (the quartic, cubic, quadratic, and linear), every one of which is solvable. In the context of this salient difference, we can ask why the [quintic] is unsolvable. Accordingly, an explanation is a proof that traces the quintic's unsolvability to some other respect in which the quintic differs from lower-order polynomials. Therefore, Galois's proof (unlike Abel's, for example) explains the result. ([Lange 2016], p. 440)

The main novel claim here is that Galois's proof of the unsolvability of the quintic, though not Abel's, works by identifying and exploiting some difference between fifth-degree and lower-degree equations. Hence only Galois's proof is explanatory. This is, I think, mistaken. I can't discuss Abel's proof in detail here—for that, see [Pesic 2013]—but the general idea is as follows. First

Abel shows that any solution formula for an equation of degree m must have the form

$$x = p + R^{\frac{1}{m}} + p_2 R^{\frac{2}{m}} + \cdots + p_{m-1} R^{\frac{m-1}{m}},$$

where '$p, p_2, \ldots$ are finite sums of radicals and polynomials and $R^{\frac{1}{m}}$ is in general an irrational function of the coefficients of the original equation' ([Pesic 2013], p. 90). (The already-known formulas for lower-degree equations can easily be seen to have this form.) Hence, if there were a solution formula for the general quintic, it would look like

$$x = p + R^{\frac{1}{5}} + p_2 R^{\frac{2}{5}} + p_3 R^{\frac{3}{5}} + p_4 R^{\frac{4}{5}}.$$

The rest of the proof shows that this can't occur. Abel's main tool here is a theorem of Cauchy, which implies that $R^{\frac{1}{m}}$ can have only two or five distinct values when the roots of the quintic are permuted. Abel shows that each of these possibilities leads to a contradiction; it follows that the general quintic is unsolvable.

Pace Lange, this argument strikes me as 'a proof that traces the quintic's unsolvability to some other respect in which the quintic differs from lower-order polynomials'. The respect in question is the satisfiability of the equation $x = p + R^{\frac{1}{m}} + p_2 R^{\frac{2}{m}} + \cdots + p_{m-1} R^{\frac{m-1}{m}}$ for $m < 5$, as compared to its satisfiability for $m \geq 5$. By Lange's criterion, then, it's unclear why Galois's proof should count as any more explanatory than Abel's.

2.3 Conclusion

The Galois theory case is interesting and important in its own right, and I think the above observations would be worth making even if nothing further was at stake. But in fact proof chauvinism has some more general philosophical risks. Let me close by mentioning a couple of them.

First, theorizing about mathematical explanation under the influence of chauvinism is likely to lead to a distorted picture of the phenomena. If one is bent on finding explanatory proofs, and there are no such proofs on offer in a particular case, then one is liable to conclude that the case contains nothing of explanatory interest. But if chauvinism is false and many mathematical explanations don't involve explanatory proofs, this approach will yield frequent false negatives. This is especially worrisome in a field in which progress still largely depends on assembling a representative stock of examples. As Mancosu suggests, the theory of mathematical explanation will advance only when we've managed to 'carefully analyze more case studies in order to get a better grasp of the varieties of mathematical explanations' ([Mancosu 2015]).

Second, proof chauvinism threatens to draw attention away from the broader epistemological contributions of other elements of mathematics. Explanation, after all, is closely linked to understanding, theory change, conceptual development, various cognitive benefits, and other phenomena of epistemological interest. So if one is convinced that proofs alone are explanatory, then one is a short step away from concluding (mistakenly) that only proofs play a meaningful role in these other phenomena.

The theory of mathematical explanation is still in its infancy. But infancy is a crucial developmental phase, and an early mishap can lead to more serious problems later on. The philosophy of mathematics would do well, then, to inoculate itself against an unhealthy fixation on explanatory proofs.

Chapter 3

Viewing-as and Dependence in Mathematical Explanation

3.1 Introduction

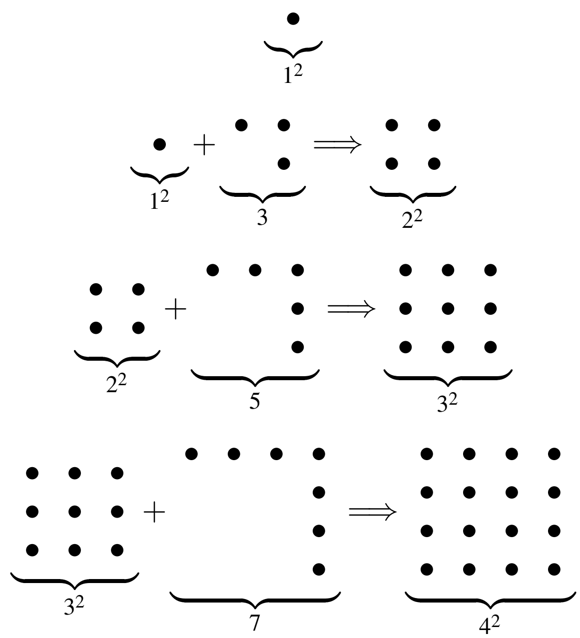
Suppose you want to show that the sum of the first n odd natural numbers is n^2. One approach is to use mathematical induction: clearly the statement is true for $n = 1$, and applying the induction hypothesis gives

$$\begin{aligned}\sum_{k=1}^{n+1}(2k-1) &= \sum_{k=1}^{n}(2k-1) + 2(n+1) - 1 \\ &= n^2 + 2n + 1 \\ &= (n+1)^2,\end{aligned}$$

as needed.

This gets the job done, logically speaking. According to the prevailing view among mathematicians and philosophers, though, inductive proofs like this one are generally unexplanatory (cf. [Lange 2009]). So if you wanted to understand *why* the sum of the first n odd numbers is n^2, you might keep looking for a more satisfying argument.

Here's a different sort of proof, the key idea of which is to *view numbers as arrangements of dots*. Start with 1, the first odd number, which can be regarded as a square array of side 1 (and hence of area 1^2). The second odd number is 3, and we can think of adding 3 to 1 as augmenting the original square array so as to make a new one of side 2 (and hence of area 2^2). Similarly, adding 5 to $1+3$ gives a square array of side 3, and so on, as in the diagram.



It's easy to see that the pattern will hold in general—for any $n-1$, the sum of the first $n-1$ odd numbers corresponds to a square array of side $n-1$ and area $(n-1)^2$, and adding the next odd number can be viewed as augmenting this array to produce a new one of side n and area n^2. Hence the diagram shows that the sum of the first n odd natural numbers is n^2. What's more, this sort of proof is often claimed to explain why the identity holds.[1]

The style of reasoning on display here is familiar, and the particular case I've used to illustrate it is simple enough. But closer examination reveals some challenging philosophical issues. For one, note that the proof turns on the idea of "viewing one mathematical object as another". Mathematicians engage in this activity all the time, but what exactly does it involve? Some cases, including the one just given, really do feature "viewing" in the relatively literal sense of perception or visualization, and so one might think that viewing-as is essentially pictorial. But not all cases are like this. (More on this point below.) So how in general should we understand the phenomenon?

[1]For instance, [Gullberg 1997] gives a similar dot-diagram proof that the sum of the first n natural numbers is $n(n+1)/2$, and claims that "the figure shows why" the identity holds (289). [Hanna 1990] contrasts the inductive proof of this identity with the dot-diagram proof, claiming that the latter but not the former is explanatory (10-11). See chapter 8 of [Giaquinto 2007] for a discussion of the epistemology of dot-diagram arguments, including a thorough defense of the claim that the images used in such arguments count as genuine proofs.

Another important issue has to do with mathematical explanation. Viewing X as Y is often a way to explain something about X, as in the above case: by viewing numbers as arrangements of dots, we're put in a position to explain why the sum of the first n odd numbers is n^2. And this identity seemed hard to account for otherwise. (Interestingly, this type of mathematical explanation was noticed already in Mark Steiner's pioneering 1978 paper—he mentions "Hardy's explanation of the lawless behavior of a certain numerical function [obtained] by regarding it, a la Ramanujan, as a 'snapshot' at each n of the resultant of infinitely many sine waves of incompatible periods and decreasing amplitudes" ([Steiner 1978a], 148). Unfortunately Steiner doesn't discuss this case, or the general phenomenon, in detail.)

Such "viewing-as explanations" seem rather mysterious at first. After all, many philosophers hold that explanatory relationships are always grounded in worldly (physical or metaphysical) dependence relations between the relevant objects—this is part of the currently ascendant "ontic conception" of explanation. But that doesn't seem to be the case here. Numbers, after all, apparently aren't arrays of dots, nor are they composed of them, constituted by them or causally influenced by them. (The two classes of object are similar in certain ways, but mere similarity isn't an explanatory relation.) So how could an observation about arrays of dots possibly explain a theorem about numbers? In general, if there are explanatory benefits to viewing X as Y, what sort of relationship must X and Y stand in to one another?

Of course, questions about the metaphysical underpinnings of mathematical explanation aren't limited to cases of viewing-as. But such cases illustrate the problem in particularly stark terms. After all, in many other cases, the relevant objects belong to the same mathematical domain, and the explanans-objects are apparently more fundamental. For instance, we can explain properties of natural numbers in terms of properties of their prime factors, and this is plausibly because all numbers are "composed of" (or somehow determined by) primes. (It's a common thought that "primes are the atoms of number theory, the basic indivisible entities of which all numbers are made" ([Ellenberg 2014], 139).) Likewise for a finite group and its simple composition factors, or a set and its elements. Examples like these arguably fit the standard ontic picture of explanation, since the explanatory relations on display seem to be underwritten by dependence or determination relations of some sort. Viewing-as cases, by contrast, typically don't feature anything that looks like dependence. So it's unclear whether and how such cases are compatible with the ontic picture. This paper aims to shed some light on these issues. I begin by giving some examples of viewing-as, and extracting some general lessons about it, in §2. In §3 I introduce the notion of frames. Drawing on recent work of Elisabeth Camp on framing phenomena in science, I suggest that viewing-as is usefully understood as a type of frame-oriented cognition. §4 turns to the relationship between viewing-as and mathematical explanation. I again start by discussing cases. Finally, I consider whether such explanations might be grounded in underlying metaphysical relations of determination or dependence. Since the prospects for such an account look dim, I conclude by sketching an alternative "cognitivist" approach in §5 that makes use of the frame-theoretic viewpoint developed earlier.

A couple of notes on terminology. I'll call X the *source* object and Y the *target* object when X is being viewed as Y. Also, mathematical writers seem to use the phrases "view X as Y", "regard X as Y", "think of X as Y", and "interpret X as Y"[2] more or less interchangeably. So I'll do the same here.

Note, however, that "seeing X as Y" traditionally has a different, distinctively perceptual meaning. Wittgenstein's duck-rabbit[3] provides the classic case of seeing-as—one can perceive the figure in the aspect of one animal or the other, but not both at once—and the literature on this phenomenon has centered on questions of attention, perceptual content and the like.[4] Although seeing-as and viewing-as are similar in some respects, and although seeing-as certainly has its place in mathematical cognition[5], I take the two phenomena to be importantly distinct. I only intend to discuss viewing-as here.

[2]Perhaps other expressions are used in a similar way, but these seem to be the most common.

[3]See Part II, section xi of [Wittgenstein 2009].

[4]For recent work on seeing-as, see for instance [Day & Krebs 2010], [Kemp & Mras 2016] and [Beaney et al. forthcoming]. Although the usage I've mentioned is standard, seeing-as is sometimes taken to encompass non-perceptual "perspectival" phenomena like metaphor, as in [Camp 2008].

[5]See for instance [Coliva 2012] on the role of seeing-as in reasoning with Euclidean diagrams, and [Floyd 2010] on Wittgenstein's ideas about seeing-as in logic, probability and other mathematical subjects.

3.2 The viewing-as phenomenon

Before getting to questions of general theory, let me offer a few more examples. Doing so should be useful in two ways. First, it will help illustrate the scope and interest of the phenomenon (and hence allay possible skepticism on this front). Second, it'll bring out some characteristic features of viewing-as that will be worth keeping in mind for later.

An obvious but important first point is that a given object can generally be thought of in a number of different ways, each of which might be advantageous for some purposes but not others. For example, Terence Tao remarks that "we can think of the number $e^{i\theta}$ both as a point in the [unit] circle and as a rotation through an angle of θ" ([Tao 2008], 207)—the former viewpoint is natural when we're doing complex analysis, or for other applications that make it convenient to visualize complex numbers as points in the plane, while the latter is more appropriate for geometric or topological purposes. Some objects, in fact, are noteworthy for the number and variety of interesting ways in which it's possible to view them. Consider ultrafilters. The standard axiomatic definition describes an ultrafilter on a set X as a family $\mathscr{U}$ of subsets of X that meets certain conditions. But an assortment of other interpretations are available in different contexts:

> Another way to view an ultrafilter $\mathscr{U}$ on X is as a map from the set 2^X to X, or more generally $M^X \to M$ for any finite set M, namely the map sending any $f : X \to M$ to the unique $m \in M$ for which $f^{-1}(\{m\}) \in \mathscr{U}$. ...A third way to view ultrafilters is as new quantifiers by defining $(\mathscr{U}x)\,\varphi(x) \Longleftrightarrow \{x \in X : \varphi(x)\} \in \mathscr{U}$. The quantifier $(\mathscr{U}x)$ can be read as "for almost all x (with respect to $\mathscr{U}$)" or "for $\mathscr{U}$-most x". ...A fourth view of ultrafilters is as uniform ways to choose limits of sequences. ...A fifth view is that an ultrafilter on X is a point in the Stone-Čech compactification βX of the discrete space X. ([Blass & Rupprecht 2015], 2)

It follows—at least on any reasonably standard views about the ontology of mathematics and the metaphysics of identity—that *viewing* X as Y doesn't entail taking X to *be* Y. Granted, issues surrounding isomorphism and other forms of equivalence sometimes make it hard to say whether two mathematical objects are identical or distinct. But if we can ever make any confident assertions on this score, then surely we can say that a map of sets, a quantifier, and a point in a topological space (for example) aren't all one and the same entity. Even a mathematical structuralist of the sort that's committed to identifying isomorphic structures would be hard pressed to disagree, since there are no obvious isomorphisms to appeal to in many such cases.

