

Ground and Explanation in Mathematics

Marc Lange

University of North Carolina at Chapel Hill

© 2019 Marc Lange

*This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 3.0 License.
<www.philosophersimprint.org/019033/>*

Increased attention has recently been paid to the fact that in mathematical practice, certain mathematical proofs but not others are recognized as explaining why the theorems they prove obtain (Mancosu 2008; Lange 2010, 2015a, 2016; Pincock 2015). Such “mathematical explanation” is presumably not a variety of causal explanation. In addition, the role of metaphysical grounding as underwriting a variety of explanations has also recently received increased attention (Correia and Schnieder 2012; Fine 2001, 2012; Rosen 2010; Schaffer 2016). Accordingly, it is natural to wonder whether mathematical explanation is a variety of grounding explanation. This paper will offer several arguments that it is not.

One obstacle facing these arguments is that there is currently no widely accepted account of either mathematical explanation or grounding. In the case of mathematical explanation, I will try to avoid this obstacle by appealing to examples of proofs that mathematicians themselves have characterized as explanatory (or as non-explanatory). I will offer many examples to avoid making my argument too dependent on any single one of them. I will also try to motivate these characterizations of various proofs as (non-) explanatory by proposing an account of what makes a proof explanatory. In the case of grounding, I will try to stick with features of grounding that are relatively uncontroversial among grounding theorists. But I will also look briefly at how some of my arguments would fare under alternative views of grounding. I hope at least to reveal something about what grounding would have to look like in order for a theorem’s grounds to mathematically explain why that theorem obtains. I hope thereby to identify some of the difficulties confronting a too-quick identification of grounding relations as underwriting explanation in mathematics.

1. Proofs that Identify Grounds versus Proofs that Explain.

Suppose you write down all of the whole numbers from 1 to 99,999. How many times would you write down the digit 7? The answer turns out to be 50,000 times. That is a striking result since (as you have probably noticed) 50,000 is half of 100,000, which is just one more than 99,999.

One way to prove this result would be to list all of the whole numbers from 1 to 99,999; to note, for each entry in the list, the number of times the digit 7 appears in it; and to add these numbers of appearances, arriving at 50,000 in total. This proof, I maintain, gives *grounds* of the fact that 7 appears 50,000 times in total from 1 to 99,999. Here, I aim to be using the notion of “ground” in the manner intended by Fine (2001, 2012) and many other grounding theorists. On this view, a fact’s grounds are whatever it is in virtue of which that fact obtains, and a truth-bearer (such as a proposition) is grounded in its truth-makers.

In particular, by regarding the fact that 7 appears 50,000 times in the list from 1 to 99,999 as grounded in facts about the individual appearances of 7’s in that list, I am making the rough presupposition that a mathematical fact is grounded by the atomic (or negated atomic) truths to which one is led if one starts with that fact and moves “downward” to logically simpler truths in an obvious way, such as from universal facts (i.e., facts expressed by generalizations) to their instances and from conjunctions to their conjuncts.¹ This presupposition seems to lie behind many plausible-looking claims about the grounds of mathematical facts. For instance, this presupposition motivates regarding the fact that every number in the sequence 31, 331, ..., 3333331 is prime as grounded by the fact that 31 is prime, the fact that ..., and the fact that 3333331 is prime. Likewise, this presupposition motivates regarding the fact that there is exactly one prime number between 12 and 15 (inclusive) as grounded in the fact 13 is prime but 12, 14, and 15 are composite. In the same way, it motivates regarding the fact that 7 appears 50,000 times in the list from 1 to 99,999 as grounded in facts about the particular locations where 7 appears (or does not appear) in that list.

Let’s return to the ponderous proof that tallies every appearance of a 7 in the list of numbers from 1 to 99,999. That the number of 7’s turns

1. My thanks to a referee who urged me to be more explicit about these presuppositions concerning grounding and offered this formulation as a candidate for doing so. In the next section, I will consider briefly how my arguments would fare under an alternative picture of the grounds of mathematical facts.

out to be almost exactly half of 99,999 is treated by this proof as if it just turned out that way as a matter of mathematical coincidence. Of course, it is mathematically necessary that the number of 7’s is almost equal to half of 99,999; this result is obviously not a matter of chance. Nevertheless, as far as this proof shows, the fact that the number of 7’s is almost equal to half of 99,999 arises from two unrelated facts: that there are 50,000 7’s and that 50,000 is half of 100,000 (one more than 99,999). As far as this proof reveals, one of these facts is not responsible for the other and there is no important common reason behind both.² The proof arrives at the number of 7’s without doing anything like taking half of 100,000. Thus, the proof makes it seem like sheer happenstance that the number of 7’s is very nearly half of 99,999.

