Aspects of Mathematical Explanation: Symmetry, Unity, and Salience

Marc Lange
University of North Carolina at Chapel Hill

1. Introduction

Compared to *scientific* explanation, with which philosophy of science has been seriously engaged for at least six decades, explanation *in mathematics* has been little explored by philosophers.[1] Its neglect is remarkable. Mathematical proofs that explain *why* some theorem holds were distinguished by ancient Greek mathematicians from proofs that merely establish *that* some theorem holds (Harari 2008), and this distinction has been invoked in various ways throughout the history of mathematics.[2] Fortunately, mathematical explanation has now begun to receive greater philosophical attention. As Mancosu (2008, 134) remarks, the topic's "recent revival in the analytic literature is a welcome addition to the philosophy of mathematics."

1. In the literature on explanation, the existence of explanations in mathematics is often acknowledged; for one example per decade, see Nagel 1961, 16; Rescher 1970, 4; Sober 1984, 96; Smart 1990, 2; and Psillos 2002, 2. But explanation in mathematics is typically not examined at length; in the five works just cited, either it is set aside as not germane to the kind of explanation being examined, or it is given cursory treatment, or it is just ignored after being acknowledged once. Among the very few earlier papers that examine explanation in mathematics at some length are Steiner 1978a; Kitcher 1984, 1989; and Resnik and Kushner 1987.

2. "Explaining why the theorem holds" is just explaining the theorem. It is distinct from explaining why we should (or do) believe the theorem. This distinction is familiar from scientific explanation.

Philosophical Review, Vol. 123, No. 4, 2014
DOI 10.1215/00318108-2749730

Published by Duke University Press

One reason that mathematical explanation has received relatively scant philosophical attention may be the temptation among philosophers to believe that when mathematicians talk about a given proof's "explanatory power," they are merely gesturing toward some aesthetically attractive quality that the proof possesses—a quality that is very much in the eye of the beholder (like a proof's being interesting, understandable, or witty). However, no such suspicion is seriously entertained with regard to *scientific* explanation,[3] and we should demand some reason why mathematical explanation deserves to be regarded differently. Of course, some aspects of explanation (whether in science or math) may well depend on the audience. Many philosophers accept that in science, one fact may explain another only in a given context—that is, only in light of the audience's interests, which help to make certain facts salient and certain styles of argument relevant.[4] But as Kitcher and Salmon (1987) have convincingly argued, this interest relativity in scientific explanation does not mean that "anything goes"—that an explanation in a given context is just whatever strikes the given audience as explanatory or whatever facts or arguments would be of interest to them.

A key issue in the philosophical study of scientific explanation is the source of explanatory asymmetry: why does one fact explain another rather than vice versa? Some philosophers have argued that explanatory priority is generally grounded in causal priority: causes explain their effects, not the reverse.[5] Likewise, some philosophers have argued that certain laws of nature are fundamental and therefore explanatorily prior

3. Such suspicions have sometimes been seriously entertained regarding historical and other social-scientific explanations insofar as a given social scientist purports to understand another person's behavior by finding (through the operation of "Verstehen" or empathetic identification) "that it parallels some personal experience of the interpreter" (Abel 1953 [1948], 686). But we might (joining Abel) regard these suspicions as suggesting instead that genuine explanations in social science do not incorporate such empathetic identification, but rather are causal.

4. For instance, Van Fraassen (1980, 125) says that a cause of some event counts as explaining it to a given person by virtue of "being salient to [that] person because of his orientation, his interests, and various other peculiarities in the way he approaches or comes to know the problem—contextual factors." Likewise, Lewis (1986, 226–29) says that a causal explanation supplies enough of the sort of information that the recipient wants regarding the explanandum's causal history.

5. For example: "Here is my main thesis: *to explain an event is to provide some information about its causal history*" (Lewis 1986, 217); "The explanation of an event describes the 'causal structure' in which it is embedded" (Sober 1984, 96).

to other, "derivative" laws.[6] By contrast, in mathematical explanations consisting of proofs, the source of explanatory priority cannot be causal or nomological (or temporal) priority. Rather, at least part of its source would seem to be that axioms explain theorems, not vice versa. Of course, there may be several different ways to axiomatize a given branch of mathematics. Perhaps only some of these axiomatizations are correct for explanatory purposes. Or perhaps any axiomatization is equally good for explanatory purposes, but a proof's status as explanatory is relative to a given axiomatization.[7]

I will be concerned with a logically prior issue. By means of several examples, I will argue that two mathematical proofs may prove the same theorem from the same axioms, though only one of these proofs is explanatory. My goal in this essay will be to identify the ground of this distinction. Accordingly, my focus will be on the course that a given proof takes *between* its premises and its conclusion. The distinction between explanatory and nonexplanatory proofs from the same premises must rest on differences in *the way* they extract the theorem from the axioms.

Furthermore, for a theorem of the form "All Fs are G," we will see that the distinction between explanatory and nonexplanatory proofs rests on differences in the ways in which the proofs manage to extract the property of being G from the property of being F. But two proofs of the same theorem that differ in this way are nevertheless alike in a respect distinct from their both proceeding from the same axioms, namely, in their both proving the theorem by extracting G-hood from F-hood (and, in this sense, by treating the theorem as "All Fs are G" rather than, for example, as "All non-Gs are non-F"). As we will see, a proof's explanatory power depends on the theorem's being understood in terms of

6. See, for example, Mill 1872, 364–69, Salmon 1989, 14–19.

7. Here, perhaps, is one way in which a proof under two axiomatizations could be explanatory under one axiomatization but not under the other. Desargues's theorem cannot be proved from the other axioms of plane (that is, two-dimensional) projective geometry. If we add as an additional axiom that the two-dimensional projective space can be embedded in a space of three (or more) dimensions, then Desargues's theorem can be given a proof that is regarded as explanatory (by, for instance, Stankova [2004, 175], who contrasts it with a proof that "doesn't give us a clue *really why* Desargues' Theorem works"). However, we could instead simply add Desargues's theorem directly as a further axiom of projective geometry, as is often done—for example, by Stillwell (2005, 117) and Gray (2007, 347). Having done so, we can derive three-dimensional embeddability and from that conclusion, in turn, we can reverse course by using the above proof to deduce Desargues's theorem. But then this proof does not explain why Desargues's theorem holds, since Desargues's theorem is an axiom. For more, see Lange, n.d.

7	8	9
4	5	6
1	2	3

Figure 1. A calculator keyboard

a "setup" or "problem" (involving F's instantiation) and a "result" or "answer" (G's instantiation). In addition, we will see that the manner in which a theorem is expressed not only involves a "setup" and "result," but also may call attention to a particular feature of the theorem, where that feature's salience helps to determine what a proof must do in order to explain why the theorem holds.

To get started, let's look briefly at an example of an explanation in mathematics (to which I will return in section 6). Take an ordinary calculator keyboard (fig. 1):

We can form a six-digit number by taking the three digits on any row, column, or main diagonal on the keyboard in forward and then in reverse order. For instance, the bottom row taken from left to right, and then right to left, yields 123321. There are sixteen such "calculator numbers" (321123, 741147, 951159, . . .). As you can easily verify (with a calculator!), every calculator number is divisible by 37. But a proof that checks each of the calculator numbers separately does not explain why every calculator number is divisible by 37. Compare this case-by-case proof to the following proof:

> The three digits from which a calculator number is formed are three integers a, $a + d$, and $a + 2d$ in arithmetic progression. Take any number formed from three such integers in the manner of a calculator number—that is, any number of the form $10^5 a + 10^4(a + d) + 10^3(a + 2d) + 10^2(a + 2d) + 10(a + d) + a$. Regrouping, we find this equal to $a(10^5 + 10^4 + 10^3 + 10^2 + 10 + 1) + d(10^4 + 2.10^3 + 2.10^2 + 10) = 111111a + 12210d = 1221(91a + 10d) = (3 \times 11 \times 37)\ (91a + 10d)$.[8]

This proof explains why all of the calculator numbers are divisible by 37; as a mathematician says, this proof (unlike the case-by-case proof) reveals the result to be "no coincidence" (Nummela 1987, 147; see also Lange

8. The example appears in an unsigned "gleaning" on page 283 of the December 1986 issue of the *Mathematical Gazette*. Roy Sorensen (n.d.) called this lovely example to my attention. He also cited Nummela 1987, from which this explanatory proof comes.

2010). Later I will propose an account of what makes this proof but not the case-by-case proof explanatory.

In section 2, I will present another example in which two proofs of the same theorem differ (according to a mathematician) in explanatory power. In section 3, I will try to spell out the difference between these proofs that is responsible for their difference in explanatory power. Roughly speaking, I will suggest that one proof is explanatory because it exploits a symmetry in the problem—a symmetry of the same kind as the symmetry that initially struck us in the result being explained. In sections 4 and 5, I will present several other, diverse examples where different proofs of the same theorem have been recognized as differing in explanatory power. In each example, I will suggest that this difference arises from a difference in whether the proofs exploit a symmetry in the problem that is like a striking symmetry in the theorem being proved. These cases will also illustrate how "brute force" proofs fail to explain and why some auxiliary constructions but not others are "artificial." In section 6, I will generalize my proposal to explanations that do not exploit symmetries. I will argue that in many cases, at least, what it means to ask for a proof that explains is to ask for a proof that exploits a certain kind of feature in the problem: the same kind of feature that is outstanding in the result being explained. The distinction between proofs explaining why some theorem holds and proofs merely establishing that the theorem holds arises only when some feature of the result is salient. In section 7, I will briefly contrast my account of mathematical explanation with those of Steiner (1978a), Kitcher (1984, 1989), and Resnik and Kushner (1987). Finally, in section 8, I will suggest that some scientific explanations operate in the same way as the mathematical explanations I have examined.

Although my concern will be the distinction between explanatory and nonexplanatory *proofs* in mathematics, I will not contend that all mathematical explanations consist of proofs. Indeed, I will give some examples of mathematical explanations that are not proofs.[9]

9. Of course, mathematical facts are often used to explain why certain contingent facts hold. Sometimes mathematical facts may even play the central role in explaining some physical fact. For recent discussions, see Baker 2005 and Lange 2013; see also Steiner 1978b. But these are not "mathematical explanations" of the kind that I will be discussing, which have mathematical theorems rather than physical facts as their target.

In a conversation, we might "explain" why (or how) some mathematical proof works (by, for example, making more explicit the transitions between steps). A textbook might "explain" how to multiply matrices. A mathematics popularizer might "explain" an

I will be paying close attention to particular examples of explanatory and nonexplanatory proofs. Proposed accounts of *scientific* explanation have long been tested against certain canonical examples. Various arguments from the history of science are paradigmatically explanatory, such as Darwin's explanations of various facts of biogeography and comparative anatomy, Wegener's explanation of the correspondence between the South American and African coastlines, and Einstein's explanation of the equality of inertial and gravitational mass. Any promising comprehensive theory of scientific explanation must deem these arguments to be explanatory (as far as we now know). By the same token, there is widespread agreement that various other arguments are not explanatory—examples so standard that I need only give their familiar monikers, such as "the flagpole," "the eclipse," "the barometer," and "the hexed salt" (Salmon 1989, 46–50). Although there are some controversial cases (such as "explanations" of the dormitive-virtue variety), philosophers who defend rival accounts of scientific explanation largely agree on the phenomena that they are trying to save.

Alas, the same cannot be said when it comes to mathematical explanation. The philosophical literature contains few, if any, canonical examples of mathematical explanation drawn from mathematical practice (or of mathematical proofs that are canonically nonexplanatory). Accordingly, I think it worthwhile to try to supply some promising candidates—especially examples where the mathematics is relatively simple and the explanatory power (or impotence) of the proofs is easily appreciated. I will be particularly interested in examples that mathematicians themselves have characterized as explanatory (or not). In this way, I will be joining others—such as Hafner and Mancosu (2005) and Tappenden (2005)—who have recently offered examples to show that explanation is an important element of mathematical practice.[10]

---

obscure theorem by unpacking it. However, none of these is the kind of "mathematical explanation" with which I will be concerned. None involves explaining why some result holds—just as Hempel (2001 [1963], 80) pointed out that an account of scientific explanation does not aim to account for what I do when I use gestures to "explain" to a Yugoslav garage mechanic how my car has been misbehaving. I am also not concerned with historical or psychological explanations of why mathematicians held various beliefs or how a given mathematician managed to make a certain discovery.