A second point, alluded to above, is that viewing-as needn't involve perception or visualization in any significant way. Set theorists often view the natural numbers as ordinals via von Neumann's pairing $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\}, \ldots$, but this correspondence doesn't obviously invite any particular visual representation.[6] What's more, even the cases that do involve visualization aren't all cut from the same cloth. It's often useful, of course, to view an object that's hard to picture as one more accessible to geometric or visual intuition. The example from the beginning of the paper, in which the sum of the first n odd numbers was interpreted as a square array of dots, is a case in point. But this isn't the only type of example. Sometimes we want to view one geometric object as another, as in the projective proof of Desargues' theorem, when we regard two triangles in the plane as "projections of triangles jutting into the third dimension" ([Lange 2015], 443). Here it's easy to picture both the original and target objects, so viewing the former as the latter doesn't involve an increase in visualizability. Finally, viewing-as sometimes takes us from an object that's easy to visualize to one that resists pictorial representation. For instance, it's common in topological contexts to regard the unit circle as the quotient group $\mathbb{R}/\mathbb{Z}$ (see e.g. [Devaney 1994], 8), although the latter is much less readily visualizable than the former. And geometers since Grothendieck have found it useful to view algebraic varieties as schemes, even though the "points" comprising a scheme "have no ready to hand geometric sense" ([Deligne 1998], 12, quoted and translated in [McLarty 2008], 319). (Similarly, Joe Harris notes that scheme theory and other recent developments have brought "a gradual but consistent trade-off of naive geometric intuition for a formal unity", which has been "met with cries of, 'It may be a pretty theory, but it's not geometry!'" ([Harris 1992], 100).)

[6]The Gödelian view that we have "something like a perception... of the objects of set theory" ([Gödel 1983], 483-4) might seem to suggest otherwise. It's not clear what exactly Gödel meant by this, or whether we indeed have such a faculty. But even if some version of the claim is true, it's not clear that the perceptual modality in question is vision-like, or that all sets are liable to being distinctly perceived in this way. In my own case, at least, even a relatively simple set like the von Neumann ordinal $3 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ calls no particular image to mind.

In sum, then, there's no simple relationship between viewing-as and visualization. It's often possible, and indeed desirable, to view one object as another when either both objects, neither object, or either one object or the other is amenable to visualizing.

A final point is that viewing-as sometimes involves precise correspondences between the relevant objects and properties, while the relationship in other cases is relatively indeterminate or open-ended. When we talk about viewing numbers as sets, for instance, we have in mind a specific system of associations: the number 0 is interpreted as the set $\emptyset$, 1 as $\{\emptyset\}$, 2 as $\{\emptyset, \{\emptyset\}\}$, and so on. (We can do the same for arithmetical functions and relations, of course; this is at least part of what it means to say that arithmetic is reducible to set theory.)

But other cases involve less exact correspondences. One type of inexactness occurs when it isn't possible to precisely characterize or identify one of the relevant objects. The so-called "field with one element" $\mathbb{F}_1$ furnishes an example. Strictly speaking, there is no such field (since the field axioms require 0 and 1 to be distinct elements). As it turns out, though, there are a number of compelling reasons why something like a field with one element "should" exist, and mathematicians in recent decades have sought to identify the key properties that $\mathbb{F}_1$ is hoped to have and to determine which legitimate mathematical object can best play this role. This search is partly motivated by the explanatory work that $\mathbb{F}_1$ is expected to do: for instance, it's been thought that "the analogy between the symmetric group S_n and the Chevalley group $\mathrm{GL}_n\left(\mathbb{F}_q\right)$... should find an explanation by interpreting S_n as a Chevalley group over the 'field of characteristic one'" ([López Peña & Lorscheid 2011], 241). Another, more elementary idea is that "In some sense 'a set is a vector space over a field with one element'" ([Bender & Goldman 1975], 800). The point here is that if $V_n(q)$ is an n-dimensional vector space over the field with q elements, and if we define $\begin{pmatrix} n \\ k \end{pmatrix}_q$ to be the number of k-dimensional subspaces of $V_n(q)$, then we can show that

$$\lim_{q \to 1} \begin{pmatrix} n \\ k \end{pmatrix}_q = \begin{pmatrix} n \\ k \end{pmatrix}.$$

But, of course, the binomial coefficient $\begin{pmatrix} n \\ k \end{pmatrix}$ is the number of k-element subsets of an n-element set. As Bender and Goldman remark, "This is the first manifestation of the relation between subspaces and subsets which is as mysterious as it is fascinating" (800).

In any case, as these examples show, applications of the field with one element often involve viewing a (familiar and well-understood) object as an object defined in terms of $\mathbb{F}_1$—symmetric groups as Chevalley groups over $\mathbb{F}_1$, say, or sets as vector spaces over $\mathbb{F}_1$. But we don't yet have a satisfactory way to say exactly what $\mathbb{F}_1$ is. So, for the time being at least, the correspondences involved in these cases are relatively imprecise and metaphorical.

3.3 Viewing-as and cognitive frames

With the observations from the previous section in hand, I now want to consider how the viewing-as phenomenon might be placed in a broader theoretical context. I think the theory of frames—especially Elisabeth Camp's recent work on the subject—provides a useful starting point.

The notion of frame has a long history in the social sciences, the general idea being that of an interpretative system that structures thought, perception or communication in some domain. Earlier research focused mostly on the (often negative) effects of framing in the social, political and economic spheres.[7] Camp's work, however, is concerned with the varieties of frames deployed in scientific inquiry and the cognitive benefits they frequently confer. Much of what Camp has to say about framing in the empirical sciences is equally plausible in the mathematical setting. And I think the frame-theoretic perspective sheds light on the issue at hand. This section will briefly explain how.

Let me start by saying a bit more about the idea of a frame. As per Camp, frames are "interpreted representational vehicles which provide an overarching principle for interpreting something else"

[7]The framing phenomena at issue here include stereotypes (cf. [Herald 2010]), "spin" in politics and journalism (cf. [Entman 2007]), psychological effects involving people's (often seemingly irrational) sensitivity to differing descriptions of a decision problem (cf. [Tversky & Kahneman 1981]), and the like.

([Camp forthcoming], 26)[8]. Pictorial representations, like diagrams, can serve as frames, as can pieces of language (e.g. slogans and metaphors) and concrete objects (e.g. "telling instances" which are taken as exemplary of their kinds). Given the diversity of kinds of things that can function as frames, it's unsurprising that particular examples differ from one another along various axes. A frame can be "more or less abstract, idealized, detailed, and affectively and experientially loaded" (26)—telling details are literal and concrete, for instance, whereas formal principles are abstract but determinate in content, and metaphors are open-ended and lacking in explicit detail. Another useful distinction is between "internal" and "external" frames. Roughly speaking, an internal frame applied to an object X works by highlighting a feature possessed (or imagined to be possessed) by X itself, whereas an external frame offers some distinct thing as a lens through which to view X. Notwithstanding these differences, all frames have common features, and they play a characteristic role in cognition. For instance, "All frames presuppose a taxonomy, which is necessarily selective and contrastive; [and] all frames determine what matters about their subject, and how" (26).

Analogies, metaphors, models and other frame-assisted ways of thinking are common in science.[9] Why is this? The details will differ from case to case, of course, but the general answer is that frames often structure our thought in beneficial ways—by imposing useful taxonomies, raising important facts to salience, introducing comparisons whose consequences can be pursued in fruitful directions, and so on. And the same is true in mathematics. Mathematical research often involves choices about how to think of objects and problems, and selecting the right frame can be a crucial step in discovering proofs, crafting good definitions, hitting upon insightful conjectures and asking incisive questions.

I want to suggest, in particular, that the practice of viewing one mathematical object as another involves a type of frame-oriented cognition. More precisely, I propose that viewing X as Y amounts to taking Y as an (external) frame for interpreting or thinking about X. Among the virtues of the

[8] Further references in this section are to this paper.

[9] See [Gentner & Jeziorski 1993] for an account of the historical transition away from excessively vague and open-ended scientific metaphors (which were common in alchemy, for instance) and toward more precise analogies.

frame-theoretic viewpoint is that it agrees with, and helps make sense of, the observations about viewing-as from the previous section. Recall that these were

1. That a given source object can typically be viewed in more than one way, with different target objects being appropriate in different contexts;
2. That viewing-as sometimes, but not always, involves depiction or visualization;
3. That the correspondences involved in viewing-as can be either strict and determinate or loose and open-ended to varying degrees.

These data are just what one would expect on the hypothesis that viewing-as involves the application of frames. As for the first point, it's clear that different frames can be usefully brought to bear on a given object in different situations. Just as it's sometimes advantageous to view traffic as a continuous fluid and sometimes as a collection of discrete bodies, or a photon as either a wave or a particle, so it's fruitful in various settings to think of numbers as sets, or arrays of dots, or geometric magnitudes. Note that, on this picture, there needn't be any particular metaphysical relationship between the source and target objects in order for one to be appropriately viewed as the other. We needn't try to figure out, say, how numbers could be identical to or constituted by dot arrays, or how the properties of numbers could be somehow objectively grounded in the properties of dots. Rather, the appropriateness of viewing numbers as dots is a matter of the dot-frame's being able to structure our numerical thinking in a productive way—e.g., by getting us to notice things about sums of odd numbers that wouldn't have been obvious otherwise. In general, when we apply external frames "we do not pretend, even temporarily, that the world really is as the representation literally describes"—i.e., that the source and target objects are identical. "Instead, we leave [the objects] distinct in imagination, and seek to identify respects in which they are alike" (20).

Regarding the second point, it's not surprising that viewing-as can either have or lack a visual component, since frames include both pictorial and non-pictorial kinds of representations. Again, this point holds in the empirical sciences too. Thinking of molecules as ball-and-stick models, or of particle interactions in terms of Feynman diagrams, is one way to organize our thinking about these subjects. But molecular formulas or explicit equations are equally good alternatives in other contexts.

Finally, instances of viewing-as can feature more or less definite correspondences between the source and target objects. As we've seen, the same is true for frames. As Camp uses the terms, *analogies* are a type of external frame that "employ precise, consistent, systematic matches" between the source and target objects (22), while *metaphors* "[rely] on tacit, vague, and otherwise inarticulable intuitions of similarity" (23). The specificity of analogies is obviously useful, but the open-ended quality of metaphors also has a distinctive type of epistemic value. Metaphors "often play a theoretically and empirically fruitful role in scientific inquiry precisely because they stand in need of clarification... This means they can guide attention and suggest hypotheses in epistemic circumstances where a more precise structural analogy would be stymied" (23).

This framework fits our earlier examples well. Cases like the "field with one element", for instance, plausibly involve metaphorical frames—since we can't say exactly what such an object would be, talk of viewing sets as vector spaces over $\mathbb{F}_1$ and the like is to some degree merely suggestive. But the attempt to clarify the metaphor has itself been a source of insight and progress.

I conclude that the theory of frames is an appropriate setting in which to place the viewing-as phenomenon. My proposal, again, is that *viewing X as Y* involves taking Y as a frame for thinking about X. Understanding viewing-as in this way makes sense of several of its key features, and it helps explain how viewing one thing as another can be cognitively beneficial.

In the next section, I turn to the relationship between viewing-as and mathematical explanation. My first goal will be to show that there is such a relationship and to highlight its philosophical interest. As I indicated at the beginning of the paper, "viewing-as explanations" have a puzzling aspect: how can viewing X as Y explain anything about X, given that X and Y are distinct, and in general apparently not related in any interesting metaphysical way? I'll consider a few possible answers to this question and finally sketch what I think is the most promising account. The "cognitivist" approach I'll describe will make further use of the frame-theoretic viewpoint introduced in this section.

3.4 Viewing-as and mathematical explanation

3.4.1 Some examples

At the outset of the paper, I presented a case in which viewing one thing (numbers) as another (arrays of dots) led to an explanatory insight (about why the sum of the first n odd numbers is n^2). That case is typical of the phenomenon I want to discuss here. But as usual, it's wise to collect more than one piece of data before turning to questions of theory. So I'll begin with some further examples, again drawn from mathematical literature, that exemplify the link between viewing-as and explanation.

Consider first a case discussed by Timothy Gowers, concerning a famous identity due to Euler:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots = \frac{\pi^2}{6}.$$

As Gowers notes, this looks mysterious at first: "What on earth, one might wonder, has π to do with adding up reciprocals of squares? This is a perfectly legitimate question" ([Gowers 2008], 261). He then shows how to answer the question using ideas from Fourier analysis, in the following way. First, given a periodic function $f : \mathbb{R} \to \mathbb{C}$ with period 2π, define the nth Fourier coefficient of f—denoted a_n—by the formula

$$a_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} dx.$$

Then the theorem known as Plancherel's identity describes the sum of the squared moduli of the Fourier coefficients:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |a_n|^2 .$$

Now let f be the function such that $f(x) = 1$ when $\left(2n - \frac{1}{2}\right)\pi \leq x \leq \left(2n + \frac{1}{2}\right)\pi$ for some integer n, and $f(x) = 0$ otherwise. (Note that f has period 2π.) Combining Plancherel's identity with some calculation and manipulation, we get the formula

$$\begin{aligned}\frac{\pi^2}{8} &= 1+\frac{1}{3^2}+\frac{1}{5^2}+\frac{1}{7^2}+\cdots\\ &= \sum_{n=1}^{\infty}\frac{1}{n^2}-\sum_{n=1}^{\infty}\frac{1}{(2n)^2}\\ &= \frac{3}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}.\end{aligned}$$

Finally, multiplying both sides by $\frac{4}{3}$ gives the original statement, $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$. After going through this argument, Gowers writes:

> Now we have a reason for the appearance of π: it comes up in the formula for the Fourier coefficients. What is more, its appearance there can be explained as well. A periodic function on $\mathbb{R}$ is more naturally thought of as a function defined on the unit circle. The Fourier coefficient a_n is a certain average defined on the unit circle, so we have to divide by the length of the circle, which is 2π. (262)

Gowers' reasoning, then, seems to be as follows. We wanted to know at first why π shows up in the identity $\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$. This turns out to involve a result from Fourier analysis: namely, the fact that Plancherel's theorem applied to a certain periodic function f gives a closely related identity involving π and reciprocals of squares. What's more, we can explain why π appears here by viewing f as a function g defined on the unit circle, and noting that the Fourier coefficients are a type of average involving the values of g. Since taking this average involves dividing by the circumference of the circle, the appearance of π is no surprise.

This looks like a case where viewing-as makes a crucial contribution to a mathematical explanation. The function f is defined to have domain $\mathbb{R}$ and range $\mathbb{C}$, and it's not obvious what π has to do with averaging a function like that. It's only when we regard f as a quite different object—the function g defined on the unit circle—that we find the explanation we were after.

A second example comes from Laptev and Rozenfel'd's account of developments in nineteenth-century geometry. The relevant passage occurs during a discussion of Lobachevky's discovery of hyperbolic geometry, and in particular the efforts by him and others to prove its consistency. Lobachevsky himself was only partially successful at this. But later authors, notably Poincaré and Minkowski, discovered an interpretation of the hyperbolic plane that not only showed it to be consistent, but also explained certain similarities between hyperbolic and spherical geometry. As Laptev and Rozenfel'd write:

> [T]he geometry of the hyperbolic plane [with Gaussian curvature $-\frac{1}{q^2}$] can be realized on a sphere of imaginary radius qi in a subspace of complex space, whose points have rectangular x- and y-coordinates that are real ($x = \bar{x}$, $y = \bar{y}$), and z-coordinates that are purely imaginary ($z = -\bar{z}$). This subspace can be regarded as a real affine space in which the distance d between the points with rectangular coordinates x_1, y_1, z_1 and x_2, y_2, z_2 is defined by the formula
>
> $$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 .$$
>
> ...The sphere of radius qi in this space has the form of a hyperboloid of two sheets (the geometry of the hyperbolic plane is realized on each nappe of such a hyperboloid...)... This interpretation, which demonstrates vividly the consistency of Lobachevskii's planimetry, explains why the formulas of hyperbolic trigonometry can be obtained from those of spherical trigonometry by replacing the radius of the sphere by qi. ([Laptev & Rozenfel'd 1996], 64)

The formulas referred to here include the law of cosines and the law of sines, which in spherical geometry take the forms

$$\cos\frac{a}{r} = \cos\frac{b}{r}\cos\frac{c}{r} + \sin\frac{b}{r}\sin\frac{c}{r}\cos A,$$

$$\frac{\sin A}{\sin(a/r)} = \frac{\sin B}{\sin(b/r)} = \frac{\sin C}{\sin(c/r)}$$

(where a, b, c are the side lengths of a triangle, A is the angle opposite a, and r is the radius of the sphere). Lobachevky of course noticed the resemblance between the spherical and hyperbolic formulas, but he lacked the resources to completely understand it—doing so requires Poncelet's later idea of enriching Euclidean space with imaginary points. In the resulting "pseudo-Euclidean space", the notion of a sphere of radius qi makes sense, and one can view the hyperbolic plane as a hemisphere of such a "pseudosphere" (which is realized as a sheet or "nappe" of a hyperboloid). This relationship holds because a sphere of radius r has Gaussian curvature $\frac{1}{r^2}$, whereas a hyperbolic plane has negative Gaussian curvature, so a sphere of imaginary radius corresponds to a hyperbolic surface. Thinking of the hyperbolic plane in this way, then, lets us explain the peculiar form taken by the laws of hyperbolic trigonometry.

One might want to suggest at this point that something more (or other) than viewing-as is going on here. After all, the usual way to describe the above type of situation is by saying that the hemisphere of a pseudosphere is a *model* of the hyperbolic plane, or of the axioms of hyperbolic geometry. Why invoke the relatively unfamiliar and obscure idea of viewing-as when we can instead just talk about modeling, which we already understand quite well? In particular, why not say that the similarity between hyperbolic and spherical trigonometry is explained by this modeling relationship?

My reply is that there's no incompatibility here. I've argued that viewing X as Y consists in taking Y as a frame for structuring one's thinking about X. And this is precisely what we do when we use the modeling relationship between X and Y as a way to draw conclusions about X. At least in this type of scenario, model-oriented cognition just is a kind of frame-oriented cognition. In general, embracing talk of frames doesn't mean rejecting other possible descriptions of the relationships between objects.

Third and finally, here's a case in which visualization plays a key role, from Marc Konvisser's linear algebra textbook. The problem Konvisser starts with is that of finding the solutions to polynomial equations of the form $x^n - 1 = 0$. For values of n less than 5, the solutions can be found by relatively simple algebraic methods (e.g., factoring or applying the quadratic formula). After displaying these solutions, Konvisser writes:

> We have now found all the roots for all equations of the form $0 = x^n - 1$ for $n = 1, 2, 3$, and 4. However, this has given us little insight into how to find the solutions of such equations for $n \geq 5$. To see what is happening, we view complex numbers as vectors. ([Konvisser 1986], 25)

Konvisser then explains how addition and multiplication of complex numbers are to be understood on this interpretation, noting in particular that multiplying a complex number by i corresponds to rotating the associated vector counterclockwise by an angle of $\pi/2$. He continues:

> Now let us see how this interpretation of multiplication by i as rotation by $\pi/2$ can help us solve our original problem of finding the roots of equations of the form $0 = x^n - 1$... In order to find complex numbers that satisfy the equation $z^n = 1$, let us see if we can find complex numbers z so that multiplication by z represents a rotation of $1/n$ way around, that is, a rotation of $2\pi/n$. (29)

Some simple trigonometry leads to the desired general result: "Let $z = \cos\theta + i\sin\theta$. Multiplication by z represents a counterclockwise rotation by the angle θ" (31). This essentially solves the original problem. We wanted to find the solutions to $x^n - 1 = 0$, i.e. the complex numbers z corresponding to rotations of $2\pi/n$. By the previous proposition, one such solution is $z = \cos(2\pi/n) + i\sin(2\pi/n)$. In fact, z is a "primitive nth root of unity", and the remaining roots are the powers $z^2, z^3, \ldots, z^{n-1}$.

Clearly this is a case where viewing one kind of thing (complex numbers) as another (vectors in the plane) yields a satisfying solution to a problem about the original objects. Although Konvisser doesn't explicitly frame the situation in terms of explanation, he does point out the lack of insight afforded by the naïve algebraic approach, and he says that viewing complex numbers as vectors allows us to better "see what is happening" with respect to the roots of $x^n - 1$. I think these remarks are plausibly understood as claims about the relative explanatoriness of the viewing-as strategy. In particular, I take it that facts about $1/n$th rotation vectors explain why the solutions to $x^n - 1 = 0$ have the form that they do.

It wouldn't be hard to add further examples to our list, and no doubt doing so would paint a more vivid picture. But hopefully the cases discussed so far are enough to give a good idea of the phenomenon. As promised, the rest of the paper will be devoted to sharpening a philosophical problem about viewing-as and mathematical explanation, and finally to outlining a solution.

3.4.2 Viewing-as, explanation and dependence

In cases like the ones just presented, we've seen that viewing a mathematical object X as an object Y can be a route to explaining something about X. These cases seem to have roughly the following structure: First, we establish that X is to be viewed as Y. Second, we obtain some fact F about Y. Finally, we use F to explain a corresponding fact F^* about X. The question I'd like to address here is: How is this possible? What sort of relationship must X and Y stand in so that viewing X as Y can be explanatory in this way?

Of course, there's nothing mysterious about many kinds of explanatory relationships. It's not surprising, for instance, that facts about thrown stones explain facts about broken windows, or that facts about threads explain facts about sweaters. That's because there's an appropriate connection between the objects involved in these cases: the stone caused the broken window, and the threads compose the sweater. And everyone agrees that relations like causation and composition can underwrite explanatory relations. Similarly, it's not surprising that facts about lumps of clay explain facts about statues, that facts about my shirt's hue explain facts about its color, or that facts about my C-fibers firing explain facts about my pain. That's because these cases involve special kinds of metaphysical relationships (material constitution, determination and function realization, respectively), and everyone agrees that relations like these can ground explanations.

Indeed, according to a popular line of thought—calling it "the consensus view" is probably not far off—*all* explanations must be underwritten by worldly (physical or metaphysical) grounding relations between the objects appearing in the explanans and those appearing in the explanandum. For example, David-Hillel Ruben says that "Explanations work only in virtue of the determinative relations that exist in the world. ...[W]e explain something by showing what makes it or what is responsible for it" ([Ruben 1990], 231). And Kathrin Koslicki writes that "an explanation, when successful, captures or represents... an underlying real-world relation of dependence of some sort which obtains among the phenomena cited in the explanation in question" ([Koslicki 2012], 212, quoted in [Pincock 2015], 7). Likewise, Jaegwon Kim: "My main proposal... is this: *explanations track dependence relations*. The relation that 'grounds' the relation between an explanans, G, and its explanatory conclusion, E, is that of dependence; namely, G is an explanans of E just in case e, the event being explained, depends on g, the event invoked" ([Kim 1994], 68). And more recently, Mark Povich: "Explanations provide information about relations of ontic dependence, causal and non-causal" ([Povich 2016], 14).

Call this principle the *Dependence Thesis*. I think most everyone can agree that it has much to recommend it. For one, the Dependence Thesis is an intuitively plausible principle supported by a wealth of scientific, philosophical and quotidian examples. Moreover, historically speaking, the emergence of the "ontic" picture of explanation with its emphasis on causal relations—initiated by Wesley Salmon and others in the 1970s[10]—was a key moment in the shift away from Hempel's deductive-nomological theory. In many respects this shift was undeniably a step in the right direction. Finally, Dependence makes life easier for theorists of explanation. The grounding criterion is a neat and convenient way to tell explanations from non-explanations, and it's not clear that there's a similarly simple alternative.

The problem is that there's an apparent conflict between the Dependence Thesis and cases of viewing-as like those discussed above. That's because, as we've seen, there generally aren't any obvious determinative relationships between the source and target objects in such cases. (Of course there aren't any determinative *physical* relations. But neither, in general, are there familiar metaphysical relations like constitution or realization. Sums of odd numbers aren't comprised of arrays of dots in any sense, nor is "array of dots" a functional role that numbers could realize, etc.—on the whole, there seems to be no sense in which dot arrays "make" or are "responsible for" numbers or their properties.) Nevertheless, the viewing-as cases seem to involve genuine explanatory connections between the source and target objects. What conclusions should we draw from this?

It seems to me that there are four main views one might take. The first possibility is that the viewing-as cases don't really involve explanation, so they aren't counterexamples to the Dependence Thesis. The second and third possibilities hold that the cases do involve explanation, but they do so in ways that don't threaten Dependence. According to the second option, the explananda in such cases are something other than the target objects. So the lack of determination relations

[10]As detailed in [Salmon 1989].

between the sources and targets is no problem for Dependence. According to the third option, it's indeed true that facts about the target objects explain facts about the source objects—however, there are also determination relations between the target and source, and hence Dependence is borne out. Finally, the fourth possibility is that the cases are genuine counterexamples, and so the Dependence Thesis is false.

This isn't the place to settle such a complex and weighty issue once and for all. Nevertheless, I think there are good reasons to take the last option seriously, in spite of its departure from orthodoxy. In the next part, I'll explain the problems with the other responses, and in the subsequent part I'll sketch an account of what it might mean to have explanations in the absence of dependence relations.

3.4.3 Defending dependence

3.4.3.1 The first possibility: denying explanatoriness

The first approach to saving the Dependence Thesis is to deny that the viewing-as cases really involve explanation. One can imagine this stance emanating either from skepticism about mathematical explanation in general, or from worries about these cases in particular. I don't intend to discuss the first possibility—although a very few voices of doubt have been raised about the existence of mathematical explanations (notably [Zelcer 2013]), the arguments offered on this score have been unconvincing, and at any rate I think the last fifteen years of research have shown that there's a phenomenon here worth theorizing about.[11] As for the second possibility, it's not clear what features of the above examples would mark them as especially problematic. There's nothing, after all, to suggest that explanation-talk is being used lightly or figuratively in the relevant passages. And in the absence of such warning signs, I think philosophers should default to taking mathematicians' accounts of their practice seriously, especially when those accounts fall into a coherent pattern indicative of a widely shared standard. So, at least pending further argument, this line of thought doesn't look promising.

[11] See [Weber & Frans 2017] for a reply to Zelcer.