10. To avoid the corrupting influence of philosophical intuitions, I have also tried to use examples from workaday mathematics rather than from logic, set theory, and other parts of mathematics that have important philosophical connections. But by focusing on proofs that mathematicians themselves recognize as explanatory, I do not mean to suggest

As we will see, mathematicians have often distinguished proofs that explain why some theorem holds from proofs that merely establish that it holds. For instance, in the *Port-Royal Logic* of 1662, Pierre Nicole and Antoine Arnauld characterized indirect proof (that is, proof of p by showing that $\sim p$ implies a contradiction) as "useful" but nonexplanatory: "Such Demonstrations constrain us indeed to give our Consent, but no way clear our Understandings, which ought to be the principal End of Sciences: for our Understanding is not satisfied if it does not know not only that a thing is, but why it is? which cannot be obtain'd by a Demonstration reducing to Impossibility" (Nicole and Arnauld 1717, 422 [part 4, chapter 9]). Nicole and Arnauld evidently took explanation to be as important in mathematics as it is in science. More recently, the mathematician William Byers (2007, 337) has characterized a "good" proof as "one that brings out clearly the reason why the result is valid." Likewise, researchers on mathematics education have recently argued empirically that students who have proved and are convinced of a mathematical result often still want to know why the result is true (Mudaly and de Villiers 2000), that students assess alternative proofs for their "explanatory power" (Healy and Hoyles 2000, 399), and that students expect a "good" proof "to convey an insight into *why* the proposition is true" even though explanatory power "does not affect the *validity* of a proof" (Bell 1976, 24). However, none of this work investigates what it is that makes certain proofs but not others explanatory. This question will be my focus.

that philosophers must unquestioningly accept the verdicts of mathematicians. Indeed, some mathematicians, such as Gale (1990), have denied that there is any distinction between explanatory and nonexplanatory proofs. (Some philosophers have agreed; see, for example, Grosholz 2000, 81. However, Gale [1991, 41] later changed his mind.) But just as an account of scientific explanation should do justice to scientific practice (without having to fit every judgment of explanatory power made by every scientist), so an account of mathematical explanation should do justice to mathematical practice. Regarding the examples that I will discuss, I have found the judgments made by working mathematicians of which proofs do (and do not) explain to be widely shared and also to be easily appreciated by nonmathematicians. It is especially important that an account of mathematical explanation fit such cases.

2. Zeitz's Biased Coin: A Suggestive Example of Mathematical Explanation

Consider this problem:[11]

> A number p between 0 and 1 is generated randomly so that there is an equal chance of the generated number's falling within any two intervals of the same size inside [0,1]. Next a biased coin is built so that p is its chance of landing heads. The coin is then flipped 2000 times. What is the chance of getting exactly 1000 heads?

How do the chances of getting various total numbers of heads, for a given bias p, combine with the chances of the coin's bias p being extreme or closer to fair so as to yield the overall chance of getting exactly 1000 heads? The mathematician Paul Zeitz gives the answer: "The amazing answer is that the probability is 1/2001. Indeed, it doesn't matter how many heads we wish to see—for any [integer r] between 0 and 2000, the probability that r heads occur is 1/2001" (Zeitz 2000, 2). That is, each of the 2001 possible outcomes (from 0 heads to 2000 heads) has the same probability. That is remarkable. It prompts us to ask, "Why is that?" (There is nothing special about 2000 tosses; the analogous result holds for any number n of tosses.)

Here is an elaboration of one proof that Zeitz sketches. (I give all of the gory details, but you may safely skim over them, if you wish.)

> If you flip a coin n times, where p is the chance of getting a head on any single flip, then the chance of getting a *particular* sequence of r heads and $(n-r)$ tails is $p^r(1-p)^{n-r}$. Now let's consider all of the sequences with exactly r heads. There are $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ different ways of arranging r heads and $(n-r)$ tails. So the chance of getting exactly r heads is $\binom{n}{r}p^r(1-p)^{n-r}$. This is the chance for a given value p of the coin's bias. But the coin's bias can assume any value from 0 to 1, so to arrive at the total chance of getting exactly r heads, we must take into account the chance of getting exactly r heads for each possible value of the coin's bias. The total chance of getting exactly r heads is the sum, taken over all possible biases p for the coin (ranging from 0 to 1), of the chance of getting r heads if the coin's bias is p, multiplied by the chance dp that p

11. From the Bay Area Math Meet, San Francisco, April 29, 2000.

is the coin's bias. This sum is an integral:

$$\int_0^1 \binom{n}{r} p^r(1-p)^{n-r}dp = \binom{n}{r}\int_0^1 p^r(1-p)^{n-r}dp.$$

The easiest way to tackle this integral is to use the textbook technique of "integration by parts," according to which $\int u\,dv = uv - \int v\,du$. To apply this formula, we express $\int p^r(1-p)^{n-r}dp$ in the form $\int u\,dv$ by letting $u = (1-p)^{n-r}$ and $dv = p^r dp$. It follows by routine differentiation of u that $du = -(n-r)(1-p)^{n-r-1}dp$ and from routine antidifferentiation of dv that $v = \frac{p^{r+1}}{r+1}$. Plugging all of this into the formula for integration by parts, we find

$$\int_0^1 p^r(1-p)^{n-r}dp = (1-p)^{n-r}\frac{p^{r+1}}{r+1}\bigg|_0^1 + \frac{n-r}{r+1}\int_0^1 p^{r+1}(1-p)^{n-r-1}dp.$$

The first term on the right side equals zero at both $p = 0$ and $p = 1$, so that term disappears. We can leave behind to be gathered later the coefficient $\frac{n-r}{r+1}$ preceding the integral in the second term on the right side. As for that integral, it takes the same form as the integral on the left side—merely with the first exponent increased by 1 and the second exponent decreased by 1. So we can in the same way apply integration by parts to this integral, yielding the same increase and decrease (respectively) by 1 in the two exponents:

$$\int_0^1 p^{r+1}(1-p)^{n-r-1}dp = \frac{n-r-1}{r+2}\int_0^1 p^{r+2}(1-p)^{n-r-2}dp.$$

Again leaving behind the coefficient $\frac{n-r-1}{r+2}$ on the right side to be gathered later, we can repeatedly integrate by parts until $(1-p)$'s exponent eventually decreases to 0. The remaining integral is much simpler to solve than its predecessors:

$$\int_0^1 p^n(1-p)^0 dp = \int_0^1 p^n dp = \frac{p^{n+1}}{n+1}\bigg|_0^1 = \frac{1}{n+1}.$$

Having finally gotten rid of all of the integrals, we can go back and gather the various coefficients we left behind that were produced by each of the integrations by parts:

$$\int_0^1 p^r(1-p)^{n-r}dp = \left(\frac{n-r}{r+1}\right)\left(\frac{n-r-1}{r+2}\right)\cdots\left(\frac{1}{n}\right)\left(\frac{1}{n+1}\right)$$
$$= \left(\frac{1}{n+1}\right)\frac{1}{\frac{n(n-1)(n-2)\ldots(r+2)(r+1)}{(n-r)(n-r-1)\ldots(2)}} = \left(\frac{1}{n+1}\right)\frac{1}{\binom{n}{n-r}}.$$

This result times $\binom{n}{r}$ — the coefficient of the first integral we tackled by parts—is the total chance of getting exactly r heads in our problem. But $\binom{n}{n-r} = \binom{n}{r}$ since the number of different arrangements of exactly $n-r$ tails in n coin tosses equals the number of different arrangements of exactly r heads in n tosses. So the total chance of getting exactly r heads in our problem (with $n = 2000$) is $\binom{n}{r}\left(\frac{1}{n+1}\right)\frac{1}{\binom{n}{n-r}} = \left(\frac{1}{n+1}\right)$—the result we sought.

Although this proof succeeds, Zeitz (2000, 2–4) says that it "shed[s] no real light on *why* the answer is what it is. . . . [It] magically produced the value $\left(\frac{1}{n+1}\right)$." Although Zeitz does not spell out this reaction any further, I think we can readily sympathize with it. This proof makes it seem like an accident of algebra, as it were, that everything cancels out so nicely, leaving us with just $\frac{1}{n+1}$.

Of course, nothing in math is genuinely accidental; the result is mathematically necessary. Nevertheless, until the very end, nothing in the proof suggested that every possible outcome (for a given n) would receive the same chance.[12] The result simply turns out to be independent of r, and this fact remains at least as remarkable after we have seen the

12. Regarding another method of tackling this integral (using generating functions), Zeitz (2007, 352) says, "The above proof was a thing of beauty. . . . Yet the magical nature of the argument is also its shortcoming. Its punchline creeps up without warning. Very entertaining, and very instructive in a general sense, but it doesn't shed quite enough light on this particular problem. It shows us *how* these n + 1 probabilities were uniformly distributed. But we still don't know *why*."

The proof I gave can be shortened: by integration by parts, as we saw, the total chance of r heads in n tosses is $\binom{n}{r}\int_0^1 p^r(1-p)^{n-r}dp = \binom{n}{r}\frac{n-r}{r+1}\int_0^1 p^{r+1}(1-p)^{n-r-1}dp = \binom{n}{r+1}\int_0^1 p^{r+1}(1-p)^{n-r-1}dp$, which is the total chance of $r+1$ heads in n tosses. In this way, we show that the chance is independent of r and so must be $\frac{1}{n+1}$. However above proof as it was before. I think that many of us would be inclined to suspect that there is some *reason why* the chance is the same for every possible outcome (given n)—a reason that eludes the above proof. (Notice how natural it becomes in this context to talk of "reasons why" the result holds.)

Zeitz (2000, 5) says that in contrast to the foregoing argument, the following argument allows us to "understand *why* the coin problem had the answer that it did" (his emphasis):

> Think of the outcomes of the n coin tosses as dictated by n further numbers generated by the *same* random-number generator that generated the coin's chance p of landing heads: a number less than p corresponds to a head, and a number greater than p corresponds to a tail. (The chance that the number will be less than p is obviously p—the chance of a head.) Thus, the same generator generates $n + 1$ numbers in total. The outcomes are all heads if the first number generated (p) is larger than each of the n subsequent numbers, all but one of the outcomes are heads if the first number generated is larger than all but one of the n subsequent numbers, and so forth. Now for the key point: if we were to rank the $n + 1$ generated numbers from smallest to largest, then the first generated number (p) has the same chance of being ranked first as it has of being ranked second, and likewise for any other position.
>
> (In the same way, suppose that we were going to flip a fair coin 100 times and then combine the first ten flips into one group, the second ten into another group, and so forth. Suppose that we were then going to rank the ten groups in order from the one having the most heads to the one having the fewest heads. [The order among groups having the same number of heads is decided randomly.] Then the group consisting of flips 1 through 10 stands the same chance of being ranked first on the list as it does of being ranked second as it does of being ranked third. . . . The group of flips 1 through 10 is not special; it is no more likely to be first than, say, eighth on the list. The same applies to the first generated number (p) among the $n + 1$ numbers generated in our example.)
>
> For k in $\{0, 1, 2, \ldots, n\}$, let A_k be the set of sequences of n coin-toss outcomes with exactly k heads, and let B_k be the set of sequences of $n + 1$ randomly chosen members of [0,1] in which exactly k members of the sequence are less than the first member. The probability that flipping the coin will produce some or another sequence in A_k equals the probability that a sequence of $n + 1$ randomly chosen members of [0,1] will be in B_k,

(as Tord Sjödin noted when he showed me this shortcut), that things suddenly work out here so neatly is no less "magical" than in the longer proof.

and the latter is the probability that the first chosen member (p) is ranked $(k+1)$st when the $n+1$ members of the sequence are ranked from smallest to largest.

Hence, every possibility (from 0 heads to n heads) is equally likely, so each has chance $\frac{1}{n+1}$.

In light of this proof, Zeitz (2000, 5) concludes that the result "is not unexpected, magical algebra. It is just simple, almost inevitable symmetry." I agree: this proof explains why every possible outcome has the same chance (and, therefore, why each possible outcome's chance is $\frac{1}{n+1}$). The proof explains this symmetry in the chances (namely, that each possible outcome has the same chance) by showing how it arises not from an algebraic miracle, but rather from a symmetry in the setup, namely, that when the $n+1$ generated numbers (from the same random-number generator) are listed from smallest to largest, each possible position on the list is equally likely to be occupied by the first number that was generated.[13] A symmetry in the setup accounts for the same symmetry in the chances of the possible outcomes.

In short, our curiosity was initially aroused by the symmetry of the result: that, remarkably, every possible outcome has the same chance. The first proof did not satisfy us because it failed to exploit any such symmetry in the setup. We suspected that there was a reason for the result—a hidden "evenness" in the setup that is responsible for the same "evenness" in the result. The second proof revealed the setup's hidden symmetry and thereby explained the result.

3. Explanation by Symmetry

The example of Zeitz's coin suggests the following proposal, for which I will argue. Often a mathematical result that exhibits symmetry of a certain kind is explained by a proof showing how it follows from a similar symmetry in the problem. Each of these symmetries consists of some sort of invariance under a given transformation; the same transformation is involved in both symmetries. For instance, in the example of Zeitz's coin, both symmetries involve invariance under a switch from one possible outcome (such as 1,000 heads and 1,000 tails) to any other (such as

13. Zeitz (2007, 353): "The probabilities were uniform because the numbers [generated randomly] were uniform, and thus their rankings [that is, the place of the first generated number among the others, as ranked from smallest to largest] were uniform. The underlying principle, the 'why' that explains this problem, is . . . Symmetry."

999 heads and 1,001 tails). In such a case, what makes a proof appealing to an underlying symmetry in the setup count as "explanatory"—in contrast to other proofs of the same result? Nothing beyond the fact that the result's symmetry was what drew our attention in the first place.

For instance, in the example of Zeitz's coin, had the result been some complicated, unremarkable function of n and r, then the question "Why is *that* the chance?" would probably have amounted to nothing more than a request for a proof. One proof might have been shorter, less technical, more pleasing or accessible to some audience, more elegant in some respect, or more fully spelled out than some other proof. But any proof would have counted as answering the question. According to the view that I will ultimately defend, there would then have been no distinction between a proof that explains the result and a proof that merely proves it.

However, the result's symmetry immediately struck us, and it was made further salient by the first proof we saw, since in that proof, the solution's symmetry emerged "magically" from out of the fog of algebra. Its origin was now especially puzzling. The symmetry, once having become salient, prompts the demand for an explanation: a proof that traces the result back to a similar symmetry in the problem. (There need not be any such proof; a mathematical fact may have no explanation.)[14] In light of the symmetry's salience, there is a point in asking for an explanation over and above a proof. In short, a proof that exploits the setup's symmetry is privileged as explanatory because the result's symmetry is especially striking.