3.4.3.2 The second possibility: relocating the explanantia

The second argument in defense of the Dependence Thesis rests on the claim that, although the viewing-as cases really do involve explanations of facts about the source objects, the explanantia in these cases are something other than the target objects. In order to preserve Dependence, of course, the true explanantia will have to be things that *do* bear determination relations to the source objects. What might these be? One option is that it's the *proofs* figuring in these cases that ultimately explain, and also metaphysically ground or determine, the relevant facts about the source objects. This view has a certain weight of tradition behind it: starting with Mark Steiner's work in the 1970s, there's been a tendency to think (or at least assume) that all cases of mathematical explanation involve explanatory proofs in some essential way. If that were true, then this strategy would have some degree of initial plausibility. As I've argued elsewhere in this dissertation, however ([D'Alessandro forthcoming]), such "proof chauvinist" assumptions are unjustified—there are many perfectly good cases of mathematical explanation in which proofs play no important role. So there's no a priori reason to expect they'll solve the dependence problem.

What's more, there's good reason to think that proofs can't play this role. The hypothesis gets the direction of dependence wrong, for one. Surely it's mathematical *objects and properties* that are metaphysically fundamental, and *representations* of those objects and properties, like proofs, that are derivative. The idea that proofs "make" or are "responsible for" facts about mathematical objects is about as plausible as the claim that quarks are made or determined by scientists' reasonings about them.

In any case, there's an alternative way to spell out the second argument. Perhaps the explanantia in the viewing-as cases are really facts about *the source objects themselves*. On this picture, facts about the target objects don't directly contribute anything to the relevant explanations. They may serve some sort of purpose—making an explanation more obvious, or easier to state, or something like that—but they aren't essential. In principle, we could give more or less the same explanation without mentioning the target objects at all. Since there's no mystery about how some properties of an object can be determined by other properties of the very same object, this scenario appears to be consistent with the Dependence Thesis.

The problem with this story is that it's not at all clear how to subtract the target objects from the explanations in the above cases. In what sense could one possibly give "the same" explanation of why the sum of the first n odd numbers is n^2 without talking about arrays of dots (or whatever)?[12] This hardly seems possible, since a key component of the explanans is the fact that certain transformations of square arrays of side n yield square arrays of side $n+1$. Similarly, how might we follow Gowers in explaining the appearance of π in the Euler identity without mentioning functions defined on the circle, when it's precisely the idea of averaging over such a function that brings π into the picture? In these and the other examples, the target objects seem not to function as convenient-but-dispensable accessories. Rather, it looks like the explanations in these cases depend in a constitutive and direct way on facts about the target objects.

3.4.3.3 <u>The third possibility: identifying dependence relations</u>

Finally, consider the third strategy for defending the Dependence Thesis. This line of thought admits that facts about the target objects explain facts about the source objects, but also insists that the former bear appropriate determinative relations to the latter, as required by Dependence.

The recent literature contains a couple of proposals along these lines. [Pincock 2015], for instance, argues at length that some proofs are explanatory because the objects appearing in the proofs stand in a particular metaphysical relation to the objects appearing in the associated theorems. Pincock calls this relation "abstract dependence". Abstract dependence, he says, is "a new kind of dependence relation" that isn't reducible to modal covariance, and "that can obtain in the absence of any essential composition relation."[13]

Pincock's main motivating example is the relationship between polynomial equations and their Galois groups, as described by Galois theory. It's long been known that a given polynomial equation

[12] Obviously there's nothing special about dots in particular—we could just as well use arrays of stars or triangles, say. What's essential to the explanation is the configuration of the arrays, not the properties of the individual elements (except, perhaps, insofar as their sizes and distances from one another affect how easily we can take in the overall configuration).

[13] [Pincock 2015], 11.

is "solvable in radicals"—i.e., that there exists a formula expressing the equation's solutions in terms of its coefficients, basic arithmetical operations and *n*th roots—just in case the polynomial's Galois group has a certain algebraic property, also known as "solvability". There's also widespread agreement that this is a case of mathematical explanation: the Galois-theoretic criterion explains *why* certain equations (but not others) are solvable in radicals. Why should this be? Pincock suggests, in quite Ruben-like terms, that "what makes a given polynomial equation solvable.... is that its Galois group is solvable."[14] That is, we have an explanatory relation in this case because objects of one type abstractly depend on objects of another type.

Pincock doesn't claim that *all* cases of mathematical explanation, or even any particular class of cases, involve abstract dependence. So whether or not the viewing-as cases are examples of this type isn't diagnostic of the rightness or wrongness of Pincock's view. But it's worth asking whether abstract dependence might be the sort of determination relation that the Dependence theorist is looking for.

So what is abstract dependence, exactly? The details of Pincock's account are somewhat involved, but what matters most for our purposes is that an object X abstractly depends on an object Y only if Y is *more abstract* than X. Relative abstractness, in turn, is analyzed in terms of instancehood, in roughly the type-token sense. For example, the word-type *cat* has particular token inscriptions of the word 'cat' as instances. So the word-type is more abstract than the word-token, on Pincock's view. Similarly, the "concrete" group $(\{0,1\},+)$ is an instance of the "abstract" cyclic group C_2. So abstract groups (in particular, the Galois groups of polynomials) count as more abstract than concrete groups (in particular, the automorphism groups of extensions of $\mathbb{C}$ associated with these polynomials). Relative abstractness is only a necessary condition for abstract dependence—a further requirement is Y's being the *least more abstract* object than X such that a certain kind of biconditional holds. But it's enough for our purposes to ask whether the viewing-as cases satisfy even the relative abstractness requirement.

It's clear that the answer is no. None of the cases discussed above involve anything like a type-token relation between the source and target objects—numbers aren't instances of dot arrays, nor

[14][Pincock 2015], 11.

are functions $\mathbb{R} \to \mathbb{C}$ instances of functions defined on the circle, nor is the hyperbolic plane an instance of the hemisphere of a pseudosphere, nor are complex numbers instances of vectors. So Pincock's abstract dependence isn't a candidate for the grounding relation that holds between the objects in these cases.

Another view that links mathematical explanation to metaphysics can be found in [Lange 2015] (see also chapter 9 of [Lange 2016]). Here Lange compares two types of proofs of Desargues' theorem on intersection points: one type which takes place in the setting of classical Euclidean geometry, and which takes the form of "a motley collection of special cases" ([Lange 2015], 438), and another type which achieves a unified treatment using methods from projective geometry. According to Lange and other commentators, only the second sort of proof is explanatory. Lange presents a theory to account for this—a version of the "explanation as coincidence elimination" approach he employs elsewhere—and he also uses the case to motivate some metaphysical conclusions. It's the latter that I want to consider here.

According to Lange, the fact that projective points can be used to explain theorems of ordinary Euclidean geometry entails that there must be some ontological connection between the two. In particular, such an explanatory relationship can obtain only if projective objects "exist in" Euclidean geometry. Lange summarizes this argument as follows:

> P1: Certain facts about points at infinity explain certain facts about Euclidean points, lines and planes. [...]
> P2: What explains a fact about some entities must be on an ontological par with those entities. (Roughly: only facts about what exists can explain facts about what exists.)
> C: Points at infinity exist in Euclidean geometry. ([Lange 2015], 461)

What does it mean for a mathematical object to "exist in" a particular mathematical domain? Lange doesn't say much about this, as he seems not to have a settled view. Indeed, he leaves open some fairly radical possibilities, including the view that "what it is for points at infinity to exist in Euclidean geometry is for them to play an explanatory role there" (462). I take it that traditional approaches to ontology regard existence as decidedly more fundamental than explanation. So this type of view would involve major revisions to much of our basic metaphysics.

For this reason, it may be worth considering a more conservative interpretation of Lange's proposal. (I don't claim that Lange would or should endorse this interpretation. The view to be described strikes me as a straightforward way of reading his remarks, and as an idea that may have some independent interest. But I don't doubt that Lange could come up with something better.) The idea is this. Mathematical objects are divided metaphysically into various collections—let's call them "domains"—such that facts about X can explain facts about Y only if X and Y belong to the same domain. (Perhaps an object can belong to more than one domain, so we won't assume that this division is a partition.) "X exists in D" and "X is on an ontological par with Y" are then domain-theoretic statements: the first names a domain that X belongs to, and the second says that X and Y belong to some common domain. This proposal may not give defenders of the Dependence Thesis everything they need, but it gives at least a necessary metaphysical condition on facts about X explaining facts about Y.

I have three comments about this interpretation of "existing-in". First, the domains it posits probably won't look much like the familiar classifications of mathematical objects (since there's so much "impure" mathematical explanation cutting across traditional subject boundaries[15]), and there probably won't be more than a few distinct domains (for the same reason). So it's unclear how interesting the resulting metaphysical structure will be. Second, even if the proposal is correct, it still gives us little insight into the nature of the determinative relation we were looking for. After all, the claim that X and Y belong to the same domain doesn't tell us anything about whether X-facts explain Y-facts, or vice versa, or both, or neither. In particular, belonging to the same domain isn't *sufficient* for instantiating relations of explanation or dependence. (Presumably the integers 33 and -2766018 belong to the same domain, but I doubt that either one determines or explains anything about the other.) Finally, the proposal raises some awkward metaphysical questions. Why should there be domains of this sort in the first place? In virtue of what does a particular mathematical object belong or not belong to a particular domain? It would be mysterious if the existence and composition of domains were brute ontological facts, but it's also unclear what

[15] As Lange notes himself: "a proof's explanatory power is distinct from its 'purity' in the rough sense of its making use of no concepts foreign to the concepts in the theorem being proved" ([Lange 2016], 292).

other facts we could appeal to for an explanation. I conclude that this way of interpreting Lange is neither very appealing nor, even if true, likely to be of much use in defending the Dependence Thesis.

So where does the third strategy stand? The idea, again, was to identify a determinative metaphysical relation between the source and target objects in the viewing-as cases discussed above. We saw previously that none of the familiar grounding relations is a good candidate. And I've just now argued that Pincock's abstract dependence and Lange's "coexistence" proposal don't seem obviously helpful either. (Again, this isn't a criticism of either author, since neither set out to deal with this particular issue.) Of course, it's possible that nobody has discovered the true relation yet. But I don't see much reason to hold out hope. There simply doesn't seem to be any interesting metaphysical theme running through the relevant cases. (What do the pairs (numbers, dot arrays), (hyperbolic planes, hemispheres of pseudospheres), and (periodic functions $\mathbb{R} \to \mathbb{C}$, functions defined on the unit circle) have in common, metaphysically speaking?)

Since it's at best unclear how to preserve Dependence in light of the viewing-as cases, the last of the four responses mentioned above seems well worth considering. The idea here, recall, is to maintain that the viewing-as cases involve explanatory relations between the source and target objects, but to deny that these explanatory relations are underwritten by metaphysical facts about grounding or determination (in the general case). This viewpoint could of course be developed in many different ways. I don't have the space to consider all the possibilities here, or to discuss even one in the level of detail it deserves. Instead, I'll just aim to outline the type of account that strikes me as particularly promising.

3.5 Viewing-as explanations without (much) metaphysics

3.5.1 A "cognitivist" view

The account to be sketched here might be called "cognitivist" by its friends, or "psychologistic" by its enemies. The idea is roughly that a viewing-as explanation of a fact F is a fact about a different object that makes us cognitively better off with respect to F. To fill in the details, I return to the language of frames introduced in §3.

As we saw, external frames work by imposing or suggesting correspondences between the source and target objects. (These correspondences are relatively detailed and determinate in the case of analogy, and relatively loose and open-ended in the case of metaphor.) In either type of case, we can often avail ourselves of the frame structure to make the following sort of inference:

1. Claim C about the source object X corresponds to claim C^* about the target object Y (under correspondence $\mathscr{R}$).
2. C^* is true of Y.
3. Therefore C is (probably) true of X.

If the correspondence $\mathscr{R}$ is a good one, then the products of such inferences will typically be true. What's more, in a well-chosen frame the target object Y will be more cognitively tractable than the source object X in the relevant respects, so that the truth of C^* is more intuitive or easier to establish than the truth of C.

I propose that this is essentially the sense in which we can explain something about one mathematical object by viewing it as another object. On this picture, we have a "viewing-as explanation" of a fact F when we can deploy an external frame that puts us in a better cognitive position with respect to F: from the "naïve" position, the truth of F is non-obvious, unexpected, or only ascertainable with difficulty (if at all), but with the help of the frame, F's truth is relatively clear and intuitive. Insofar as this use of frames reduces our surprise with respect to F or improves our cognitive handle on it, it seems not entirely unreasonable to call it an explanation.

This proposal matches what we've witnessed in the earlier examples. Naïvely, one can prove by induction that the sum of the first n odd numbers is n^2, but this approach takes some work and fails to make the identity seem any more inevitable or obvious. By contrast, it's quite clear in the dot-diagram setting that a single dot is a square array of side 1, and that adding the nth odd number of dots to a square array of side $n-1$ yields a square array of side n. Combining this fact with the correspondence induced by viewing numbers as dot arrays, one can obtain the original identity with similar ease.

The other cases involve similar "transfers of obviousness". It's initially surprising that π would appear in the formula for the sum of reciprocals of squares, but unsurprising that averaging a function defined on the circle might involve division by 2π. Once we're able to think about the former in terms of the latter, however, the unsurprisingness of the second fact can be transported to the original setting, making the first fact seem similarly natural. Likewise, we would expect the law of cosines to take the form

$$\cos\frac{a}{qi} = \cos\frac{b}{qi}\cos\frac{c}{qi} + \sin\frac{b}{qi}\sin\frac{c}{qi}\cos A$$

on a sphere of radius qi, but it's not immediately clear that it should look this way in the hyperbolic plane. What makes it clear is the fact that there's a correspondence between the two types of object—a sphere of radius r has Gaussian curvature $\frac{1}{r^2}$, whereas a hyperbolic plane has negative Gaussian curvature, so a sphere of imaginary radius can be thought of as a hyperbolic surface—in virtue of which the result follows easily. One can tell the same sort of story about the Konvisser case, and I predict that similar remarks will apply to other instances of viewing-as explanation.

3.5.2 Clarifications and objections

Let me briefly try to clarify some things about this view and address a couple objections.

First I should acknowledge that the view isn't quite as unprecedented and heretical as I've perhaps made it out to be. Although the ontic conception of explanation has long held sway in metaphysics and philosophy of science, there's a longstanding alternative tradition that emphasizes the relationship between explanation and understanding, prediction, question-answering, and other epistemic or cognitive goods. Salmon called Hempel's deductive-nomological view an "epistemic" theory of explanation, since it holds that explanations are essentially the same as predictions and retrodictions; cf. [Hempel 1958]. (For a recent defense of a prediction-centered view of explanation, see [Douglas 2009].) Meanwhile, "pragmatic" accounts (e.g. [van Fraassen 1980] and [Achinstein 1983]) claim that what counts as an explanation depends on features of agents and their contexts—for example, on "whether provision of a certain body of information to some audience produces understanding or a sense of intelligibility or is appropriate or illuminating for that audience" ([Woodward 2014]). Others have argued that understanding in particular is an essential feature or goal of explanations (e.g. [De Regt & Dieks 2005], [Grimm 2010], [Turner 2013]). So I'm not the first to claim that some explanations are not best thought of in terms of worldly dependence relations.

The association of explanation with something like "surprise reduction" may be the historical view most similar in spirit to my proposal. As William Dray wrote, "To explain a thing is sometimes merely to show that it need not have caused surprise" ([Dray 1957], 157). Ernst Mach makes similar remarks in *The Science of Mechanics*: "In the infinite variety of nature many ordinary events occur; while others appear uncommon, perplexing, astonishing, or even contradictory to the ordinary run of things. As long as this is the case we do not possess a well-settled and unitary conception of nature." Our goal, Mach suggests, is to reach a state where the once-puzzling facts "appear to us as things that are familiar; we are no longer surprised, there is nothing new or strange to us in the phenomena, we feel at home with them, they no longer perplex us, they are *explained*" ([Mach 1960], 7; italics in original). In addition to these kinds of informal remarks, some theories of explanation have featured surprise reduction as a constitutive element (e.g. [Gärdenfors 1980]). Of course, my proposal is about increases in cognitive tractability, a relatively general phenomenon of which I take surprise reduction to be a special case. In the examples discussed so far, some of the explananda likely do (or should) elicit a feeling of surprise, and some of the explanations help mitigate this feeling. But this doesn't seem true in general. For instance, the fact that the solutions to $x^n - 1 = 0$ for the first few n are $\{1\}$, $\{\pm 1\}$, $\left\{1, \frac{-1+i\sqrt{3}}{2}, \frac{-1-i\sqrt{3}}{2}\right\}$, $\{\pm 1, \pm i\}$ doesn't strike me as particularly strange or unexpected. So I doubt that surprise reduction is the right story here. Nevertheless, viewing complex numbers as vectors certainly does something to improve one's cognitive handle on these facts (and on roots of unity in general)—it does so, for example, by making it more obvious that the solutions had to take this form, making it easier to derive some of them, and providing a better understanding of what the solutions have in common. It seems to me that these and other factors are present to differing degrees in different cases. So it would be unduly narrow to focus on any particular way of being cognitively better off.

The account I've sketched here also differs in other ways from the views of recent critics of the ontic approach to explanation. For instance, many of these philosophers hold that explanations are mental or linguistic items (e.g. arguments), or else human activities of some sort (e.g. speech acts). This isn't my view. I think viewing-as explanations are mind-independent entities, like the fact that the law of cosines takes a certain form on the sphere of radius qi. (My disagreement with Dependence theorists is about what *makes* such facts explanations—they aspire to point to underlying determinative relations, while I point to the cognitive effects of frames.) Conversely, most "epistemicists" about explanation generally aren't motivated by the thought that some (true) explanations lack metaphysical grounds, as I am. Rather, they're impressed by such things as the idea that "many if not most of the explanations that are advanced in science turn out to be false" ([Bechtel 2008], 18), which suggests that explanations are representational items rather than facts or objects in the world. In spite of its opposition to the ontic conception, then, it's not particularly illuminating to associate the current proposal with the mainstream epistemicist tradition in philosophy of science.

Let me now consider a few objections. First, a common worry about "psychologistic" philosophical theories is that they make the target phenomenon out to be subjective and arbitrary. I'm not sure that this is always inherently bad, but in any case this type of criticism would be largely misplaced here. On the view I've sketched, whether or not an instance of viewing-as is explanatory depends on whether or not the relevant frame makes some fact about the source object more obvious or intuitive than it would otherwise be. Here, "frame $\mathscr{M}$ makes fact F more obvious" should be understood to mean roughly "frame $\mathscr{M}$ is disposed to make fact F more obvious when deployed

[16]Except to the extent that what counts as "ordinary" should be sensitive to variations in which concepts are available, which facts are common knowledge, and so on. If the "ordinary" condition seems unappealing or unworkable, another option is to go subjectivist and allow that different things count as explanatory to different people at different times. Determining what ordinariness should be, and which of these options is ultimately preferable, is an important task that I won't be able to execute here.

by ordinary subjects under ordinary circumstances". And whether or not a particular frame has such a disposition is an objective and principled matter of fact. It's objective, at least, insofar as it doesn't vary (much) across persons and times.[16] And it's principled insofar as it's based on lawful cognitive regularities. Perhaps this won't be enough to satisfy diehard opponents of psychologism, but at least it means that the phenomenon is sufficiently stable, intelligible and well-grounded to be worth theorizing about.

A second objection[17]: I've characterized viewing-as explanations in terms of "transfers of obviousness" (or of other kinds of cognitive tractability). The idea is that we can combine an obvious fact about a target object with a frame-induced principle of correspondence to render a fact about a source object more obvious. But, one might ask, what about the obviousness of the correspondence principle itself? If such a principle lacks some positive cognitive status, can it still play a role in conferring that positive status on something else? It would be worrisome if the answer were no. After all, many of the correspondences in the above examples aren't particularly apparent and intuitive.

In fact, however, non-obvious correspondence principles can be (and often are) used to make other things more obvious. A non-mathematical case might help make this clear. Suppose I'm a well-trained radiologist. As I stand here looking at John's X-ray slides, it's clear to me that there's a certain sort of dark spot in the bottom right corner of the third image. In virtue of my radiological expertise, I immediately conclude—with a similar and equally warranted sense of intuitive obviousness—that John has a tumor in his lymph node.

I'm able to draw this conclusion because I know there's a certain correspondence between patterns on X-ray slides and tissue conditions. But does this correspondence also have to be obvious to me? I don't think so. After all, learning radiology requires mastering a lot of subtle and tricky rules about slide interpretation. These rules definitely aren't a priori obvious, and they need not even be a posteriori obvious—I might well know that the dark spot means a tumor without knowing why this correspondence holds, or having a strong intuitive sense that it should hold, or anything of the sort. As long as I know that the correspondence *does* hold, however, I can use it to "transfer

[17]Thanks to Marc Lange for raising this issue.

obviousness" from facts about slides to facts about John's condition. In general, given propositions P and Q, it seems that Q can come to be obvious (or whatever) as long as P is obvious and $P \leftrightarrow Q$ is merely *known*. So there's no problem with the correspondence principles lacking in cognitive tractability in the viewing-as cases.

A third issue[18] is whether there are legitimate viewing-as explanations in the sciences. If not, then this difference needs to be accounted for somehow—why should a certain style of explanation be permissible in one realm but not in the other? On the other hand, if scientists do accept viewing-as explanations, how does this acceptance square with the widespread endorsement of the Dependence Thesis in the philosophy of science?

Perhaps the most natural way to approach this question is through the lens of models. For one thing, there's an existing literature on whether and how scientific models can be explanatory. And it's plausible, as Camp suggests, that "we can illuminate the utility and effects of at least many models by treating them as frames" ([Camp forthcoming], 15)—after all, understanding a phenomenon in terms of a model is just using the model to structure one's thinking about the phenomenon. So questions about viewing-as explanations in science (understood as frame-based explanations) are at least closely related to questions about the explanatoriness of scientific models.

As for whether models explain, an emerging consensus holds that they can, at least in certain cases. Some kinds of model are claimed to be explanatory because they represent the actual causes or mechanisms responsible for the phenomena in question (cf. [Craver 2006]). This sort of model-explanation is obviously compatible with the Dependence Thesis. Other models are thought to explain by identifying various kinds of noncausal dependence relations. For instance, [Saatsi & Pexton 2013] claims that the model of Brown, Enquist and West gives a (potential) geometric explanation of Kleiber's allometric scaling law: "The model seems to explain by virtue of... [showing] how the scaling exponent counterfactually varies with the dimensionality of organisms. But this explanatory modal information is not easily construed as causal dependence" (620).[19] De-

[18] I'm grateful to Kenny Easwaran for prompting me to say more about this point.

[19] See also [Berger 1998] for an early account of a geometric-style modeling explanation, and [Rice 2015] for a discussion of noncausal "optimality explanations".

pendence theorists should have no problem with these cases either, since they still involve worldly determination relations.

If Dependence is true, then we should expect all explanatory models to belong to one of these two general types. This is the view espoused by [Povich 2016], who writes that "[models] are explanatory because they accurately represent the relevant dependence relations, that is, the objective features of the world on which the explanandum phenomenon counterfactually depends" (4).[20] However, explanatory power has sometimes been claimed for other, less Dependence-friendly kinds of models. For instance, [Bokulich 2011] argues that some "fictional models" are explanatory, although such models fail to represent the actual world even in an approximate or idealized way. According to Bokulich, a fictional model "explains the explanandum by showing that the counterfactual structure of the model is isomorphic (in the relevant respects) to the counterfactual structure of the phenomenon" (43).

Take Bohr's model of the atom. Although the Bohr model with its classical-style electron orbits fundamentally misrepresents reality, Bokulich thinks the model nevertheless explains the spectrum of the hydrogen atom. This is because there's "a pattern of counterfactual dependence of the emission spectrum of hydrogen on the elements represented in Bohr's model" (43).[21] The similar counterfactual structures of Bohr's model and of actual atoms clearly aren't grounded in any objective worldly dependence. In no sense do the orbits represented by the Bohr model cause, or make, or otherwise metaphysically determine the spectrum of hydrogen (or other features of atoms)—plainly they can't, since nothing like Bohr's orbits exists in reality. The most that can

[20] Povich is talking specifically about the "minimal models" of [Batterman & Rice 2014] here, but the rest of the paper makes it clear that the thesis applies to all models, and indeed to scientific explanations in general.

[21] I find this claim a little puzzling, since it's unclear to me in what sense the spectrum of hydrogen counterfactually depends on the features represented by the Bohr model. Taken at face value, the claim seems false, since the key elements represented by the model don't even exist. Perhaps what Bokulich means is that, *if* the model had been an accurate representation of atomic structure, *then* the spectrum would have counterfactually depended on the properties of electron orbits and the like. This is true, but of course we can say the same about any model of anything—if the phenomenon in question had been like the model, then the properties of the phenomenon would have depended on the properties of the elements represented in the model. So this can't be a way to identify explanatory models in particular. Alternatively, Bokulich might mean that if the model were different, then this would have been because the relevant observational data (including the spectrum of hydrogen) had been different—i.e., Bohr wouldn't have proposed the model if it hadn't been empirically adequate. But this is a "backtracking" interpretation of the counterfactual and hence can't be used to determine counterfactual dependence. (Cf. [Lewis 1973].) I think the likeliest story is that Bokulich's official view is given by the earlier statement about "isomorphic counterfactual structures", and the current claim is an accidental misstatement.

be said is that "Bohr's model is able to correctly answer a number of 'what-if-things-had-been-different questions,' such as how the spectrum would change if the orbits were elliptical rather than circular... This shows that Bohr's model is not simply an ad hoc fitting of the model to the empirical data, as would be the case in a merely phenomenological model" (43).

Is Bokulich right that fictional models can genuinely explain, or must all explanatory scientific models respect Dependence, as Povich suggests? I won't try to settle this question here, but I think either alternative is broadly compatible with my proposal about viewing-as. On the one hand, if models like Bohr's are explanatory, then the scientific and mathematical realms are alike in allowing viewing-as explanations that aren't grounded in dependence relations.[22] This would make the phenomenon I've described in this paper seem less peculiarly parochial. On the other hand, there's a sense in which Povich is justified in wanting to exclude viewing-as explanations that violate Dependence. After all, one of the distinctive goals of scientific explanation is gaining control over the explanandum phenomenon, or at least knowing what such control requires and to what extent it's possible in principle. (This, of course, is the main idea behind mechanism-based theories of scientific explanation.) While viewing-as explanations improve understanding, reduce surprise, and confer other cognitive benefits, they conspicuously fail to tell us how to manipulate, control or intervene.

This is no problem in the mathematical setting, since control is a nonissue. But it gives scientists good reason to be dubious. Indeed, since viewing-as explanations are unable to perform one of the core functions of scientific explanations—and since there are generally alternative causal or metaphysical explanations that do this job better—it may be reasonable to deny that viewing-as cases even count as members of that category. This is an area where scientists and mathematicians

[22] Bokulich attributes this explanatoriness to similarity of counterfactual structure rather than the possibility of transferring cognitive tractability. The relationship between these two proposals is worth further discussion. I'm inclined to think that counterfactual similarity by itself doesn't amount to very much if it fails to yield any epistemic or cognitive benefits. A sketch of an argument: Suppose we want to study a complex phenomenon P, and we propose to do so by modeling P with a model M. As it turns out, M and P have similar counterfactual structures in certain respects. But M is extremely complex itself—it's so hard to work with that it leads to no new predictions, no better understanding, and so on. Moreover, the elements of M in no way accurately represent the actual components of P—it's just a coincidence that the two agree in certain ways.