My proposal predicts that mathematical practice contains many other examples where an explanation of some result is distinguished from a mere proof of it only in view of the result's exhibiting a puzzling symmetry—and where only a proof exploiting such a symmetry in the problem is recognized as explaining why the solution holds. I will now present several examples of this phenomenon.

4. A Theorem Explained by a Symmetry in the Unit Imaginary Number

Consider this theorem (first proved by d'Alembert in 1746):

14. For example, after asking *why* a given Taylor series fails to converge, Spivak (1980, 482) says, "Asking this sort of question is always dangerous, since we may have to settle for an unsympathetic answer: it happens because it happens—that's the way things are!"

> If the complex number $z = a + bi$ (where a and b are real) is a solution to $z^n + a_{n-1}z^{n-1} + \ldots + a_0 = 0$ (where the a_i are real), then z's "complex conjugate" $\bar{z} = a - bi$ is also a solution.

Why is that?

We can prove this theorem directly by evaluating $\bar{z}^n + a_{n-1}\bar{z}^{n-1} + \ldots + a_0$.

> First, we show by calculation that $\bar{z}\bar{w} = \overline{zw}$:
> Let $z = a + bi$ and $w = c + di$. Then $\bar{z}\bar{w} = (a - bi)(c - di) = ac - bd + i(-bc - ad)$ and $zw = (a + bi)(c + di) = ac - bd + i(bc + ad)$ so $\overline{zw} = ac - bd - i(bc + ad) = \bar{z}\bar{w}$.
> Hence, $\bar{z}^2 = \bar{z}\,\bar{z} = \overline{zz} = \overline{z^2}$, and likewise for all other powers. Therefore, $\bar{z}^n + a_{n-1}\bar{z}^{n-1} + \ldots + a_0 = \overline{z^n} + a_{n-1}\overline{z^{n-1}} + \ldots + a_0$.
> Now we show by calculation that $\bar{z} + \bar{w} = \overline{z + w}$:
> Let $z = a + bi$ and $w = c + di$. Then $\bar{z} + \bar{w} = (a - bi) + (c - di) = a + c + i(-b - d)$ and $\overline{z + w} = \overline{a + bi + c + di} = a + c - i(b + d) = \bar{z} + \bar{w}$.
> Thus, $\overline{z^n} + a_{n-1}\overline{z^{n-1}} + \ldots + a_0 = \overline{z^n + a_{n-1}z^{n-1} + \ldots + a_0}$, which equals $\bar{0}$ and hence 0 if z is a solution to the original equation.

Although this proof shows d'Alembert's theorem to be true, it pursues what mathematicians call a "brute force" approach. That is, it simply calculates everything directly, plugging in everything we know and grinding out the result. The striking feature of d'Alembert's theorem is that the equation's nonreal solutions all come in pairs where one member of the pair can be transformed into the other by the replacement of i with $-i$. Why does exchanging i for $-i$ in a solution still leave us with a solution? This symmetry just works out that way ("magically") in the above proof. But we are inclined to suspect that there is some reason for it. In other words, the symmetry in d'Alembert's theorem puzzles us, and in asking for the theorem's "explanation," we are seeking a proof of the theorem from some similar symmetry in the original problem—that is, a proof that exploits the setup's invariance under the replacement of i with $-i$.

The sought-after explanation is that $-i$ could play exactly the same roles in the axioms of complex arithmetic as i plays. Each has exactly the same definition: each is exhaustively captured as being such that its square equals -1. There is nothing more to i (and to $-i$) than that characterization. Of course, i and $-i$ are not equal; each is the negative of the other. But neither is intrinsically "positive," for instance, since neither is greater than (or less than) zero. They are distinct, but they are no different in their relations to the real numbers. Whatever the axioms of complex arithmetic say about one can also be truly said about the other. Since the axioms remain true under the replacement of i with $-i$, so must the theorems—for example, any fact about the roots of a polynomial with real coefficients. (The coefficients must be real so that the transformation of i into $-i$ leaves the polynomial unchanged.) The symmetry expressed by d'Alembert's theorem is thus grounded in the same symmetry in the axioms.

Here we have another example where a proof is privileged as explanatory because it exploits a symmetry in the problem—a symmetry of the same kind as initially struck us in the fact being explained. Furthermore, this is a good example with which to combat the impression that a proof's being explanatory is no more objective (no less "in the eye of the beholder") than a proof's being understandable, being of interest, or being sufficiently spelled out. Mathematicians largely agree on whether or not a proof is aptly characterized as "brute force," and I suggest that no "brute force" proof is explanatory. A brute force approach is not selective. It sets aside no features of the problem as irrelevant. Rather, it just "ploughs ahead" like a "bulldozer" (Atiyah 1988, 215), plugging everything in and calculating everything out. (The entire polynomial, not just some piece or feature of it, was used in the first proof above of d'Alembert's theorem.) In contrast, an explanation must be selective. It must pick out a particular feature of the setup and deem it responsible for (and other features irrelevant to) the result being explained. (Shortly we will see another example in which a brute-force proof is explanatorily impotent.) Mathematicians commonly say that a brute-force solution supplies "little understanding" and fails to show "what's going on" (as in Levi 2009, 29–30).

Suppose that we had begun not with d'Alembert's theorem, but instead merely with some particular instances of it (as d'Alembert might have done). The solutions of $z^3 + 6z - 20 = 0$ are 2, $-1 + 3i$, and $-1 - 3i$. The solutions of $z^2 - 2z + 2 = 0$ are $1 - i$ and $1 + i$. In both examples, the solutions that are not real numbers are pairs of complex conjugates. Having found many examples like these, one might ask, "Why is it that in all of the cases we 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?" One possible answer to this why question is that they are not all like that; we have simply gotten lucky by having examined an unrepresentative group of examples. Another possible answer is that our examples were unrepresentative in some systematic way: all polynomials of a certain kind (for instance, with powers less than 4) have their nonreal solutions coming in complex-conjugate pairs, and all of the polynomials we examined were of that kind. In fact, as we have seen, 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.

However, it is not the case that just any broader mathematical theorem at all that subsumes the examples to be explained would suffice to account for them. After all, we could have subsumed those two cases under this gerrymandered theorem: For any equilateral triangle or equation that is either $z^3 + 6z - 20 = 0$ or $z^2 - 2z + 2 = 0$ (the two cases above), either the triangle is equiangular or the equation's nonreal solutions all form complex-conjugate pairs. This theorem does not explain why the two equations have the given feature. (Neither does a theorem covering just these two cases.) Plausibly, whether a theorem can be used to explain its instances depends on whether that theorem has a certain kind of explanation. (I argue for this view in Lange 2010.) In any case, my concern in this essay is with the way that a certain *proof* of some theorem can explain why that theorem holds rather than with mathematical explanations where a *theorem* explains why one or more of its instances hold.

Here is another example (also discussed by Kitcher 1989, 425–26) of a proof that is widely respected as possessing explanatory power because it derives a result exhibiting a salient symmetry from a similar symmetry in the setup. It had been well known before Lagrange that a cubic equation of the form $x^3 + nx + p = 0$, once transformed by $x = y - n/3y$, becomes a sixth-degree equation $y^6 + py^3 + n^3/27 = 0$ (the "resolvent") that, miraculously, is quadratic in y^3. Lagrange aimed to determine why: "I gave reasons why [*raison pourquoi*] this equation, which is always of a degree greater than that of the given equation, can be reduced" (Lagrange 1826 [1808], 242). Lagrange showed that exactly the resolvent's solutions y can be generated by taking $1/3\ (a_1 + \omega a_2 + \omega^2 a_3)$ and replacing a_1, a_2, and a_3 with the cubic's three solutions x_1, x_2, and x_3 in every possible order—where $\omega = (-1 + \sqrt{3i})/2$, one of the cube roots of unity. But the three solutions generated by even permutations of x_1, x_2, and x_3—namely, $1/3\ (x_1 + \omega x_2 + \omega^2 x_3)$, $1/3\ (x_2 + \omega x_3 + \omega^2 x_1)$, and $1/3\ (x_3 + \omega x_1 + \omega^2 x_2)$—all have the same cubes (since $1 = \omega^3 = (\omega^2)^3$)—and likewise for the three solutions generated by odd permutations. Since y^3 takes on only two values, y^3 must satisfy a quadratic equation. So "this is why the equation that y satisfies proves to be a quadratic in y^3"

(Kline 1972, 602). The symmetry of $1/3\ (a_1 + \omega a_2 + \omega^2 a_3)$ under permutations of the three x_i explains the symmetry that initially strikes us regarding the sixth-degree equation (that three of its six roots are the same, and the remaining three are, too). As mathematicians commonly remark, y^3 "assumes two values under the six permutations of the [x_i]. It is for this reason that the equation of degree six which [y] satisfies is in fact a quadratic in $[y]^3$" (Kiernan 1971, 51).

5. Two Geometric Explanations That Exploit Symmetry

Proofs in geometry can also explain by exploiting symmetries. Consider the theorem: If ABCD is an isosceles trapezoid as shown in figure 2 (AB parallel to CD, AD = BC) such that AM = BK and ND = LC, then ML = KN.

A proof could proceed by brute-force coordinate geometry: first let D's coordinates be (0,0), C's be (0,c), A's be (a,s), and B's be (b,s), and then solve algebraically for the two distances ML and KN, showing that they are equal. A more inventive, Euclid-style option would be to draw some auxiliary lines (fig. 3) and to exploit the properties of triangles:

> Draw the line from N perpendicular to CD; call their intersection P (see fig. 3); likewise draw line LS. Consider triangles DNP and CLS: angles D and C are congruent (since the trapezoid is isosceles), ND = LC (given), and the two right angles are congruent. Hence, by having two angles and the nonincluded side congruent, ΔDNP = ΔCLS, so their corresponding sides NP and LS are congruent. They are also parallel (being

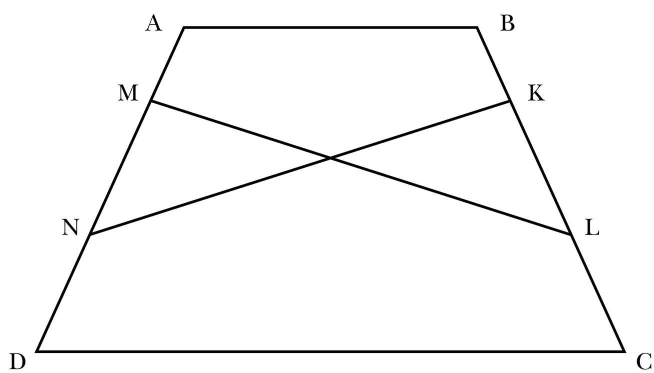

Figure 2. An isosceles trapezoid

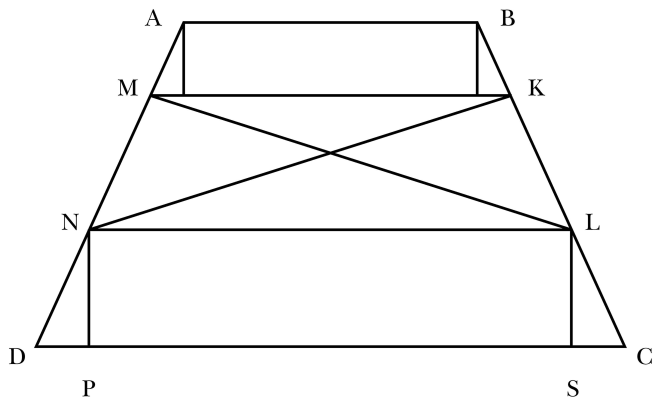

Figure 3. A Euclid–style proof

perpendicular to the same line). That these two opposite sides are both congruent and parallel shows PNLS to be a parallelogram. Hence, NL is parallel to DC. By the same argument with two new auxiliary lines, AB is parallel to MK. Therefore, MK and NL are parallel (since they are parallel to lines that are parallel to each other), so MKLN is a trapezoid. Since $MN = AD - AM - ND$, $KL = BC - BK - LC$, $AM = BK$, $AD = BC$, and $ND = LC$, it follows that $MN = KL$. As corresponding angles, $\measuredangle KLN = \measuredangle LCS$; since $\Delta CLS = \Delta DNP$, $\measuredangle LCS = \measuredangle NDP$; as corresponding angles, $\measuredangle NDP = \measuredangle MNL$. Therefore, $\measuredangle KLN = \measuredangle MNL$. From this last (and that $NL = NL$, $MN = KL$), it follows (by having two sides and their included angle congruent) that $\Delta MNL = \Delta KLN$, and so their corresponding sides ML and KN are the same length.

This proof succeeds, but only by using a construction that many mathematicians would regard as artificial or "clever." (See, for instance, Vinner and Kopelman 1998, from whom I have taken this example.) The construction is artificial because the proof using it seems forced to go to elaborate lengths—all because it fails to exploit the feature of the figure that most forcibly strikes us: its symmetry with respect to the line between the midpoints of the bases (fig. 4).