Would anyone be tempted to say that M explains P? Presumably not. To the extent that the Bohr model is plausibly explanatory, this is surely at least in part because it's more vivid, intuitive and easy to work with than other ways of thinking about atomic structure, not just because it makes certain (actually and counterfactually) correct predictions.

have importantly different interests and expectations. So it's not surprising that the explanatory status of fictional models and the like is controversial, and it's no threat to my proposal if viewing-as explanations are accepted in one domain but not the other.

3.5.3 An application: intertheoretic reduction and explanation

I want to conclude by briefly mentioning an application. The issue I have in mind is that of *intertheoretic reduction* in mathematics. This is a particularly natural setting for viewing-as: after all, if theory S is reducible to theory T, then one can view the S-objects as T-objects according to the correspondence induced by the reduction.

Elsewhere in this dissertation ([D'Alessandro 2017]), I've argued that some mathematical reductions have explanatory value, while others don't. In particular, the reduction of arithmetic to set theory is unexplanatory (and similarly for a number of other set-theoretic reductions). In terms of viewing-as, this means that "viewing the natural numbers and arithmetical operations as set-theoretic constructs is a useful expedient for certain purposes, but doing so has no particular explanatory value" ([D'Alessandro 2017], 7-8). By contrast, I suggested that the reduction of classical algebraic geometry to scheme theory is explanatory, in the sense that viewing classical varieties as schemes allows us to explain some things about the former.

These are useful test cases for the account of "viewing-as explanation" I've proposed here. If viewing numbers as sets is unexplanatory, this should mean that no arithmetic facts are made more obvious, more intuitive or less surprising by applying the frame induced by set theory. On the other hand, if viewing varieties as schemes is explanatory, we should expect the scheme-theoretic frame to make us cognitively better off with respect to some facts about varieties.

Are these predictions borne out? I think so. Viewing numbers as sets doesn't make any arithmetical facts easier to prove or otherwise more cognitively tractable. (I can't defend this claim at length here, but see §2.4 of [D'Alessandro 2017] for a detailed discussion of one case that might have seemed like a promising counterexample.) By contrast, viewing varieties as schemes does make some classic results in algebraic geometry more clear and intuitive. For instance, schemes are often invoked to make better sense of intersections, so that results like Bézout's theorem can be stated and proved in a simpler and clearer way. As Perrin says, "Even if X and Y are varieties, the scheme structure on the intersection is fundamentally important: it explains intersection multiplicities and all contact phenomena"[23]. Similarly, Eisenbud and Harris write:

> How do schemes... arise in practice? One way is as intersections of curves. For example, when we want to work with the intersection of a line L and a conic C that happen to be tangent, it is clearly unsatisfactory to take their intersection in the purely set-theoretic sense; a line and a conic should meet twice. Nor is it completely satisfactory to describe $C \cap L$ as their point of intersection "with multiplicity two": for example, the intersection should determine L, as it does in the nontangent case. The satisfactory definition is that $C \cap L$ is the subscheme of [the affine plane] $\mathbb{A}^2_K$ defined by the sum of the ideals I_C and I_L so that, for example, the line $y = 0$ and the parabola $y = x^2$ will intersect in the subscheme $X_{0,1} = \operatorname{Spec} K[x,y] / (x^2, y)$. This does indeed determine L, as the unique line in the plane containing $X_{0,1}$.[24]

This section has been brief and plenty of details remain to be filled in. But I think these examples give some reason for optimism that the account I've offered is on the right track.

3.6 Conclusion

This paper has tried to shed some light on the phenomenon of viewing-as in mathematics. As I hope is now clear, there are major challenges and rewards involved in this undertaking, extending across fields including the philosophy of mathematical practice, cognitive theory, metaphysics and philosophy of science. In particular, viewing-as is worth studying for its connection to mathematical explanation and the questions it raises about explanation in general.

To summarize, the claims I've defended here include the following. First, I've suggested that viewing-as is usefully understood as a sort of frame-theoretic cognitive phenomenon. In particular, viewing an object X as an object Y amounts to using Y as an external frame for structuring one's thinking about X. Second, I've tried to show that mathematicians often use facts about target objects as explanations of facts about source objects in cases of viewing-as. Third, I've noted that

[23] [Perrin 2008], 212.

[24] [Eisenbud & Harris 2000], 59-60.

there's a prima facie puzzle about how viewing X as Y could explain anything about X, given the apparent lack of familiar dependence relations between mathematical objects of different types. It's not immediately clear whether this puzzle is to be solved by locating an appropriate dependence relation or by rejecting the idea that such a relation needs to be found. Finally, I've argued that the latter option is more appealing than the former. There are no plausible determination relations in sight, but understanding viewing-as explanations in terms of the cognitive effects of frames leads to a promising-looking theory.

Appendix

Springer permission form for “Arithmetic, Set Theory, Reduction and Explanation”

The Springer license document is included on the subsequent pages. The following statement on permissions for reuse in a thesis appears on page 2 below:

“If your request is for reuse in a Thesis, permission is granted free of charge under the following conditions:

This license is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published, permission needs to be reobtained.

- includes use in an electronic form, provided it is an author-created version of the thesis on his/her own website and his/her university’s repository, including UMI (according to the definition on the Sherpa website: http://www.sherpa.ac.uk/romeo/);

- is subject to courtesy information to the co-author or corresponding author.”

SPRINGER LICENSE
TERMS AND CONDITIONS

Nov 08, 2017

This Agreement between Mr. William D'Alessandro ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4224310934316
License date	Nov 08, 2017
Licensed Content Publisher	Springer
Licensed Content Publication	Synthese
Licensed Content Title	Arithmetic, set theory, reduction and explanation
Licensed Content Author	William D’Alessandro
Licensed Content Date	Jan 1, 2017
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of your thesis / dissertation	Dimensions of Mathematical Explanation
Expected completion date	Aug 2017
Estimated size(pages)	66
Requestor Location	Mr. William D'Alessandro 1624 S. Fairfield Ave. Apt. 1F CHICAGO, IL 60608 United States Attn: Mr. William D'Alessandro
Billing Type	Invoice
Billing Address	Mr. William D'Alessandro 1624 S. Fairfield Ave. Apt. 1F CHICAGO, IL 60608 United States Attn: Mr. William D'Alessandro
Total	0.00 USD
Terms and Conditions	

Introduction
The publisher for this copyrighted material is Springer. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions

shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.

Other conditions:

V 12AUG2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Cited Literature

[Achinstein 1983] Achinstein, Peter. 1983. *The Nature of Explanation*. Oxford University Press: New York.

[Artebani & Kondo 2011] Artebani, Michael and Shigeyuki Kondo. 2011. The moduli of curves of genus six and K3 surfaces. *Transactions of the American Mathematical Society* 363, 1445-1462.

[Avigad 2008] Avigad, Jeremy. 2008. "Understanding proofs." In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford University Press: New York.

[Baker 2010] Baker, Alan. 2010. Mathematical induction and explanation. *Analysis* 70, 681-689.

[Baker 2012] Baker, Alan. 2012. Science-driven mathematical explanation. *Mind* 121, 243-267.

[Balaguer 1998] Balaguer, Mark. 1998. "Non-uniqueness as a non-problem." *Philosophia Mathematica* 6, 63-84.

[Batterman & Rice 2014] Batterman, Robert and Collin C. Rice. 2014. "Minimal model explanations." *Philosophy of Science* 81, 349-376.

[Beaney et al. forthcoming] Beaney, Michael, Brenda Harrington and and Dominic Shaw (eds.). Forthcoming. *Aspect Perception After Wittgenstein: Seeing-as and Novelty*. Routledge: New York.

[Bechtel 2008] Bechtel, William. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Routledge: New York.

[Benacerraf 1965] Benacerraf, Paul. 1965. "What numbers could not be." *Philosophical Review* 74, 47-73.

[Bender et al. 2009] Bender, Carl M., Daniel W. Hook and Karta Singh Kooner. 2009. Complex elliptic pendulum. In *Asymptotics in Dynamics, Geometry and PDEs; Generalized Borel Summation*, eds. O. Costin, F. Fauvet, F. Menous and D. Sauzin, 1-18. Pisa: Edizioni Della Normale.

[Bender & Goldman 1975] Bender, E.A. and J.R. Goldman. "On the applications of Mobius inversion in combinatorial analysis." *American Mathematical Monthly* 82, 789-803.

[Berger 1998] Berger, Ruth. 1998. "Understanding science: Why causes are not enough." *Philosophy of Science* 65, 306-332.

[Blass & Rupprecht 2015] Blass, Andreas and Nicholas Rupprecht. 2015. "Ultrafilters and cardinal characteristics of the continuum." Lectures notes from the 2010 meeting of the Appalachian Set Theory Workshop. Online at <http://www.math.cmu.edu/~eschimme/Appalachian/BLASS.pdf>.

[Bokulich 2011] Bokulich, Alisa. 2011. "How scientific models can explain." *Synthese* 180, 33-45.

[Bolza 1891] Bolza, Oskar. 1891. On the theory of substitution-groups and its application to algebraic equations [continued]. *American Journal of Mathematics* 13, 97-144.

[Bos 1974] Bos, H.J.M. 1974. "Differentials, higher-order differentials and the derivative in the Leibnizian calculus." *Archive for the History of the Exact Sciences* 14, 1-90.

[Bos 2001] Bos, Henk J.M. 2001. *Redefining Geometrical Exactness: Descartes' Transformation of the Early Modern Concept of Construction*. Springer-Verlag: Heidelberg.

[Brown 1997] Brown, James Robert. 1997. Proofs and pictures. *British Journal for the Philosophy of Science* 48, 161-180.

[Camp 2008] Camp, Elisabeth. 2008. "Showing, telling and seeing: Metaphor and 'poetic' language." *Baltic International Yearbook of Cognition, Logic and Communication* 3, 1-24.

[Camp forthcoming] Camp, Elisabeth. Forthcoming. "Imaginative frames for scientific inquiry: Metaphors, telling facts and just-so stories." In Peter Godfrey-Smith and Arnon Levy (eds.), *The Scientific Imagination*. Oxford University Press: New York.

[Carter 2008] Carter, Jessica. 2008. "Visualization and understanding in mathematics." In Bharath Sriraman, Claus Michelsen, Astrid Beckmann and Viktor Freiman (eds.), *Interdisciplinary Educational Research in Mathematics and Its Connections to the Arts and Sciences*. Information Age Publishing: Charlotte, NC.

[Cellucci 2008] Cellucci, Carlo. 2008. The nature of mathematical explanation. *Studies in History and Philosophy of Science* 39, 202-210.

[Coliva 2012] Coliva, Annalisa. 2012. "Human diagrammatic reasoning and seeing-as." *Synthese* 186, 121-148.

[Correia & Schnieder 2012] Correia, Fabrice and Benjamin Schnieder (eds.). 2012. *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press: Cambridge.

[Corry 2004] Corry, Leo. 2004. *Modern Algebra and the Rise of Mathematical Structures* (2nd edition). Basel: Birkhäuser Verlag.

[Craver 2006] Craver, Carl F. 2006. "When mechanistic models explain." *Synthese* 153, 355-376.

[D'Alessandro 2017] D'Alessandro, William. 2017. "Arithmetic, set theory, reduction and explanation." *Synthese*, doi:10.1007/s11229-017-1450-8.

[D'Alessandro forthcoming] D'Alessandro, William. Forthcoming. "Mathematical explanation beyond explanatory proof." *British Journal for the Philosophy of Science*.

[Dauben 1979] Dauben, Joseph. 1979. *Georg Cantor: His Mathematics and Philosophy of the Infinite*. Princeton University Press: Princeton.

[Dauben 1988] Dauben, Joseph. 1988. "Abraham Robinson and nonstandard analysis: History, philosophy, and foundations of mathematics."

Minnesota Studies in Philosophy of Science 11, 177-200.

[Day & Krebs 2010] Day, William and Victor J. Krebs (eds.). 2010. *Seeing Wittgenstein Anew: New Essays on Aspect-Seeing*. Cambridge University Press: New York.

[Deligne 1998] Deligne, Pierre. 1998. "Quelques idées maîtresses de l'oeuvre de A. Grothendieck." In *Materiaux pour l'Histoire des Mathematiques au XX^e Siecle*. Société Mathématique de France: Paris.

[De Regt & Dieks 2005] De Regt, Henk W. and Dennis Dieks. 2005. "A contextual approach to scientific understanding." *Synthese* 144, 137-170.

[Descartes 1954] Descartes, René. 1954. *The Geometry of René Descartes*. Translated by David Eugen Smith. Dover: New York.

[Devaney 1994] Devaney, Robert L. 1994. "The complex dynamics of quadratic polynomials." In Robert L. Devaney (ed.), *Complex Dynamical Systems: The Mathematics Behinds the Mandelbrot and Julia Sets*. American Mathematical Society: Providence, RI.

[Dieudonné 1985] Dieudonné, Jean. 1985. *History of Algebraic Geometry*. Translated by Judith D. Sally. Wadsworth Advanced Books & Software: Monterey, CA.

[Dipert 1982] Dipert, Randall R. 1982. "Set-theoretical representations of ordered pairs and their adequacy for the logic of relations." *Canadian Journal of Philosophy* 12, 353-374.

[Douglas 2009] Douglas, Heather. 2009. "Reintroducing prediction to explanation." *Philosophy of Science* 76, 444-463.

[Dray 1957] Dray, William. 1957. *Laws and Explanations in History*. Oxford University Press: Oxford.