The theorem (that $ML = KN$) "makes sense" in view of the figure's overall symmetry. Intuitively, a proof that fails to proceed from the figure's symmetry strikes us as failing to focus on "what is really going on": that we have here the same figure twice, once on each side of the line of symmetry. Folding the figure along the line of symmetry, we find that NO coincides with LO and that MO coincides with KO, so that MO +

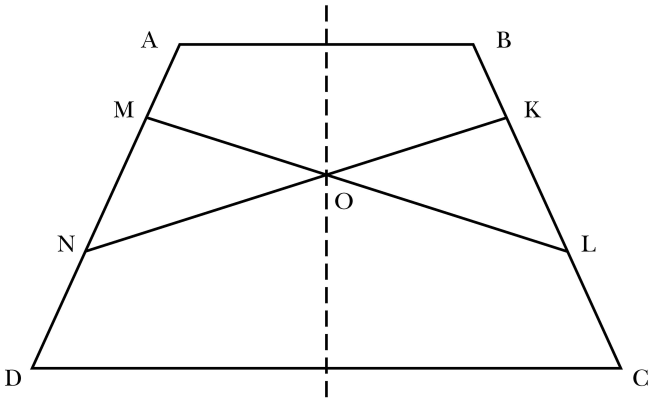

Figure 4. A proof by symmetry

OL = KO + ON, and hence ML = KN. Of course, to make this proof complete, we must first show that the point at which ML intersects KN lies on the line of symmetry. But that is also required by the figure's overall symmetry: if they intersect off of the line of symmetry, then the setup will be symmetrical only if there is another point of intersection at the mirror-image location on the other side of the line of symmetry, but two lines (ML and KN) cannot intersect at more than one point.

Of course, this proof exploits a very simple symmetry: mirror reflection across a line. A proof in geometry can explain by virtue of exploiting a more intricate symmetry in the setup. For instance, consider one direction of Menelaus's theorem: If the three sides of triangle ABC are intersected by a line l_1 (fig. 5), where C$'$ (A$'$,B$'$) is the point where l_1 intersects line AB (BC, CA), then

$$\frac{AC'}{BC'}\frac{BA'}{CA'}\frac{CB'}{AB'} = 1$$

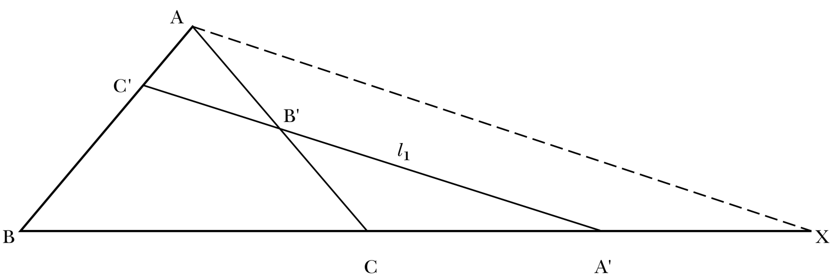

Figure 5. Menelaus's theorem

We are immediately struck by the symmetry on the left side of this equation. (Indeed, we inevitably use the symmetry to get a grip on what the left side is all about.) The left side consists of a framework of primed points $\frac{\square C'}{\square C'}\frac{\square A'}{\square A'}\frac{\square B'}{\square B'}$ exhibiting an obvious symmetry: “A,” “B,” and “C” all play the same role around the primes (modulo the order from left to right, which makes no difference to the product of the three terms). Within this framework, the three unprimed points are arranged so that “A,” “B,” and “C” all again play the same role: each appears once on the top and once on the bottom, and each is paired once with each of the other two letters primed. These constraints suffice to fix the expression modulo the left-right order, which does not matter to their product, and modulo the inversion of top and bottom, which does not matter since the equation sets the top and bottom equal. In short, the left-hand expression is invariant (modulo features irrelevant to the equation) under any systematic interchange of “A,” “B,” and “C” around the other symbols. In the literature, the primed points are almost always named as I have named them here (for instance, with C$'$ as the point where l_1 intersects line AB) in order to better display the expression’s symmetry.

Having recognized this symmetry in the theorem, we regard any proof of the theorem that ignores the symmetry as failing to explain why the theorem holds. For instance, consider this proof:

> Draw the line through A parallel to l_1 (the dotted line in fig. 5); let X be its point of intersection with line BC. As corresponding angles, $\measuredangle$BA$'$C$'$ = $\measuredangle$BXA. Therefore (since they also share $\measuredangle$B), ΔBC$'$A$'$ is similar to ΔBAX, so their corresponding lengths are in a constant proportion. In particular, $\frac{AC'}{BC'} = \frac{XA'}{BA'}$. Likewise, as corresponding angles, $\measuredangle$CB$'$A$'$ = $\measuredangle$CAX. Therefore (since they also share $\measuredangle$B$'$CA$'$), ΔACX is similar to ΔB$'$CA$'$, so their corresponding lengths are in a constant proportion. In particular, $\frac{AB'}{CB'} = \frac{XA'}{CA'}$. Solving this for XA$'$ = (AB$'$)(CA$'$)/CB$'$ and substituting the resulting expression for XA$'$ into the earlier equation yields $\frac{AC'}{BC'} = \frac{AB'}{BA'}\frac{CA'}{CB'}$. The theorem follows by algebra.

Einstein (Luchins and Luchins 1990, 38) says that this proof is “not satisfying” and Bogomolny (n.d.) agrees. Both cite the fact that (in Einstein’s words) “the proof favors, for no reason, the vertex A [since the auxiliary line is drawn from that vertex], although the proposition [to be proved] is symmetrical in relation to A, B, and C.” The point here is that although the proof could have been carried out with an auxiliary line parallel to l_1 drawn through vertex B or C rather than through A, the choice of any one vertex through which to draw the line breaks the symmetry between

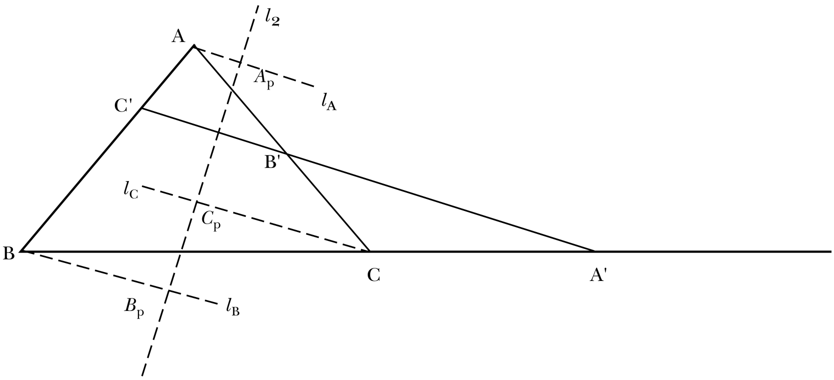

Figure 6. Another proof of Menelaus's theorem

A, B, and C in what had been an entirely symmetric arrangement.[15] That symmetry is then restored in the resulting equation. Consequently, this proof depicts the symmetric result as arising "magically," whereas to explain why the theorem holds, we must proceed entirely from the figure's symmetries over A-B-C.

Following Brunhes (1991 [1920], 84), Bogomolny (n.d.) offers such a proof, which I now elaborate (fig. 6):

> Add a line l_2 perpendicular to l_1. Project A onto l_2 by a line l_A from A parallel to l_1; let A_p be the point on l_2 to which A is projected. Perform the same operation on B and C, adding lines l_B and l_C, points B_p and C_p. Of course, A$'$, B$'$, and C$'$ all project to the same point on l_2 (since l_1 is their common line of projection), which can equally well be called "A'_p", "B'_p", or "C'_p". The following equation exhibits the same symmetry as the theorem:
>
> $$\frac{A_pC'_p}{B_pC'_p}\ \frac{B_pA'_p}{C_pA'_p}\ \frac{C_pB'_p}{A_pB'_p} = 1$$
>
> This equation is true (considering that A'_p, B'_p, and C'_p are the same point, allowing a massive cancellation) and is strikingly invariant (modulo features irrelevant to the equation: left-right order and inversion of top and

15. The arrangement may seem asymmetric in one respect: vertex A differs from B and C in being opposite to the only side of ΔABC that is intersected by l_1 along its extension outside the triangle rather than between the triangle's vertices. However, the configuration shown in figure 5 is a special case of Menelaus's theorem since l_1 need not intersect any side of ΔABC between the vertices; l_1 may intersect all three sides along their extensions.

bottom) under systematic interchange of “A,” “B,” and “C” around the primes and subscript “p”s—the same symmetry that the theorem possesses. To arrive at the theorem, all we need to do is to find a way to remove the subscript “p”s from this equation, which is easily done. For any side of the triangle, l_1 and the two lines projecting its endpoints onto l_2 are three parallel lines, and all three are crossed by the side and by l_2. Now for a lemma: Whenever two transversals cross three parallel lines, the two segments into which the three parallels cut one transversal stand in the same ratio as the two segments into which the three parallels cut the other transversal. That is, the ratio of one transversal’s segments is preserved in the ratio of their projections onto the other transversal. By projecting each side onto l_2, we find

$$\frac{A_pC'_p}{B_pC'_p} = \frac{AC'}{BC'} \quad \frac{B_pA'_p}{C_pA'_p} = \frac{BA'}{CA'} \quad \frac{C_pB'_p}{A_pB'_p} = \frac{CB'}{AB'}$$

Thus the theorem is proved by an argument that begins with an equation that treats A, B, and C identically and in each further step treats them identically. This proof reveals how features of the setup that are A-B-C symmetric are responsible for the theorem’s symmetry. The result’s symmetry does not just come out of nowhere. Indeed, the proof’s general strategy is to project A′, B′, and C′ onto the very same point (thereby treating them identically) by projecting the triangle’s three sides onto the same line.

This explanation also shows that a proof’s auxiliary lines need not be “artificial”; that is, the use of auxiliary lines does not suffice to make the proof nonexplanatory. Although the auxiliary lines in the Euclid-style proof of the trapezoid theorem were mere devices to get the theorem proved, and likewise for line AX in the first proof of Menelaus’s theorem, the scheme of auxiliary lines in the second proof of Menelaus’s theorem is A-B-C symmetric (Bogomolny, n.d.).

6. Generalizing the Proposal: Mathematical Explanations That Do Not Exploit Symmetries

I have now given several examples of mathematical explanations consisting of proofs that exploit symmetries. However, I do not mean to suggest that only a proof that appeals to some symmetry can explain why a mathematical theorem holds. Rather, I am using proofs by symmetry to illustrate the way in which certain proofs manage to become privileged as explanatory. Symmetries are not somehow intrinsically explanatory in mathematics. Rather, some symmetry in a mathematical result is often salient to us, and consequently, in those cases, a proof that traces the result back to a similar symmetry in the problem counts as explaining why the result holds. Some feature of a mathematical result other than its symmetry could likewise be salient, prompting a why question answerable by a proof deriving the result from a similar feature of the given. What it *means* to ask for a proof that explains is to ask for a proof that exploits a certain kind of feature in the setup—the same kind of feature that is salient in the result. The distinction between proofs that explain why some theorem holds and proofs that merely establish that it holds exists only when some feature of the result being proved is salient. That feature's salience makes certain proofs explanatory. A proof is accurately characterized as an explanation (or not) only in a context where some feature of the result being proved is salient.

My proposal predicts that if the result exhibits no noteworthy feature, then to demand an explanation of why it holds, not merely a proof that it holds, makes no sense. There is nothing that its explanation over and above its proof would amount to until some feature of the result becomes salient.[16]

This prediction is borne out. For example, there is nothing that it would be for some proof to explain why, not merely to prove that, $\int_1^3 (x^3 - 5x + 2)dx = 4$. Nothing about this equation calls for explanation.

My proposal also predicts that if a result exhibits some noteworthy feature, but no proof traces that result to a similar feature in the setup, then the result has no explanation. This prediction is also borne out. Take the following example of a "mathematical coincidence" given by the mathematician Timothy Gowers (2007, 34): "Consider the decimal expansion of *e*, which begins 2.718281828.... It is quite striking that a pattern of four digits should repeat itself so soon—if you choose a random sequence of digits, then the chances of such a pattern appearing would be one in several thousand—and yet this phenomenon is universally regarded as an amusing coincidence, a fact that does not demand an explanation" (see also Baker 2009, 140). I take it, then, that mathemati-

16. There may also be cases where the result exhibits a feature that is only slightly salient. If some proofs but not others exploit a similar feature in the problem, this difference would ground only a slight distinction between proofs that explain why and proofs that merely establish that some theorem holds. Another way for intermediate cases to arise is for a certain feature to be salient in the result, but for proofs to exploit to varying degrees a similar feature in the setup—rather than for any proof to proceed entirely from such a feature. See also note 21.

cians regard this fact as having no explanation. Of course, there are many ways to derive *e*'s value, and thus to derive that the third-through-sixth digits of its base-ten representation are repeated in the seventh-through-tenth digits. For example, we could derive this result from the fact that *e* equals the sum of $(1/n!)$ for $n = 0, 1, 2, 3 \ldots$ However, such a proof does not *explain why* the seventh-through-tenth digits repeat the third-through-sixth digits. It merely proves that they do. On my view, that is because the expression $(1/0!) + (1/1!) + (1/2!) + \ldots$ from which the proof begins does not on its face exhibit any feature similar to the repeated sequence of digits in *e*'s decimal expansion. (None of the familiar expressions for calculating *e* makes any particular reference to base 10.) There is, I suggest, no reason why that pattern of digits repeats. It just does.

Let's now return to the example from section 1: that every "calculator number" is divisible by 37. Is this fact a coincidence? (This is the question asked by the title of the *Mathematical Gazette* article that contributed this example to the mathematical literature.)

The striking thing about this result is that it applies to every single calculator number. In other words, the result's "unity" is salient. A proof that simply takes each calculator number in turn, separately showing each to be divisible by 37, treats the result as if it were a coincidence. That is, it fails to explain why all of the calculator numbers are divisible by 37. Indeed, a case-by-case proof merely serves to highlight the fact that the result applies to every single calculator number. Especially in light of this proof, an explanation would be a proof that proceeds from a property common to each of these numbers (where this property is a genuine respect in which these numbers are similar, that is, a mathematically "natural" property—unlike the property of being either 123321 or 321123 or . . .) and that is common to them precisely because they are calculator numbers.[17] I gave such a proof in section 1. I took it from a later *Mathematical Gazette* article (Nummela 1987, 147) entitled "No Coincidence." That proof exploits the fact that every calculator number

17. Elsewhere (Lange, n.d.), I am more explicit about what it takes for a mathematical predicate to denote a mathematically natural property—a genuine respect of similarity.

Strictly speaking, that a given result identifies a property common to every single case of a certain sort is just a symmetry in the result—for example, that under a switch from one calculator number to any other, divisibility by 37 is invariant. Accordingly, I have already argued that a proof that works by exploiting the same sort of symmetry in the setup counts as explanatory. My view does not require that a theorem's displaying a striking symmetry be sharply distinguished from a theorem's being striking for its treating various cases alike.

can be expressed as $10^5 a + 10^4(a + d) + 10^3(a + 2d) + 10^2(a + 2d) + 10(a + d) + a$ where a, $a + d$, $a + 2d$ are three integers in arithmetic progression. These three integers, of course, are the three digits on the calculator keypads that, taken forward and backward, generate the given calculator number. Hence, this proof traces the fact that every calculator number is divisible by 37 to a property that they have in common by virtue of being calculator numbers.

In short, an explanation of this result consists of a proof that treats every calculator number in the same way.[18] This unified treatment makes the proof explanatory because what strikes us as remarkable about the result, especially in light of the case-by-case proof, is its unity: that it identifies a property common to every calculator number. The point of asking for an explanation is then to ask for a proof that exploits some other feature common to them because they are calculator numbers.[19]

18. Elsewhere (Lange 2010, n.d.), I am more explicit about what it takes to treat them all "in the same way."

19. If a result is remarkable for identifying something common to each case of an apparently diverse lot, then those "cases" may themselves be general results—as when each case is a theorem and the result identifies something common to each of them. Entirely dissimilar proofs of two theorems would fail to explain why those theorems involve a common element. On the other hand, proofs of each theorem may explain this pattern if the proofs themselves exploit a common element. Note the explanatory language in this remark from a mathematician:

> There are lots of 'meta-patterns' in mathematics, i.e. collections of seemingly different problems that have similar answers, or structures that appear more often than we would have expected. Once one of these meta-patterns is identified it is always helpful to understand what is responsible for it. . . . To give a trivial example, years ago while the author was writing up his PhD thesis he noticed in several places the numbers 1, 2, 3, 4 and 6. For instance, $\cos(2\pi r) \in \mathbb{Q}$ for $r \in \mathbb{Q}$ iff the denominator of r is 1, 2, 3, 4 or 6. Likewise, the theta function $\Theta[\mathbb{Z} + r](\tau)$ for $r \in \mathbb{Q}$ can be written as $\Sigma\, a_i\theta_3(b_i\, \tau)$ for some $a_i, b_i \in \mathbb{R}$ iff the denominator of r is 1,2,3,4 or 6. This pattern is easy to explain: they are precisely those positive integers n with Euler totient $\varphi(n) \leq 2$, that is there are at most two positive numbers less than n coprime to n. The various incidences of these numbers can usually be reduced to this $\varphi(n) \leq 2$ property. (Gannon 2006, 168)

We can understand the general idea here even if we do not know what Jacobi's theta function is or how $\varphi(n) \leq 2$ is connected to these two theorems. The general idea is that proofs of the two theorems (one about cos, the other about $\Theta[\mathbb{Z} + r]$) can explain why {1,2,3,4,6} figures in both if each of those proofs exploits exactly the same feature of {1,2,3,4,6}—for example, nothing about {1,2,3,4,6} save that this set contains all and only the positive integers where $\varphi(n) \leq 2$. (A proof exploits more if, for instance, it determines which numbers n are such that $\varphi(n) \leq 2$, and then proceeds case-wise from there.) Since the striking feature of the two theorems (when presented as Gannon pres-

A proof does not have to treat each instance separately in order to fail to treat them all together. Consider this proof by mathematical induction of the fact that the product of any three consecutive nonzero natural numbers is divisible by 6:

> The product of 1, 2, and 3 is 6, which is divisible by 6.
>
> Suppose that the product of $(n-1)$, n, and $(n+1)$ is divisible by 6. Let's show that the product of n, $(n+1)$, and $(n+2)$ is divisible by 6. By algebra, that product equals $n^3+3n^2+2n=(n^3-n)+3n(n+1)$. Now $(n^3-n)=(n-1)n(n+1)$, so by hypothesis, it is divisible by 6. And $n(n+1)$ is even, so $3n(n+1)$ is divisible by 3 and by 2, and therefore by 6. Hence, the original product is the sum of two terms, each divisible by 6. Hence, that product is divisible by 6.

This proof fails (by a whisker—see note 21) to treat every product of three consecutive nonzero natural numbers alike. Instead, it divides them into two classes: the product of 1, 2, and 3; and all of the others. This proof, then, does not supply a common reason why all of the products of three consecutive nonzero natural numbers are divisible by 6. Rather, it treats the first case as a special case. Insofar as we found the theorem remarkable for identifying a property common to every triple of consecutive nonzero natural numbers, our point in asking for an explanation was to ask for a proof that treats all of the triples alike. This feature of the theorem is made especially salient by a proof that does treat all of the triples alike:

> Of any three consecutive nonzero natural numbers, at least one is even (that is, divisible by 2) and exactly one is divisible by 3. Therefore, their product is divisible by $3 \times 2 = 6$.

This proof proceeds entirely from a property possessed by every triple.[20] Like the explanation of the fact that every calculator number is divisible by 37, this proof traces the result to a property common to every instance

---

ents them) is that {1,2,3,4,6} figures in both, the point of asking why the two theorems hold is plainly to ask for proofs of the two theorems where (nontrivially) both proofs exploit exactly the same feature of {1,2,3,4,6}. That both theorems involve exactly the positive integers n where $\varphi(n) \leq 2$ explains why both theorems involve {1,2,3,4,6}, and hence involve exactly the same integers.

20. Zeitz (2000, 1) mentions this theorem in passing but neither discusses it nor gives any proof of it. I wonder if he had in mind either of the proofs I discuss.

and so (when the unity of the result is salient) explains why the result holds.[21]

A proof may focus our attention on a particular feature of the result that would not otherwise have been salient. The proof may even call our attention to this feature because the proof conspicuously fails to exploit it. When this happens, the proof fails to qualify as explanatory. For instance, take one standard proof of the formula for the sum S of the first n nonzero natural numbers $1 + 2 + \ldots + (n - 1) + n$.

> There are two cases.
>
> When n is even, we can pair the first and last numbers in the sequence, the second and second-to-last, and so forth. The members of each pair sum to $n + 1$. No number is left unpaired, since n is even. The number of pairs is $n/2$ (which is an integer, since n is even). Hence, $S = (n + 1)n/2$.

21. It follows that in a context where the result's unity is salient, a proof by mathematical induction cannot explain the result, since a proof by mathematical induction always treats the first instance as a special case. However, the inductive proof of the triplet theorem is a far cry from the proof of the calculator-number theorem that treats each of the sixteen calculator numbers separately. The inductive proof *nearly* treats every triple alike. It gives special treatment only to the base case; all of the others receive the same treatment. Therefore, although this inductive proof is not an explanation (when the result's unity is salient), it falls somewhere *between* an explanation and a proof utterly lacking in explanatory power. (Explanatory power is a matter of degree; it is not all or nothing.) A result (displaying unity as its striking feature) having such a proof by induction, but (unlike the triplet theorem) having no proof that treats every case alike, has no fully qualified explanation but is not an utter mathematical coincidence either, since the inductive proof ties all but one of its cases together. On my account, the triplet theorem's inductive proof is inferior in explanatory power to the fully unifying proof, but nevertheless retains some measure of explanatory significance.

In a proof by course-of-values induction (a.k.a. "strong induction," "complete induction"), there is no base case. Such an argument uses this rule of inference:

For any property P,
if for any natural number k, if P(1) and . . . and P(k − 1), then P(k),
then for any natural number n, P(n).

Since there is no base case, it might appear that a course-of-values induction explains (in a context where the result's unity is salient), by my lights, since it does not give special treatment to a base case. However, even though a course-of-values induction contains only the "inductive step," typically (as when the triplet theorem is proved in this way) it must prove the inductive step by treating k = 1 as a special case, since the antecedent "if P(1) and . . . and P(k − 1)" is empty for k = 1. So although any result provable by ordinary induction can be proved by course-of-values induction, the latter has no automatic explanatory advantage over the former.

> When n is odd, we can pair the numbers as before, except that the middle number in the sequence is left unpaired. Again, the members of each pair sum to $n+1$. But now there are $(n-1)/2$ pairs, since the middle number $(n+1)/2$ is unpaired. The total sum is the sum of the paired numbers plus the middle number: $S = (n+1)(n-1)/2 + (n+1)/2$. This simplifies to $(n+1)n/2$—remarkably, the same as the expression we just derived for even n.

Before having seen this proof, we would not have found it remarkable that the theorem finds that the same formula applies to both even n and odd n. However, this feature of the result strikes us forcibly in light of this proof. We might then well wonder: is it a coincidence that the same formula emerges in both cases? This proof depicts it as an algebraic miracle. Accordingly, in this context, to ask for the reason why the formula holds, not merely a proof that it holds, is to ask for the feature (if any) common to both of these cases from which the common result follows.

Indeed there is such a feature; the result is no coincidence. Whether n is even or odd, the sequence's midpoint is half of the sum of the first and last numbers: $(1+n)/2$. Furthermore, all sequences of both kinds consist of numbers balanced evenly about that midpoint. In other words, for every number in the sequence exceeding the midpoint by some amount, the sequence contains a number less than the midpoint by the same amount. Allowing the excesses to cancel the deficiencies, we have each sequence containing n numbers of $(1+n)/2$ each, yielding our formula. This is essentially what happens in the standard proof of the formula, where each member in the sequence having an excess is paired with a member having an equal deficiency:

> $S = 1 + 2 + \ldots + (n-1) + n$
>
> $S = n + (n-1) + \ldots + 2 + 1$
>
> If we pair the first terms, the second terms, and so forth, in each sum, then each pair adds to $(n+1)$, and there are n pairs. So $2S = n(n+1)$, and hence $S = n(n+1)/2$.

This proof is just slightly different from the proof that deals separately with even n and odd n. The use of two sequences could be considered nothing but a trick for collapsing the two cases. Yet it is more than that. It brings out something that the earlier proof obscures: that the common result arises from a common feature of the two cases. For this reason, the earlier proof fails to reveal that the formula's success for both even n and odd n is no coincidence. What allows the second proof to show that this is no coincidence? It traces the result to a property common to the two cases: that the terms are balanced around $(1 + n)/2$. (Whether or not a term in the sequence actually occupies that midpoint is irrelevant to this balancing.) The earlier proof does not exploit this common feature. Rather, it simply works its way through the two cases and magically finds itself with the same result in both.[22]

Similar phenomena arise in some explanations that are not proofs. For example, mathematicians such as Cardano and Euler had developed various tricks for solving cubic and quartic equations. One of Lagrange's tasks in his monumental 1770–71 memoir "Reflections on the Solution of Algebraic Equations" was to explain why his predecessors' various methods all worked ("pourquoi ces méthods réussissent") (Lagrange 1869 [1770–71], 206; see also Kline 1972, 601). By showing that these different procedures all amount fundamentally to the same method, Lagrange showed that it was no coincidence that they all worked. In other words, faced with the fact that Cardano's method works, Euler's method works, Tschirnhaus's method works, and so forth, one feature of this fact is obviously salient: that it identifies something common to each of these methods (namely, that each works). Lagrange's explanation succeeds by tracing this common feature to other features that these methods have in common—in sum, that they are all fundamentally the same method: "These methods all come down to the same general principle" (Lagrange 1869 [1770–71], 355, quoted in Kiernan 1971, 45). Part of the point in asking *why* these methods all work is to ask whether their success can be traced to a feature common to them all or is merely a coincidence. (Lagrange then asked why this common method works; earlier we saw his explanation of the "resolvent.")

Simplicity is another feature that generally stands out when a mathematical result possesses it.[23] An especially simple result typically

22. Steiner (1978a, 136) deems the second proof "more illuminating" than an inductive proof. Steiner does not discuss mathematical coincidences or contrast the second proof with separate proofs for the even and odd cases.

23. Is it unilluminating to characterize a "salient feature" as a feature that gives content (in the manner I have described) to a demand for mathematical explanation, while in turn characterizing "mathematical explanation" in terms of a salient feature? I do not think that my proposal is thereby rendered trivial or unilluminating. We can say plenty about salience other than through its role in mathematical explanation (by way of paradigm cases such as those given in this essay, the kinds of features that are typically salient, how features become salient, and so forth), and we can recognize a feature as salient apart from identifying a proof as explanatory. We can likewise say plenty about mathematical

cries out for a proof that exploits some similar, simple feature of the setup. In contrast, a proof where the result, in all of its simplicity, appears suddenly out of a welter of complexity—through some fortuitous cancellation or clever manipulation—tends merely to heighten our curiosity about why the result holds. Such a proof leaves us wanting to know where such a simple result came from. After such a proof is given, there often appears a sentence like the following: "The resulting answer is extremely simple despite the contortions involved to obtain it, and it cries out for a better understanding" (Stanley 2012, 14). Then another proof that traces the simple result to a similar, simple feature of the problem counts as explaining why the result holds.