[Eisenbud & Harris 1992] Eisenbud, David and Joe Harris. 1992. *Schemes: The Language of Modern Algebraic Geometry*. Wadsworth & Brooks/Cole Advanced Books & Software. Pacific Grove, CA.

[Eisenbud & Harris 2000] Eisenbud, David and Joe Harris. 2000. *The Geometry of Schemes*. Springer-Verlag: New York.

[Eisenbud & Harris 2013] Eisenbud, David and Joe Harris. 2013. *3264 and All That: Intersection Theory in Algebraic Geometry*. Online at <https://scholar.harvard.edu/files/joeharris/files/000-final-3264.pdf>.

[Ellenberg 2014] Ellenberg, Jordan. 2014. *How Not to Be Wrong: The Power of Mathematical Thinking*. Penguin Books: New York.

[Entman 2007] Entman, Robert. 2007. "Framing bias: Media in the distribution of power." *Journal of Communication* 57, 163-173.

[Everitt 2007] Everitt, Brent. 2007. *Symmetries of Equations: An Introduction to Galois Theory*. Online at <http://www-users.york.ac.uk/~bje1/galnotes.pdf>.

[Farjoun et al. 2007] Farjoun, Emmanuel Dror, Rüdiger Göbel and Yoav Segev. 2007. Cellular covers of groups. *Journal of Pure and Applied Algebra* 208, 61-76.

[Feferman 1960] Feferman, Solomon. 1960. "Arithmetization of mathematics in a general setting." *Fundamenta Mathematicae* 49, 35–92.

[Feferman 2009] Feferman, Solomon. 2009. "Conceptions of the continuum." *Intellectica* 51, 169-189.

[Floyd 2010] Floyd, Juliet. 2010. "On being surprised: Wittgenstein on aspect-perception, logic, and mathematics." In William Day and Victor J. Krebs (eds.), *Seeing Wittgenstein Anew: New Essays on Aspect-Seeing*. Cambridge University Press: New York.

[Gärdenfors 1980] Gärdenfors, Peter. 1980. "A pragmatic approach to explanations." *Philosophy of Science* 47, 404-423.

[Gasarch 2014] Gasarch, William. 2014. Classifying problems into complexity classes. In *Advances in Computers, Volume 95*, ed. Atif Memon, 239-292. Cambridge, MA: Academic Press.

[Gentner & Jeziorski 1993] Gentner, Dedre and Michael Jeziorski. 1993. "The shift from metaphor to analogy in Western science." in A. Ortony (ed.), *Metaphor and Thought* (second edition), Cambridge University Press: Cambridge, 447- 480.

[Giaquinto 2007] Giaquinto, Marcus. 2007. *Visual Thinking in Mathematics: An Epistemological Study*. Oxford University Press: New York.

[Gillett 2016] Gillett, Carl. 2016. *Reduction and Emergence in Science and Philosophy*. Cambridge: Cambridge University Press.

[Gödel 1983] Gödel, Kurt. 1983. "What is Cantor's continuum problem?" In Paul Benacerraf and Hilary Putnam (eds.), *Philosophy of Mathematics*. Cambridge University Press: Cambridge.

[Gowers 2008] Gowers, Timothy. 2008. "π." In Timothy Gowers, June Barrow-Green and Imre Leader (eds.), *The Princeton Companion to Mathematics*. Princeton University Press: Princeton.

[Gowers et al. 2008] Gowers, Timothy, June Barrow-Green and Imre Leader (eds.). 2008. *The Princeton Companion to Mathematics*. Princeton: Princeton University Press.

[Gowers 2011] Gowers, Timothy. November 13, 2011. Why isn't the fundamental theorem of arithmetic obvious? *Gowers's Weblog*, URL = <https://gowers.wordpress.com/2011/11/13/why-isnt-the-fundamental-theorem-of-arithmetic-obvious/>.

[Grimm 2010] Grimm, Stephen R. 2010. "The goal of explanation." *Studies in History and Philosophy of Science* 41, 337-344.

[Gullberg 1997] Gullberg, Jan. 1997. *Mathematics: From the Birth of Numbers*. Norton: New York.

[Hafner & Mancosu 2005] Hafner, Johannes and Paolo Mancosu. 2005. "The varieties of mathematical explanation." In Paolo Mancosu, Klaus Frovin Jørgensen and Stig Andur Pedersen (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Berlin: Springer, 215-250.

[Hallett 1984] Hallett, Michael. 1984. *Cantorian Set Theory and Limitation of Size*. Oxford University Press: New York.

[Hanna 1990] Hanna, Gila. 1990. "Some pedagogical aspects of proof." *Interchange* 21, 6-13.

[Harris 1992] Harris, Joe. 1992. "Developments in algebraic geometry." In Felix E. Browder (ed.), *Mathematics into the Twenty-first Century: 1988 Centennial Symposium, August 8-12*. American Mathematical Society: Providence, RI.

[Hempel 1958] Hempel, Carl G. 1958. "The theoretician's dilemma." In H. Feigl, M. Scriven, and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. II. University of Minnesota Press: Minneapolis, pp. 37–98.

[Hempel & Oppenheim 1948] Hempel, Carl G. and Paul Oppenheim. 1948. Studies in the logic of explanation. *Philosophy of Science* 15, 135-175.

[Herald 2010] Herald, Marybeth. 2010. "Situations, frames, and stereotypes: Cognitive barriers on the road to nondiscrimination." *Michigan Journal of Gender and Law* 17, 39-55.

[Kemp & Mras 2016] Kemp, Gary and Gabriele M. Mras (eds.). 2016. *Wollheim, Wittgenstein, and Pictorial Representation: Seeing-as and Seeing-in*. Routledge: New York.

[Keshav 2012] Keshav, Srinivasan. 2012. *Mathematical Foundations of Computer Networking*. Upper Saddle River, NJ: Addison-Wesley.

[Khalifa 2012] Khalifa, Kareem. 2012. "Inaugurating understanding or repackaging explanation?" *Philosophy of Science* 79, 15-37.

[Khovanskii 2014] Khovanskii, Askold. 2014. *Topological Galois Theory: Solvability and Unsolvability of Equations in Finite Terms.* Heidelberg: Springer.

[Kim 1994] Kim, Jaegwon. 1994. "Explanatory knowledge and metaphysical dependence." *Philosophical Issues* 5, 51-69.

[Kitcher 1978] Kitcher, Philip. 1978. "The plight of the Platonist." *Noûs* 12, 119-136.

[Klein 2009] Klein, Colin. 2009. "Reduction without reductionism: A defence of Nagel on connectability." *Philosophical Quarterly* 59, 39-53.

[Kleiner 1989] Kleiner, Israel. 1989. "Evolution of the function concept: A brief survey." *The College Mathematics Journal* 20, 282-300.

[Kollár 2008] Kollár, János. 2008. "Algebraic geometry." In Timothy Gowers, June Barrow-Green and Imre Leader (eds.), *The Princeton Companion to Mathematics.* Princeton University Press: Princeton, NJ.

[Konvisser 1986] Konvisser, Marc W. 1986. *Elementary Linear Algebra with Applications*. Ardsley House: New York.

[Koslicki 2012] Koslicki, Kathrin. 2012. "Varieties of ontological dependence." In Fabrice Correia and Benjamin Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press: New York.

[Ladyman & Presnell 2014] Ladyman, James and Stuart Presnell. 2014. "Does Homotopy Type Theory provide a foundation for mathematics?" Unpublished MS.

[Lange 2009] Lange, Marc. 2009. Why proofs by mathematical induction are generally not explanatory. *Analysis* 69, 203-211.

[Lange 2010] Lange, Marc. 2010. "What are mathematical coincidences (and why does it matter)?" *Mind* 119, 307-340.

[Lange 2014] Lange, Marc. 2014. "Aspects of mathematical explanation: Symmetry, unity, and salience." *Philosophical Review* 123, 485-531.

[Lange 2015] Lange, Marc. 2015. "Explanation, existence and natural properties in mathematics—A case study: Desargues' theorem." *Dialectica* 69, 435-472.

[Lange 2016] Lange, Marc. 2016. *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. New York: Oxford University Press.

[Laptev & Rozenfel'd 1996] Laptev, B.L. and B.A. Rozenfel'd. 1996. "Geometry." In A.N. Kolmogorov and A.P. Yushkevich (eds.), *Mathematics of the 19th Century: Geometry, Analytic Function Theory*. Translated by Roger Cooke. Birkhäuser Verlag: Basel.

[Lewis 1973] Lewis, David. 1973. "Causation." *Journal of Philosophy* 70, 556-567.

[Linnebo & Pettigrew 2011] Linnebo, Øystein and Richard Pettigrew. 2011. "Category theory as an autonomous foundation." *Philosophia Mathematica* 19, 227-254.

[Lipton 2011] Lipton, Peter. 2011. "Mathematical understanding." In John Polkinghorne (ed.), *Meaning in Mathematics*. Oxford University Press: New York.

[Lombrozo 2006] Lombrozo, Tania. 2006. "The structure and function of explanations." *Trends in Cognitive Sciences* 10, 464-470.

[Lombrozo 2012] Lombrozo, Tania. 2012. "Explanation and abductive inference." In Keith J. Holyoak and Robert G. Morrison (eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press: New York.

[López Peña & Lorscheid 2011] López Peña, Javier and Oliver Lorscheid. 2011. "Mapping $\mathbb{F}_1$-land: An overview of geometries over the field with one element." In Caterina Consani and Alain Connes (eds.), *Noncommutative Geometry, Arithmetic, and Related Topics: Proceedings of the Twenty-First Meeting of the Japan-U.S. Mathematics Institute*. The Johns Hopkins University Press: Baltimore.

[Mach 1960] Mach, Ernst. 1960. *The Science of Mechanics: A Critical and Historical Account of Its Development* (6th edition). Translated by Thomas J. McCormack. Open Court: La Salle, IL.

[Maddy 1981] Maddy, Penelope. 1981. "Sets and numbers." *Noûs* 15, 495-511.

[Maddy 1990] Maddy, Penelope. 1990. *Realism in Mathematics*. Oxford University Press: New York.

[Maddy 2011] Maddy, Penelope. 2011. "Set theory as a foundation." In Giovanni Sommaruga (ed.), *Foundational Theories of Classical and Constructive Mathematics.* Springer: Dordrecht.

[Maddy 2017] Maddy, Penelope. 2017. "Set-theoretic foundations." In Andrés Eduardo Caicedo, James Cummings, Peter Koellner and Paul B. Larson (eds.), *Contemporary Mathematics, Volume 690: Foundations of Mathematics*. American Mathematical Society: Providence, RI.

[Mancosu 1999] Mancosu, Paolo. 1999. "Bolzano and Cournot on mathematical explanation." *Revue d'Histoire des Sciences* 52, 429-455.

[Mancosu 2000] Mancosu, Paolo. 2000. "On mathematical explanation." In Emily Grosholz and Herbert Breger (eds.), *The Growth of Mathematical Knowledge*, Dordrecht: Kluwer, 103-119.

[Mancosu 2001] Mancosu, Paolo. 2001. "Mathematical explanation: Problems and prospects." *Topoi* 20, 97-117.

[Mancosu 2008] Mancosu, Paolo. 2008. "Mathematical explanation: Why it matters." In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*, New York: Oxford University Press, 134-150.

[Mancosu 2015] Mancosu, Paolo. 2011. Explanation in mathematics. *The Stanford Encyclopedia of Philosophy* (Summer 2015 edition), ed. Edward N. Zalta, URL = <https://plato.stanford.edu/archives/sum2015/entries/mathematics-explanation/>.

[Maor 2007] Maor, Eli. 2007. *The Pythagorean Theorem: A 4,000-Year History*. Princeton: Princeton University Press.

[Mayberry 1994] Mayberry, John. 1994. "What is required of a foundation for mathematics?" *Philosophia Mathematica* 3, 16-35.

[McLarty 2008] McLarty, Colin. 2007. "The rising sea: Grothendieck on simplicity and generality". In Jeremy J. Gray and Karen Hunger Parshall (eds.), *Episodes in the History of Modern Algebra (1800-1950)*. American Mathematical Society: Providence, RI.

[Medvedev 1998] Medvedev, F.A. 1998. "Nonstandard analysis and the history of classical analysis." Translated by Abe Shenitzer. *American Mathematical Monthly* 105, 659-664.

[Moschovakis 2006] Moschovakis, Yiannis. 2006. *Notes on Set Theory*, 2nd edition. Springer: New York.

[Nadis & Yau 2013] Nadis, Steve and Shing-Tung Yau. 2013. *A History in Sum: 150 Years of Mathematics at Harvard (1825-1975)*. Cambridge: Harvard University Press.

[Nagel 1961] Nagel, Ernest. 1961. *The Structure of Science*. Routledge and Kegan Paul: London.

[Newman 2012] Newman, Stephen C. 2012. *A Classical Introduction to Galois Theory*. Hoboken, NJ: John Wiley & Sons.

[Niebergall 2000] Niebergall, Karl-Georg. 2000. "On the logic of reducibility: Axioms and examples." *Erkenntnis* 53, 27-61.

[Pantea et al. 2014] Pantea, Casian, Ankur Gupta, James B. Rawlings and Gheorghe Craciun. 2014. The QSSA in chemical kinetics: as taught and practiced. In *Discrete and Topological Models in Molecular Biology*, eds. Nataša Jonoska and Masahico Saito, 419-442. Berlin: Springer.

[Perrin 2008] Perrin, Daniel. 2008. *Algebraic Geometry: An Introduction*. Translated by Catriona Maclean. Springer-Verlag: London.

[Pesic 2013] Pesic, Peter. 2013. *Abel's Proof: An Essay on the Sources and Meaning of Mathematical Unsolvability*. Cambridge, MA: MIT Press.