Examples of this kind often arise in proofs of "partition identities." The number $p(n)$ of "partitions of n" is the number of ways that the nonnegative integer n can be expressed as the sum of one or more positive integers (irrespective of their order in the sum). For instance, $p(5) = 7$, since 5 can be expressed in seven ways: as 5, $4 + 1$, $3 + 2$, $3 + 1 + 1$, $2 + 2 + 1$, $2 + 1 + 1 + 1$, and $1 + 1 + 1 + 1 + 1$. (By convention, $p(0) = 1$.) Here is a "partition identity" (proved by Euler in 1748): The number $p_O(n)$ of partitions of n into exclusively odd numbers ("O-partitions") equals the number $p_D(n)$ of partitions of n into parts that are distinct (that is, that are all unequal—"D-partitions"). For instance, $p_O(5) = 3$ since 5, $3 + 1 + 1$, and $1 + 1 + 1 + 1 + 1$ are the O-partitions of 5, and $p_D(5) = 3$ since 5, $4 + 1$, and $3 + 2$ are the D-partitions of 5.

There are two standard ways of proving partition identities: either with "generating functions" or with "bijections." By definition, the "generating function" f(q) for a sequence $a_0, a_1, a_2 \ldots$ is $a_0q^0 + a_1q^1 + a_2q^2 + \ldots = a_0 + a_1q + a_2q^2 + \ldots$. It does not matter whether the sum in the generating function converges because the generating function is merely a device for putting the sequence on display; "q^n" does not stand for some unknown quantity, but just marks the place where a_n appears. For example, the generating function for the sequence 1, −1,

explanation apart from its connection to salience. The connection my account alleges between salience and mathematical explanation does no more to trivialize my account than Van Fraassen's pragmatic approach to scientific explanation is trivialized by the fact that it characterizes a "scientific explanation" as an answer to a question defined in terms of a contrast class, while it characterizes the "contrast class" in a given case as consisting of the possible occurrences that are understood to be playing a certain role in a given question demanding a scientific explanation (Van Fraassen 1980, 127).

$1, -1, 1 \ldots$ is $1 - q + q^2 - q^3 + q^4 - \ldots$. This series results from "long division" of $1 + q$ into 1; that is, it is the expansion of $1/(1 + q)$. (Again, it does not matter whether this sum converges.) Such a "formal power series" can generally be manipulated in precisely the same manner as a genuine power series. For example, we can multiply $(1/(1 + q))$ by 2 to yield $(2/(1 + q)) = 2 - 2q + 2q^2 - 2q^3 + 2q^4 - \ldots$, the generating function for the sequence $2, -2, 2, -2, 2 \ldots$.

Now let's find a generating function for the sequence of partitions of n: $p(0), p(1), p(2), \ldots$. Let's "consider (or as that word often implies, 'look out, here comes something from left field')" (Wilf 2000, 6) this formidable-looking generating function:[24]

$$(1 + q + q^2 + q^3 + \ldots)(1 + q^2 + q^4 + q^6 + \ldots)(1 + q^3 + q^6 + q^9 + \ldots)$$

$$(1 + q^4 + q^8 + q^{12} + \ldots)\ldots$$

The product of the factors in this expression is a long polynomial $a_0 + a_1q + a_2q^2 + a_3q^3 + \ldots$—the generating function for the sequence $a_0, a_1, a_2 \ldots$. What sequence is this? As an example, let's compute a_3. Since multiplying q^m by q^n involves adding their exponents ($q^m q^n = q^{m+n}$), the coefficient of q^3 will be increased by 1 for every combination of terms where the exponents add to 3, with exactly one term in the combination drawn from each factor in the above expression. For instance, q^3 from the first factor $(1 + q + q^2 + q^3 + \ldots)$ multiplied by 1's from each other factor $(1 = q^0)$ will increase a_3 by 1. Another 1 is contributed to a_3 by q from the first factor multiplied by q^2 from the second factor and 1's from each of the other factors. Another, final 1 comes from q^3 from the third factor multiplied by 1's from each other factor. Those three 1's make $a_3 = 3$; each combination of exponents adding to 3 contributes 1 to a_3. Now the exponents in the first factor $(1 + q + q^2 + q^3 + \ldots)$ increase by 1, those in the second factor $(1 + q^2 + q^4 + \ldots)$ increase by 2, and so forth. Thus, we can think of the first factor's exponents as counting by 1's, the second factor's exponents

24. In referring to the following expression as a "generating function," I am adopting (as a referee put it) "standard loose mathematician-speak." A "generating function" (as the referee went on to say) is really an infinitary symbol—that is, a string of expressions that takes the form $a_0 + a_1q + a_2q^2 + \ldots$. Therefore, the infinite product of generating functions that is about to appear in the main text is not, strictly speaking, a generating function. Rather, it is a string of strings that computes to a generating function in the manner described in the main text. Accordingly, mathematicians standardly characterize it as a "generating function."

as counting by 2's, and so on. Therefore, when we add the exponents from a combination of terms where exactly one term q^f was drawn from the first factor, one term q^s from the second factor, and so forth, we can think of this sum as representing the sum of f 1's plus the sum of (s/2) 2's and so forth. Each combination that sums to n contributes 1 to q^n's coefficient a_n. For instance, as we just saw in the case of q^3, the product of q^3 from the first factor (representing three 1's) with 1's from every other factor (zero 2's, zero 3's . . .) contributes 1 to a_3; another 1 comes from the product of q from the first factor (one 1) by q^2 from the second factor (one 2) and 1's from every other factor (zero 3's, zero 4's . . .); and the final 1 comes from the product of q^3 from the third factor (one 3) with 1's from each of the other factors (zero 1's, zero 2's, zero 4's . . .). Thus, each of these combinations contributing 1 to a_3 represents one of the three partitions of 3: its partitions are $1+1+1$, $1+2$, and 3. That is, a_3 equals the number of partitions of 3. The same thing happens for every coefficient a_n. Thus, the above expression is $p(n)$'s generating function: $p(0) + p(1)q + p(2)q^2 + \ldots$. Since by "long division," $1 = (1+q+q^2+q^3+\ldots)(1-q) = (1+q^2+q^4+\ldots)(1-q^2) = (1+q^3+q^6+\ldots)(1-q^3)$, the above product is

$$\left(\frac{1}{1-q}\right)\left(\frac{1}{1-q^2}\right)\left(\frac{1}{1-q^3}\right)\cdots$$

Now let's derive the generating functions for the numbers of odd-partitions and distinct-partitions of n. In the generating function for $p(n)$

$$(1+q+q^2+q^3+\ldots)(1+q^2+q^4+q^6+\ldots)(1+q^3+q^6+q^9+\ldots)$$
$$(1+q^4+q^8+q^{12}+\ldots)\ldots,$$

the first factor (in counting by 1's) represents the number of 1's in the partition, the second factor (in counting by 2's) represents the number of 2's, and so on. By including only the factors corresponding to the number of 1's, number of 3's, number of 5's, and so on, we produce

$$(1+q+q^2+q^3+\ldots)(1+q^3+q^6+q^9\ldots)(1+q^5+q^{10}+q^{15}\ldots)\ldots$$
$$=\left(\frac{1}{1-q}\right)\left(\frac{1}{1-q^3}\right)\left(\frac{1}{1-q^5}\right)\cdots$$

the generating function for $p_O(n)$. Returning to the m^{th} factor $(1+q^m+q^{2m}+q^{3m}\ldots)$ of $p(n)$'s generating function, we see that the terms beyond q^m allowed m to appear two or more times in the partition, so their removal yields $p_{\mathrm{D}}(n)$'s generating function

$$(1+q)(1+q^2)(1+q^3)\ldots$$

By manipulating the generating functions in various ways (justified for infinite products by taking to the limit various manipulations for finite products), we can show (as Euler first did) that the two generating functions are the same, and hence that $p_{\mathrm{O}}(n) = p_{\mathrm{D}}(n)$:

$$\begin{aligned}
&\left(\frac{1}{1-q}\right)\left(\frac{1}{1-q^3}\right)\left(\frac{1}{1-q^5}\right)\cdots \\
&=\left(\frac{1}{1-q}\right)\left(\frac{1-q^2}{1-q^2}\right)\left(\frac{1}{1-q^3}\right)\left(\frac{1-q^4}{1-q^4}\right)\left(\frac{1}{1-q^5}\right)\left(\frac{1-q^6}{1-q^6}\right)\cdots \\
&= \left(\frac{1-q^2}{1-q}\right)\left(\frac{1-q^4}{1-q^2}\right)\left(\frac{1-q^6}{1-q^3}\right)\left(\frac{1-q^8}{1-q^4}\right)\cdots \\
&=\left(\frac{(1-q)(1+q)}{1-q}\right)\left(\frac{(1-q^2)(1+q^2)}{1-q^2}\right)\left(\frac{(1-q^3)(1+q^3)}{1-q^3}\right)\left(\frac{(1-q^4)(1+q^4)}{1-q^4}\right)\cdots \\
&=(1+q)(1+q^2)(1+q^3)\ldots
\end{aligned}$$

Wilf (2000, 10) terms this "a very slick proof," which is to say that it involves not only an initial generating function "from left field" but also a sequence of substitutions, manipulations, and cancellations having no motivation other than that, miraculously, it works out to produce the simple result in the end. Proofs of partition identities by generating functions, although sound and useful, in many cases "begin to obscure the simple patterns and relationships that the proof is intended to illuminate" (Bressoud 1999, 46).

In contrast, "a common feeling among combinatorial mathematicians is that a simple bijective proof of an identity conveys the deepest *understanding* of why it is true" (Andrews and Eriksson 2004, 9; italics in the original). A "bijection" is a $1-1$ correspondence; a bijective proof that $p_{\mathrm{O}}(n) = p_{\mathrm{D}}(n)$ finds a way to pair each O-partition with one and only one D-partition. Let us look at a bijective proof (from Sylvester 1882) that $p_{\mathrm{O}}(n) = p_{\mathrm{D}}(n)$. Display each O-partition as an array of dots, as in the representation in figure 7 of the partition $7+7+5+5+3+1+1+1$ of 30:

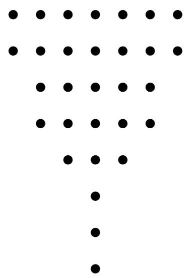

Figure 7. A partition of 30

Each row has the number of dots in a part of the partition,[25] with the rows weakly decreasing in length and their centers aligned. (Each row has a center dot since each part is odd.) Here is a simple way to transform this O-partition into a D-partition. The first part of the new partition (fig. 8) is given by the dots on a line running from the bottom up along the center column and turning right at the top—11 dots. The next part is given by the dots on a line running from the bottom, up along a column one dot left of center, turning left at the top—7 dots. The next part runs from the bottom upward along a column one dot right of center, turning right at the last available row (the second row from the top)—6 dots. This pattern leaves us with a fish-hook diagram (fig. 8).

The result is a D-partition $(11 + 7 + 6 + 4 + 2)$, and the reverse procedure on that partition returns the original O-partition. With this bijection between O-partitions and D-partitions, there must be the same number of each. The key to the proof is that by "straightening the fish hooks," we can see the same diagram as depicting both an O-partition and a D-partition.

Because the bijection is so simple, this proof traces the simple relation between $p_{\mathrm{O}}(n)$ and $p_{\mathrm{D}}(n)$ to a simple relation between the O-partitions and the D-partitions. Moreover, the simple feature of the setup that the proof exploits is similar to the result's strikingly simple feature: the result is that $p_{\mathrm{O}}(n)$ and $p_{\mathrm{D}}(n)$ are the same, and the simple bijection reveals that n's O-partitions are essentially the same objects as n's D-partitions, since one can easily be transformed into the other. It is then no wonder that $p_{\mathrm{O}}(n) = p_{\mathrm{D}}(n)$, since effectively the same objects

25. The "members" or "summands" of a partition are known as its "parts." For instance, the partition of 30 into $7 + 7 + 5 + 5 + 3 + 1 + 1 + 1$ has eight parts (three 1's, two 7's, two 5's, and one 3).

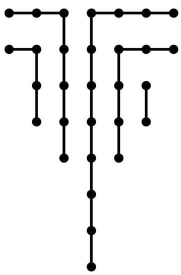

Figure 8. Another partition of 30, displayed as fish hooks

are being counted twice. This is the source of the explanatory power of a simple bijective proof. As Wilf (personal communication) puts it, a simple bijective proof reveals that "in a sense the elements [of the two sets of partitions] are the same, but have simply been encoded differently."[26]

That is, the bijective proof shows that in counting O-partitions or D-partitions, we are effectively counting the same set of abstract objects. Each of these objects can be represented by a dot diagram; the same diagram can be viewed as representing an O-partition and a D-partition. To emphasize this point (following Andrews and Eriksson 2004, 16–17), consider this partition identity: the number of partitions of n with exactly m parts equals the number of partitions of n having m as their largest part. For instance, figure 9 gives representations of the partitions of 7 with exactly three parts:[27] Figure 10 gives each of these representations seen from a different vantage point—namely, after being rotated one-quarter turn counterclockwise and then reflected across a horizontal line above it: These (figure 10) represent the partitions of 7 that have 3 as their largest part. Clearly, then, both sets of partitions are represented by the same set of 4 arrays. A partition of one kind, seen from another

26. Perhaps (contrary to the passages I have quoted) some combinatorial mathematicians regard the proof from generating functions as explanatory just like the simple bijective proof. Perhaps, then, I should restrict myself to identifying what it is that makes the bijective proof explanatory and not argue that the generating-function proof lacks this feature, but merely that it is perceived as lacking this feature by those who regard it as less explanatorily powerful than the bijective proof. (See the remark from Stanley in note 28.)