[Pincock 2015] Pincock, Christopher. 2015. "The unsolvability of the quintic: A case study in abstract mathematical explanation." *Philosophers' Imprint* 15 (3), 1-19.

[Potter 2004] Potter, Michael D. 2004. *Set Theory and Its Philosophy: A Critical Introduction.* Oxford University Press: New York.

[Povich 2016] Povich, Mark. 2016. "Minimal models and the generalized ontic conception of scientific explanation." *British Journal for the Philosophy of Science* 0, 1-21.

[Quine 1960] Quine, Willard Van Orman. 1960. *Word and Object*. MIT Press: Cambridge, MA.

[Rantala 1992] Rantala, Veikko. 1992. "Reduction and explanation: Science vs. mathematics." In Javier Echeverria, Andoni Ibarra and Thomas Mormann (eds.), *The Space of Mathematics: Philosophical, Epistemological, and Historical Explorations*, Walter de Gruyter: New York, 47-59.

[Rav 1999] Rav, Yehuda. 1999. Why do we prove theorems? *Philosophia Mathematica* 7, 5-41.

[Reck 2016] Reck, Erich. 2016. "Dedekind's Contributions to the Foundations of Mathematics." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), URL = <https://plato.stanford.edu/archives/win2016/entries/dedekind-foundations/>.

[Resnik & Kushner 1987] Resnik, Michael D. and David Kushner. 1987. "Explanation, independence and realism in mathematics." *British Journal for the Philosophy of Science* 38, 141-158.

[Rice 2015] Rice, Collin C. 2015. "Moving beyond causes: Optimality models and scientific explanation." *Noûs* 49, 589-615.

[Robinson 1974] Robinson, Abraham. 1974. *Non-standard Analysis*. Princeton University Press: Princeton.

[Robinson 2003] Robinson, Derek J.S. 2003. *An Introduction to Abstract Algebra*. Berlin: Walter de Gruyter.

[Ruben 1990] Ruben, David-Hillel. 1990. *Explaining Explanation*. Routledge: New York.

[Saatsi & Pexton 2013] Saatsi, Juha and Mark Pexton. 2013. "Reassessing Woodward's account of explanation: Regularities, counterfactuals, and noncausal explanations." *Philosophy of Science* 80, 613-624.

[Salmon 1989] Salmon, Wesley. 1989. *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.

[Sandborg 1998] Sandborg, David. 1998. "Mathematical explanation and the theory of why-questions." *British Journal for the Philosophy of Science* 49, 603-624.

[Schröder 1895] Schröder, Ernst. 1895. *Vorlesungen über die Algebra der Logik*, vol. 3. Teubner: Leipzig.

[Scriven 1962] Scriven, Michael. 1962. Explanations, predictions, and laws. In *Scientific Explanation, Space, and Time* (Minnesota Studies in the Philosophy of Science: Vol. 3), eds. H. Feigl and G. Maxwell, 170–230. Minneapolis: University of Minnesota Press.

[Seymour 2016] Seymour, Paul. 2016. Hadwiger's conjecture. In *Open Problems in Mathematics*, eds. J. Nash and M. Rassias, 417-437. Berlin: Springer.

[Shapiro 2000] Shapiro, Stewart. 2000. "Set-theoretic foundations." *The Proceedings of the Twentieth World Congress of Philosophy* 6, 183-196.

[Sierpinska 1994] Sierpinska, Anna. 1994. *Understanding in Mathematics.* Falmer Press: Washington, D.C.

[Steiner 1978a] Steiner, Mark. 1978. Mathematical explanation. *Philosophical Studies* 34, 135-151.

[Steiner 1978b] Steiner, Mark. 1978. Mathematics, explanation, and scientific knowledge. *Noûs* 12, 17-28.

[Steinhart 2002] Steinhart, Eric. 2002. “Why numbers are sets.” *Synthese* 133, 343-361.

[Stetkær 2013] Stetkær, Henrik. 2013. *Functional Equations on Groups*. Singapore: World Scientific Publishing.

[Stillwell 1995] Stillwell, John. 1995. Eisenstein’s footnote. *Mathematical Intelligencer* 17, 58-62.

[Strevens 2013] Strevens, Michael. 2013. “No understanding without explanation.” *Studies in History and Philosophy of Science* 44, 510-515.

[Tao 2008] Tao, Terence. 2008. “The Fourier transform.” In Timothy Gowers, June Barrow-Green and Imre Leader (eds.), *The Princeton Companion to Mathematics*. Princeton University Press: Princeton, NJ.

[Tao 2012] Tao, Terence. September 25, 2012. A partial converse to Bezout’s theorem. *What’s New*. Retrieved from <https://terrytao.wordpress.com/tag/bezouts-theorem/>.

[Tao 2015] Tao, Terence. December 28, 2015. “Polymath proposal: explaining identities for irreducible polynomials.” *The Polymath Blog*. Retrieved from <http://polymathprojects.org/2015/12/28/polymath-proposal-explaining-identities-for-irreducible-polynomials/>.

[Tappenden 2005] Tappenden, Jamie. 2005. “Proof style and understanding in mathematics I: Visualization, unification and axiom choice.” In P. Mancosu, K. Jørgensen and S. Pedersen (eds.), *Visualization, Explanation and Reasoning Styles in Mathematics*, Springer: Berlin.

[Taylor 1993] Taylor, R. Gregory. 1998. “Zermelo, reductionism, and the philosophy of mathematics.” *Notre Dame Journal of Formal Logic* 34, 539-563.

[Turner 2013] Turner, Stephen. 2013. "Where explanation ends: Understanding as the place the spade turns in the social sciences." *Studies in History and Philosophy of Science* 44, 532-538.

[Tversky & Kahneman 1981] Tversky, Amos and Daniel Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science* 211, 453-458.

[Vakil 2015] Vakil, Ravi. 2015. *The Rising Sea: Foundations of Algebraic Geometry*. Retrieved from <http://math.stanford.edu/~vakil/216blog/FOAGoct2415public.pdf>.

[van Fraassen 1980] van Fraassen, Bas C. 1980. *The Scientific Image*. Oxford: Oxford University Press.

[van Riel 2011] van Riel, Raphael. "Nagelian reduction beyond the Nagel model." *Philosophy of Science* 78, 353-375.

[Warner 1990] Warner, Seth. 1990. *Modern Algebra*. Dover: Mineola, NY.

[Weber & Frans 2017] Weber, Erik and Joachim Frans. 2017. "Is mathematics a domain for philosophers of explanation?" *Journal for General Philosophy of Science* 48, 125-142.

[Weber & Verhoeven 2002] Weber, Erik and Liza Verhoeven. 2002. "Explanatory proofs in mathematics." *Logique & Analyse* 179-180, 299-307.

[White 1974] White, Nicholas P. 1974. "What numbers are." *Synthese* 27, 111-124.

[Wiener 1914] Wiener, Norbert. 1914. "A simplification of the logic of relations." *Proceedings of the Cambridge Philosophical Society* 17, 387-390.

[Wittgenstein 2009] Wittgenstein, Ludwig. 2009. *Philosophical Investigations* (fourth edition). Translated by G.E.M. Anscombe, P.M.S. Hacker and Joachim Schulte. Wiley-Blackwell: Oxford.

[Woit 2006] Woit, Peter. 2006. *Not Even Wrong: The Failure of String*

Theory And the Search for Unity in Physical Law. Basic Books: New York.

[Woodward 2014] Woodward, James. 2014. "Scientific explanation." In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017 edition), <https://plato.stanford.edu/archives/spr2017/entries/scientific-explanation/>.

[Wright 2012] Wright, Cory. 2012. "Mechanistic explanation without the ontic conception." *European Journal for the Philosophy of Science* 2, 375-394.

[Zelcer 2013] Zelcer, Mark. 2013. "Against mathematical explanation." *Journal for General Philosophy of Science* 44, 173-192.

[Zermelo 1909a] Zermelo, Ernst. 1909. "Sur les ensembles finis et le principe de l'induction complète." *Acta Mathematica* 32, 185-193.

[Zermelo 1909b] Zermelo, Ernst. 1909. "Über die Grundlagen der Arithmetik." In G. Castelnuovo (ed.), *Atti del IV Congresso Internazionale dei Matematici* (vol. 2), 8-11. Academia dei Lincei: Rome.

Vita

Education

University of Illinois-Chicago

- Ph.D., Philosophy (dissertation defended August 2017; degree to be conferred December 2017)
- M.S. candidate, Pure Mathematics (expected completion Spring 2017)

Canisius College, Buffalo, New York

- B.A. (Honors), magna cum laude, Philosophy

Employment

North Central College, Naperville, Illinois

- Adjunct faculty, Fall 2017

Areas of Specialization

Philosophy of mathematics, philosophy of science, metaphysics

Areas of Competence

Epistemology, aesthetics, logic

Publications

- "Mathematical Explanation Beyond Explanatory Proof", *British Journal for the Philosophy of Science*, forthcoming
- "Arithmetic, Set Theory, Reduction and Explanation", *Synthese*, forthcoming (DOI: 10.1007/s11229-017-1450-8)
- "Explicitism about Truth in Fiction", *British Journal of Aesthetics* 56, 53-65 (2016)

Presentations

- To comment on Conor Mayo-Wilson, "Formalization, Reliabilism, and Justification", at the Central Division meeting of the American Philosophical Association, February 2018
- "Grounding, dependence and mathematical explanation", at the Eastern Division meeting of the American Philosophical Association, Baltimore, January 2017 (recipient of the APA's 2016 Sanders Graduate Student Award)

- “Mathematical objects as abstract artifacts”, at the Society for Exact Philosophy, Miami, May 2016
- “Intertheoretic reduction and explanation in mathematics”, at the Mathematical Aims Beyond Justification Workshop, Royal Flemish Academy of Belgium for Science and the Arts, Brussels, December 2015
- “‘Where there were two, let there now be one’: CSMP and mainstream philosophy of mathematical practice”, at the Case Studies in Mathematical Practice Workshop, Université Paris Diderot, June-July 2015
- Commented on Logan Douglass and Paul Shephard, “Confidence level invariantism”, at the 2015 Chicagoland Philosophy Graduate Conference
- “Truth in fiction”, to UIC Philosophy Club, Fall 2014
- “Explicitism about truth in fiction”, at the 2013 Annual Meeting of the American Society for Aesthetics (graduate student travel award recipient)
- “Formalisms and interpretations: The case of quantum logic”, at the University of Western Ontario 2011 Philosophy of Physics, Mathematics and Logic Graduate Conference
- “Mathematical explanation, symmetry and psychology”, at the University of Western Ontario 2010 Philosophy of Physics, Mathematics and Logic Graduate Conference
- “How to understand ethical theories”, at the Michigan State University 2010 Philosophy Graduate Conference

Honors and Awards

- Excellence in Undergraduate Mentoring Award, honorable mention, UIC, 2017
- 2016 Sanders Graduate Student Award for “Grounding, Dependence and Mathematical Explanation”—a $1000 prize given to “each of the three best papers in mind, metaphysics, epistemology or ethics submitted for the annual APA Eastern Division meeting by graduate students”
- Institute for the Humanities Graduate Resident Fellowship, UIC, 2016-2017
- Dean’s Scholar dissertation fellowship, UIC, 2015-2016—the largest and most competitive fellowship awarded to UIC graduate students
- Chosen as UIC’s sole nominee for the Midwestern Association of Graduate Schools 2016 Excellence in Teaching Award
- Graduate Student Teaching Award, UIC Philosophy Department, 2015
- Junior Participant in Case Studies in Mathematical Practice Workshop, Université Paris Diderot, 2015 (with full stipend for travel and accommodations)

- Second-prize winner in 2014 British Society of Aesthetics Essay Prize Contest, for "Explicitism about Truth in Fiction"
- Ruth Marcus Award for Outstanding Graduate Student, UIC Philosophy Department, 2014
- Chancellor's Interdisciplinary Graduate Research Fellowship recipient, UIC, 2013-2014 and 2014-2015
- J. Clayton Murray, S.J. Award for Outstanding Student in Philosophy, Canisius College, 2006

Teaching Experience

At North Central College:

- PHIL 110: Ethics, Fall 2017

As primary instructor at UIC:

- PHIL 201: Theory of Knowledge, Spring 2017 & Fall 2013
- PHIL 203: Metaphysics, Fall 2016, Spring 2014 & Fall 2011
- PHIL 204: Philosophy of Science, Spring 2015
- PHIL 210: Symbolic Logic, Fall 2014 & Spring 2012
- PHIL 202: Philosophy of Psychology, Summer 2013
- PHIL 101: Introduction to Philosophy, Spring 2013
- PHIL 102: Introduction to Logic, Summer 2014, Summer 2012 & Fall 2012

As teaching assistant at UIC:

- PHIL 102: Introduction to Logic, with Prof. Jon Jarrett, Spring 2011
- PHIL 110: Philosophy of Love and Sex, with Prof. Anne Eaton, Fall 2010
- PHIL 115: Philosophy of Death and Dying, with Prof. Neal Grossman, Spring 2010
- PHIL 100: Introduction to Philosophy, with Prof. Georgette Sinkler, Fall 2009

Service

- Served as referee for *Synthese* and the *British Journal of Aesthetics*
- Co-founded and co-organized the Chicagoland Philosophy Graduate Conference, 2015-2016
- Revived and helped run the UIC undergraduate Philosophy Club, 2015-2016
- UIC Graduate Student Council representative, 2015-2016
- Graduate Student Representative at Philosophy Department meetings, 2013-2014 and 2014-2015
- Co-organized Philosophy Department Works in Progress talk series, 2013-2014