27. The next dot diagrams all have their rows weakly decreasing in length (just like the earlier diagram), but (unlike the earlier diagram) their rows do not have their centers aligned. (Since the parts are not all odd, some rows do not have center dots.) Rather, each row is aligned on the left. Clearly, the same partition may be displayed in various dot diagrams.

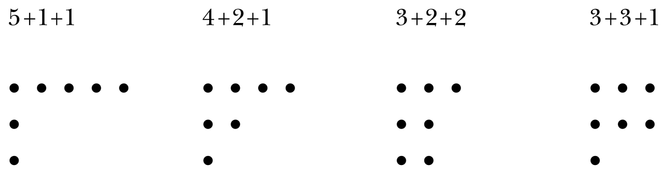

Figure 9. The partition of 7 having exactly three parts

vantage point, is a partition of the other kind. The number of partitions of one kind is the same as the number of partitions of the other kind because the same abstract object can be represented as either kind of partition, and the number of those abstract objects is the same no matter how we represent them.

Whereas the generatingfunctionology "seems like something external to the combinatorics" (Andrews, personal communication) that just miraculously manages to yield the simple result, a simple bijective proof shows that there is the same number of partitions of two kinds because partitions of those two kinds, looked at abstractly, are the same things seen from different perspectives. Such a proof "makes the reason for the simple answer completely transparent" (Stanley 2012, 15); "it provides a 'natural' explanation . . . unlike the generating function proof which depended on a miraculous trick" (Stanley and Björner 2010, 24).[28]

7. Comparison to Other Proposals

I will briefly contrast my approach to proofs that explain with several others in the recent literature.

28. I do not contend that any bijective proof (regardless of the bijection's complexity or artificiality) is explanatory, that no generating function proof is explanatory, or that there is never a good reason to seek a generating-function proof once a simple, explanatory bijective proof has been found. Each kind of proof may have some value; for instance, a generating-function proof may be much shorter or help us to find proofs of new theorems. Also bear in mind that "the precise border between combinatorial [that is, bijective] and non-combinatorial proofs is rather hazy, and certain arguments that to an inexperienced enumerator will appear non-combinatorial will be recognized by a more facile counter as combinatorial, primarily because he or she is aware of certain standard techniques for converting apparently non-combinatorial arguments into combinatorial ones" (Stanley 2012, 13).

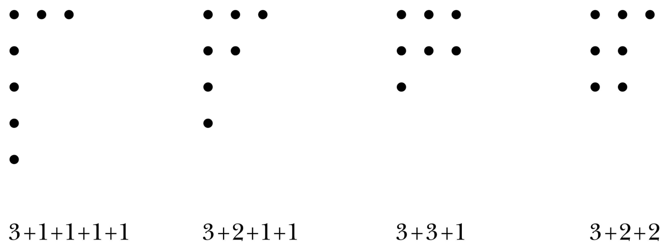

Figure 10. The previous dot diagrams, rotated

According to Steiner (1978a), a proof explains why all S_1s are P_1 if and only if it reveals how this theorem depends on S_1's "characterizing property"—that is, on the property essential to being S_1 that is just sufficient to distinguish S_1s from other entities in the same "family" (for example, to distinguish triangles from other kinds of polygons). To reveal the theorem's dependence on the characterizing property, the proof must be "generalizable." That is, if S_1's characterizing property is replaced in the proof by the characterizing property for another kind S_2 in the same family (but the original "proof idea" is maintained), then the resulting "deformation" of the original proof proves that all S_2s are P_2 for some property P_2 incompatible with being P_1. Thus, the theorem's explanation helps to show that there are different, but analogous, theorems for different classes in the same family.

Steiner's proposal nicely accommodates some of the explanatory proofs that I have examined (at least under a natural reading of the relevant "family" and "proof idea"). For instance, the proof I presented as explaining why the product of any three consecutive nonzero natural numbers is divisible by 6 ($=1 \times 2 \times 3$) could be deformed to prove that the product of any four consecutive nonzero natural numbers is divisible by 24 ($=1 \times 2 \times 3 \times 4$). However, this four-number result could also be proved by mathematical induction—by a proof that is a "deformation" of the proof by mathematical induction of the three-number result. Yet (I have argued) the inductive proof of the three-number result is not explanatory (in a context where the result's treating every case alike is salient), and its "generalizability" in Steiner's sense does not at all incline me to reconsider that verdict.[29] The proof treats the first triple of nonzero

29. Steiner (1978a, 151) says that "inductive proofs usually do not allow deformation" and hence are not explanatory, because by replacing S_1's characterizing property with S_2's in the original inductive proof, we do not automatically replace the original theorem

natural numbers differently from every other triple, rather than identifying a property common to every triple that makes each divisible by 6.

Furthermore, some of the explanatory proofs that I have identified simply collapse rather than yield new theorems when they are deformed to fit a different class in what is presumably the same "family." For instance, the proof from the isosceles trapezoid's symmetry does not go anywhere when we shift to a nonisosceles trapezoid, since the symmetry then vanishes.[30] For some of the other explanatory proofs I have presented, it is unclear what the relevant "family" includes. What, for instance, are the other classes in the family associated with d'Alembert's theorem that roots come in complex-conjugate pairs? Even if we could ultimately discover such a family, we do not need to find it in order to recognize the explanatory power of the proof exploiting the setup's invariance under the replacement of i with $-i$. (I will return to this point momentarily.)

I turn now to Kitcher (1984, esp. 208–9, 227; 1989, esp. 423–26, 437), who offers a unified account of mathematical and scientific explanation; in fact, he sees explanation of all kinds as involving unification. Roughly speaking, Kitcher says that an explanation unifies the fact being explained with other facts by virtue of their all being derivable by arguments of the same form. Explanations instantiate argument schemes in the optimal collection ("the explanatory store")—optimal in that arguments instantiating these schemes manage to cover the most facts with the fewest different argument schemes placing the most stringent constraints upon arguments. An argument instantiating an argument scheme excluded from the explanatory store fails to explain.

at the start of the inductive step with the new theorem to be proved—and in an inductive argument, the theorem must be introduced at the *start* of the inductive step. It seems to me, however, that if we take the inductive proof that the product of any three consecutive nonzero natural numbers is divisible by 6 ($= 1 \times 2 \times 3$) and replace the initial reference to three consecutive nonzero natural numbers with a reference to four consecutive nonzero natural numbers, then the first step of the inductive proof is automatically that the first case of four consecutive nonzero natural numbers is obviously divisible by their product ($1 \times 2 \times 3 \times 4 = 24$), and this gives us immediately the new theorem to be proved (namely, that the product of any four consecutive nonzero natural numbers is divisible by 24) for use at the start of the second step. So in this case, the inductive proof permits deformation.

30. Resnik and Kushner (1987, 147–52) likewise argue that the standard proof of the intermediate value theorem (which, they say, is explanatory if any proofs are) simply collapses if any attempt is made to shift it to cover anything besides an interval over the reals (such as over the rationals or a disjoint pair of intervals over the reals).

Thus, Kitcher sees a given mathematical proof's explanatory power as arising from the proof's relation to other proofs (such as their all instantiating the same scheme or their covering different facts). My account contrasts with Kitcher's (and with Steiner's) in doing justice to the fact that (as we have seen in various examples) we can appreciate a proof's explanatory power (or impotence) just from examining the details of that proof itself, without considering what else could be proved by instantiating the same scheme (or "proof idea") or how much coverage the given proof adds to whatever is covered by proofs instantiating other schemes. In addition, Kitcher regards all mathematical explanations as deriving their explanatory power from possessing the same virtue: the scheme's membership in the "explanatory store." It seems to me more plausible (especially given the diversity of our examples) to expect different mathematical explanations to derive their explanatory power by virtue of displaying different traits. On my approach, different traits are called for when the result being explained has different salient features.

A typical proof by "brute force" uses a "plug and chug" technique that is perforce applicable to a very wide range of problems. Presumably, then, its proof scheme is likely to belong to Kitcher's "explanatory store." (Not every brute-force proof instantiates the same scheme, but a given brute-force proof instantiates a widely applicable scheme.) For example, as I mentioned in section 5, we could prove the theorem regarding isosceles trapezoids by expressing the setup in terms of coordinate geometry and then algebraically grinding out the result. The same strategy could be used to prove many other geometric theorems. Nevertheless, these proofs lack explanatory power. This brute-force proof of the trapezoid theorem is unilluminating because it begins by expressing the entire setup in terms of coordinate geometry and never goes on to characterize various particular features of the setup as irrelevant. Consequently, it fails to pick out any particular feature of isosceles trapezoids (such as their symmetry) as the feature responsible for the theorem.

A new proof technique can explain why some theorem holds even if that technique allows no new theorems to be proved. My approach can account for this feature of mathematical explanation. It is more difficult to accommodate on Kitcher's proposal, since any explanatory argument scheme must earn its way into the "explanatory store" by adding coverage (without unduly increasing the number of schemes or decreasing the stringency of their constraints).[31]

I turn finally to Resnik and Kushner (1987), who doubt that any proofs explain *simpliciter.* They contend that a proof's being "explanatory" to a given audience is nothing more than its being the kind of proof that the audience wants—perhaps in view of its premises, its strategy, its perspicuity, or the collateral information it supplies (or perhaps any proof whatsoever of the theorem would do). Contrary to Resnik and Kushner, I do not think that whenever someone wants a certain kind of proof, for whatever reason, then such a proof qualifies for them as explaining why the theorem being proved holds. Rather, mathematical practice shows that an explanatory proof requires some feature of the result to be salient and requires the proof to exploit a similar noteworthy feature in the problem. Thus, the demand for an explanation is not simply the demand for a certain kind of proof; the demand arises from a certain feature of the result and is satisfied only by a proof that involves such a feature from the outset. For example, we may want to see a proof of the "calculator number" theorem that proceeds by checking each of the 16 calculator numbers individually. But this proof merely heightens our curiosity, motivating us to seek the reason why all of the calculator numbers are divisible by 37. It is not the case that any kind of proof we happen to want counts as an explanation when we want it. Moreover, explanatory power is just one of many properties the possession of which would make a proof desirable.

8. Conclusion

I have tried to identify the basis on which certain proofs but not others are explanatory. Symmetry, unity, and simplicity are among the aspects of mathematical results that are commonly salient. But many other, less common features are sometimes outstanding in mathematical results and may thereby inform the distinction between proofs that explain and proofs that do not. Admittedly, several fairly elastic notions figure in my idea of a proof's exploiting the same kind of feature in the problem as was salient in the result. This elasticity allows my proposal to encompass a wide range of cases (as I have shown). Insofar as the notions figuring in my proposal have borderline cases, there will correspondingly be room

31. Tappenden (2005) offers a similar objection to this "winner take all" feature of Kitcher's account.

for mathematical proofs that are borderline explanatory.[32] However, the existence of such cases would not make a proof's explanatory power rest merely "in the eye of the beholder."

As I noted at the outset, it is challenging to find a source of explanatory asymmetry for mathematical explanations, since the usual suspects in scientific explanation (such as causal and nomological priority) are unavailable. In response, I have gestured toward the priority that axioms in mathematics have over theorems, but I have also emphasized mathematical explanations that operate in connection with "problems," each of which is characterized by a "setup" and a "result." This structure of setup and result adds an asymmetry that enables mathematical explanation to get started by allowing why questions to be posed (as in our first example, where Zeitz [2000, 5] declares that he wants to "understand *why* the coin problem [that is, the setup] had the answer [that is, the result] that it did"). When we consider some proof that extracts the property of being G from the property of being F, whether the proof is explanatory may depend on whether we take the setup and result as involving being F and being G, respectively, or instead as (for example) being non-G and being non-F. Whether the proof is explanatory may also depend on which features of the theorem are salient to us. But despite these context-sensitive features, it does not follow that a proof is explanatory merely by virtue of striking its audience as explanatory.

Of course, if some extraterrestrials differ from us in which features of a given theorem they find salient, then it follows from my account that those extraterrestrials will also differ from us in which proofs they ought to regard as explanatory. I embrace this conclusion. In different contexts, *we* properly regard different proofs of the same theorem to be explanatory—namely, in contexts where different features of the theorem are salient. Furthermore, if some extraterrestrials differed from us so much that they never regarded symmetry, unity, and simplicity as salient, then even if we and they agreed on the truth of various mathematical theorems, their practices in seeking and refining proofs of these theorems would differ so greatly from ours that it would be a strain to characterize them as doing mathematics.[33] As cases such as the embrace of imaginary

32. See notes 16 and 21.

33. I have specified that the extraterrestrials *never* regard symmetry as salient, not merely that they fail to regard a particular symmetry in a given theorem as salient, because our attention can reasonably be drawn to one theorem's symmetry by that theorem's being compared to another theorem exhibiting a similar symmetry that is already salient.

numbers illustrate, the search for mathematical explanations often drives mathematical discovery and innovation. Such extraterrestrials would not be seeking the same things as we do in doing mathematics.

The role that salience plays in my account does not suggest that when mathematicians talk about a given proof's "explanatory power," they are merely gesturing toward some aesthetically attractive quality that the proof possesses—a quality that is very much in the eye of the beholder. Salience's role does not make mathematical explanations differ sharply from all scientific explanations. On the contrary, some scientific explanations operate in the same way as the mathematical explanations I have examined. For instance, consider the notorious (or, if you prefer, "mildly famous") puzzle, "Why are mirror images reversed sideways but not up and down?" (Bennett 1970, 181; see also Block 1974).[34] This problem involves a setup: typically (as in Martin 2002, 176), you are standing before a full-length mirror, wearing a ring on your left hand. It also involves a result: your mirror image. The result's ring is on its right hand, but it is not standing on its head. The why question is asking for a derivation in which the result, with its salient asymmetry between left-right and up-down, is traced to a similar asymmetry in the setup. The demand is not for a causal or nomological explanation. Indeed, a causal account of the reflection as given by geometrical optics will not suffice to answer this why question: "Where does the asymmetry come from? . . . To explain horizontal but not vertical reversal with optical ray diagrams is doomed to failure, for they are symmetrical and equally valid when held in any orientation" (Gregory 1987, 492; see also Block 1974, 267). That is, a mirror reflects light rays symmetrically about the normal to the surface at the point of incidence. No asymmetry lurks here. The why question requires an answer that uncovers an asymmetry in the setup that is similar to the salient asymmetry in the result. Here is such an answer to the why question:

> Why do you count that mirror image as right/left reversed? Because you imagine turning yourself so that you would face in the same direction as your mirror-image now faces . . . and moving in back of the mirror to the place where it appears your image now stands. Having turned and moved, your hand with the ring on it is in the place where the un-ringed hand of your mirror-image is. . . . In other words, when you turn this way . . . [your

34. I am grateful to John Roberts for suggesting the mirror example as an apt comparison.

> mirror image is] reversed left-to-right, compared to you. (Martin 2002, 176–77; see also Block 1974, Gregory 1987)

The setup's asymmetry, responsible for the asymmetrical privileging of one dimension (left-right) over the others, lies in a hidden feature of the setup: how you imagine moving from facing the mirror to facing out from behind it, namely, by rotating yourself around the vertical axis. This feature of the setup privileges that axis over the others. It answers the why question by tracing the result, with its salient asymmetry, back to a similar asymmetry in the setup. After all, we could imagine a different operation by which you could go from facing the mirror to facing out from behind it: by rotating yourself around a horizontal axis. Then you would be on your head behind the mirror and so reversed vertically (but not left-right) relative to your mirror image. Relative to that means of getting behind the mirror, the image is reversed vertically but not left-right. So mirrors reverse left-right rather than up-down only given the particular way we imagine getting behind them—that is, only by virtue of the salience of one axis for turning ourselves behind a mirror.[35]

Mathematicians do occasionally reflect upon explanation in mathematics. For instance, Timothy Gowers (2000, 73) writes, "[Some] branches of mathematics derive their appeal from an abundance of mysterious phenomena that demand explanation. These might be striking numerical coincidences suggesting a deep relationship between areas that appear on the surface to have nothing to do with each other, arguments which prove interesting results by brute force and therefore do not satisfactorily explain them, proofs that apparently depend on a series of happy accidents." I hope that this essay has managed to unpack some of these provocative remarks.

References

Abel, Theodore. 1953 [1948]. "The Operation Called *Verstehen*." In *Readings in the Philosophy of Science*, ed. Herbert Feigl and May Brodbeck, 677–87. New York: Appleton-Century-Crofts.

35. I do not need to claim that this answer to the why question about mirrors is actually correct. (Perhaps the correct answer instead rejects the explanandum: mirrors do not flip left-right or up-down, but rather front-back.) But even if the common answer I have mentioned is incorrect, its error is not in the *way* that it purports to explain—and my point is simply to show that some respectable purported scientific explanations aim to work in the same way as the explanations in mathematics that I have examined.

Andrews, George E., and Kimmo Eriksson. 2004. *Integer Partitions.* Cambridge: Cambridge University Press.

Atiyah, Michael. 1988. "How Research Is Carried Out." In *Collected Works,* vol. 1, 211–6. Oxford: Clarendon.

Baker, Alan. 2005. "Are There Genuine Mathematical Explanations of Physical Phenomena?" *Mind* 114: 223–38.

———. 2009. "Mathematical Accidents and the End of Explanation." In *New Waves in Philosophy of Mathematics,* ed. Otávio Bueno and Øystein Linnebo, 137–59. Houndmills, UK: Palgrave Macmillan.

Bell, A. W. 1976. "A Study of Pupils' Proof-Explanations in Mathematical Situations." *Educational Studies in Mathematics* 7: 23–40.

Bennett, Jonathan. 1970. "The Difference Between Right and Left." *American Philosophical Quarterly* 7: 175–91.

Block, Ned. 1974. "Why Do Mirrors Reverse Left/Right but Not Up/Down?" *Journal of Philosophy* 71: 259–77.

Bogomolny, A. n.d. "Menelaus Theorem: Proofs Ugly and Elegant: A. Einstein's View." www.cut-the-knot.org/Generalization/MenelausByEinstein.shtml#F (accessed April 12, 2009).

Bressoud, David M. 1999. *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture.* Cambridge: Cambridge University Press.

Brunhes, Edmond [Frère Gabriel-Marie]. 1991 [1920]. *Exercices de Géometrié.* 6th ed. Paris: Editions Jacques Gabay.

Byers, William. 2007. *How Mathematicians Think.* Princeton: Princeton University Press.

Gale, David. 1990. "Proof as Explanation." *Mathematical Intelligencer* 12, no. 1: 4.

———. 1991. "Mathematical Entertainments." *Mathematical Intelligencer* 13, no. 1: 40–2.

Gannon, Terry. 2006. *Moonshine beyond the Monster.* Cambridge: Cambridge University Press.

Gowers, W. T. [Timothy]. 2000. "The Two Cultures of Mathematics." In *Mathematics: Frontiers and Perspectives,* ed. V. Arnold, M. Atiyah, P. Lax, and B. Mazur, 65–78. Providence, RI: American Mathematical Society.

———. 2007. "Mathematics, Memory, and Mental Arithmetic." In *Mathematical Knowledge,* ed. Mary Leng, Alexander Paseau, and Michael Potter, 33–58. Oxford: Oxford University Press.

Gray, Jeremy. 2007. *Worlds Out of Nothing.* London: Springer.

Gregory, Richard L. 1987. "Mirror Reversal." In *The Oxford Companion to the Mind,* ed. Richard L. Gregory, 491–93. Oxford: Oxford University Press.

Grosholz, Emily. 2000. "The Partial Unification of Domains, Hybrids, and the Growth of Mathematical Knowledge." In *The Growth of Mathematical Knowledge,* ed. Emily Grosholz and Herbert Breger, 81–91. Dordrecht: Kluwer.

Hafner, Johannes, and Paolo Mancosu. 2005. "The Varieties of Mathematical Explanation." In *Visualization, Explanation and Reasoning Styles in Mathemat-*

ics, ed. P. Mancosu, K. F. Jørgensen, and S. A. Pedersen, 215–50. Dordrecht: Springer.

Harari, Orna. 2008. "Proclus' Account of Explanatory Demonstrations in Mathematics and Its Context." *Archiv für Geschichte der Philosophie* 90, no. 2: 137–64.

Healy, Lulu, and Celia Hoyles. 2000. "A Study of Proof Conceptions in Algebra." *Journal for Research in Mathematics Education* 31: 396–428.

Hempel, Carl G. 2001 [1963]. "Explanation and Prediction by Covering Laws." In *The Philosophy of Carl G. Hempel*, ed. James Fetzer, 69–86. New York: Oxford University Press.

Kiernan, B. Melvin. 1971. "The Development of Galois Theory from Lagrange to Artin." *Archive for History of Exact Sciences* 8: 40–154.

Kitcher, Philip. 1984. *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

———. 1989. "Explanatory Unification and the Causal Structure of the World." In *Minnesota Studies in the Philosophy of Science*. Vol. 13, *Scientific Explanation*, ed. Philip Kitcher and Wesley Salmon, 410–505. Minneapolis: University of Minnesota Press.

Kitcher, Philip, and Wesley Salmon. 1987. "Van Fraassen on Explanation." *Journal of Philosophy* 84: 315–30.

Kline, Morris. 1972. *Mathematical Thought from Ancient to Modern Times*. New York: Oxford University Press.

Lagrange, Joseph-Louis. 1826 [1808]. "Note XIII: Sur la Résolution des Équations Algébriques." In *Traité de la Résolution des Équations Numériques de Tous les Degrés*, 242–72, 3rd ed. Paris: Bachelier.

———. 1869 [1770–71]. "Refléxions sur la Résolution Algébrique des Équations." In *Œuvres de Lagrange*, vol. 3, ed. J. -A. Serret, 205–421. Paris: Gauthier-Villars.

Lange, Marc. 2010. "What Are Mathematical Coincidences (and Why Does It Matter)?" *Mind* 119: 307–40.

———. 2013. "What Makes a Scientific Explanation Distinctively Mathematical?" *British Journal for the Philosophy of Science* 64: 485–511.

———. n.d. "Explanation, Existence, and Natural Properties in Mathematics—A Case Study: Desargues' Theorem." Unpublished manuscript.

Levi, Mark. 2009. *The Mathematical Mechanic*. Princeton: Princeton University Press.

Lewis, David. 1986. "Causal Explanation." In *Philosophical Papers*, vol. 2, 214–40. New York: Oxford University Press.

Luchins, Abraham S., and Edith H. Luchins. 1990. "The Einstein-Wertheimer Correspondence on Geometric Proofs and Mathematical Puzzles." *Mathematical Intelligencer* 12, no. 2: 35–43.

Mancosu, Paolo. 2008. “Mathematical Explanation: Why It Matters.” In *The Philosophy of Mathematical Practice*, ed. Paolo Mancosu, 134–50. Oxford: Oxford University Press.

Martin, Robert M. 2002. *There Are Two Errors in the the Title of This Book*. Peterborough, ON: Broadview Press.

Mill, John Stuart. 1872. *A System of Logic*. Vol. 1. London: Longmans, Green.

Mudaly, Vimolan, and Michael de Villiers. 2000. “Learners’ Needs for Conviction and Explanation within the Context of Dynamic Geometry.” *Pythagoras* 52: 20–3.

Nagel, Ernest. 1961. *The Structure of Science*. New York: Harcourt, Brace.

Nicole, Pierre, and Antoine Arnauld. 1717. *Logic; or the Art of Thinking*. London: William Taylor.

Nummela, Eric. 1987. “No Coincidence.” *Mathematical Gazette* 71, no. 456: 147.

Psillos, Stathis. 2002. *Causation and Explanation*. Chesham: Acumen.

Rescher, Nicholas. 1970. *Scientific Explanation*. New York: Free Press.

Resnik, Michael, and David Kushner. 1987. “Explanation, Independence and Realism in Mathematics.” *British Journal for the Philosophy of Science* 38: 141–58.

Salmon, Wesley. 1989. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.

Smart, J. J. C. 1990. “Explanation—Opening Address.” In *Explanation and Its Limits*, ed. Dudley Knowles, 1–20. Cambridge: Cambridge University Press.

Sober, Elliott. 1984. *The Nature of Selection*. Chicago: University of Chicago Press.

Sorensen, Roy. n.d. “Mathematical Coincidences.” Unpublished manuscript.

Spivak, Michael. 1980. *Calculus*. 2nd ed. Berkeley: Publish or Perish.

Stankova, Zvezdelina. 2004. “Geometric Puzzles and Constructions—Six Classical Geometry Theorems.” In *Mathematical Adventures for Students and Amateurs*, ed. David Hayes and Tatiana Shubin, 169–84. Washington, DC: Mathematical Association of America.

Stanley, Richard P. 2012. *Enumerative Combinatorics*. 2nd ed. Vol. 1. Cambridge: Cambridge University Press.

Stanley, Richard P., and Anders Björner. 2010. *A Combinatorial Miscellany*. Geneva: L’Enseignement Mathématique.

Steiner, Mark. 1978a. “Mathematical Explanation.” *Philosophical Studies* 34: 135–51.

———. 1978b. “Mathematics, Explanation, and Scientific Knowledge.” *Noûs* 12: 17–28.

Stillwell, John. 2005. *The Four Pillars of Geometry*. New York: Springer.

Sylvester, J. J. 1882. “A Constructive Theory of Partitions, Arranged in Three Acts, an Interact and an Exodion.” *American Journal of Mathematics* 5: 251–330.

Tappenden, Jamie. 2005. “Proof Style and Understanding in Mathematics I: Visualization, Unification, and Axiom Choice.” In *Visualization,* *Explanation, and Reasoning Styles in Mathematics*, ed. P. Mancosu, K. F. Jørgensen, and S. A. Pedersen, 147–214. Dordrecht: Springer.

Van Fraassen, Bas. 1980. *The Scientific Image*. Oxford: Clarendon.

Vinner, Shlomo, and Evgeny Kopelman. 1998. "Is Symmetry an Intuitive Basis for Proof in Euclidean Geometry?" *Focus on Learning Problems in Mathematics* 20, no. 2–3: 14–26.

Wilf, Herbert. 2000. "Lectures on Integer Partitions." www.math.upenn.edu/~wilf/PIMS/PIMSLectures.pdf (accessed September 24, 2009).

Zeitz, Paul. 2000. "Graph Theory." Paper presented at the Berkeley Math Circle, October 8.

———. 2007. *The Art and Craft of Problem Solving*. 2nd ed. Hoboken: Wiley.