By contrast, consider this proof of the same result:

Include 0 among the numbers under consideration — this will not change the number of times the digit 7 appears. Suppose all the whole numbers from 0 to 99,999 (100,000 of them) are written down with five digits each, e.g., 1306 is written as 01306. All possible five-digit combinations are now written down, once each. Because every digit will take every position equally often, every digit must occur the same number of times overall. Since there are 100,000 numbers with five digits each — that is, 500,000 digits — each of the 10 digits appears 50,000 times. That is, $100,000 \times 5/10 = 50,000$. (Dreyfus and Eisenberg 1986, 3)³

This proof does not specify where in the list of numbers each 7 appears; it does not give the result’s ground. Nevertheless, this proof is more illuminating than the first one in that it *explains why* the number of

2. There are many examples of mathematical coincidences, such as the fact that (in base 10) the 13th digit of π is equal to the 13th digit of e . For that example and more on “mathematical coincidence”, see Lange (2010, 2016).
3. Thanks to Many Sundström for calling this article to my attention.

7's is almost exactly half of 99,999. It reveals where the one-half of 100,000 comes from: There are 100,000 numbers being written down, there are 5 digits in each number, there are 10 options (in base 10) for each of these digits, and 5 divided by 10 is one-half. By virtue of this proof, I suggest, it is no mathematical coincidence that the total number of 7's is almost exactly half of 99,999.

In section 3, I will have more to say more about what *makes* the second proof explanatory, by contrast with the first proof. For now, my aim is simply to prompt you to appreciate that the first proof (unlike the second) does *not* explain why the result holds, despite giving the result's ground, whereas the second proof explains why the result holds despite failing to identify the result's ground.

It might be objected that although the first proof displays the result's grounds, this fact does not preclude the second proof from also doing so; a given fact can have many complete sets of grounds. For instance, a given fact's grounds can themselves have grounds, and the latter may then qualify as grounds of the given fact.⁴ I have three replies to this objection. Firstly, part of what I regard this example as showing is that the grounds of a mathematical theorem do not automatically (that is, simply by virtue of being its grounds) explain it. Even if the second proof also gives the result's grounds, that the first proof gives its grounds without explaining why the result holds suffices to show that a theorem's grounds do not automatically explain it. If mathematical explanation worked by tracing the explanandum to its grounds, then it would be puzzling why the first proof does not qualify as a mathematical explanation. Secondly, it is not obvious why one should insist that the second proof gives the result's grounds. For instance, it is not the case that the second proof gives grounds of any of the various facts (namely, that the digit 7 appears nowhere in 1, ... appears twice in 7007, ...) that the first proof displays as grounding the result. Likewise, although plausibly the disjunction of two truths

4. Virtually all accounts of grounding regard grounding as transitive (Correia and Schnieder 2012, 32; Fine 2012, 56)—though see note 16 for one circumstance in which (according to Schaffer 2016) transitivity may be violated.

has distinct full grounds in each of the disjuncts (e.g., Correia and Schnieder 2012, 17; Fine 2012, 58), that precedent does little to suggest that the second proof and the first proof cite distinct full grounds, since the fact being explained is not a disjunction. The second proof might be construed as describing the result's grounds (the decimal expressions of the whole numbers from 1 to 99,999) at a high level of abstraction: in terms of the distribution of each digit in the list, without specifying where in the list each appears. But in the next section, as my third reply to the objection, we will see examples where explanatory proofs arguably do not even describe the result's grounds at some high level of abstraction. (That is because the explanatory proofs in those examples are "impure".)

In this paper, I will argue that the phenomenon on display in this example is widespread: Many mathematical proofs that explain why some result holds fail to present the grounds of the theorem being proved and many mathematical proofs that present the grounds of the theorem being proved fail to explain why that theorem obtains. I will identify several reasons why grounds and explanations tend to come apart in mathematics.

To begin to show why this is typical, I will give another example.⁵ Consider these two Taylor series:

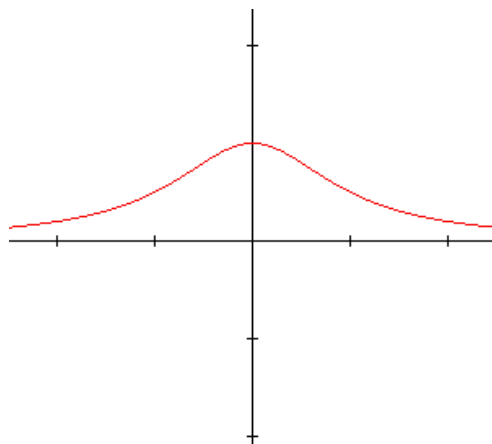
$$1 / (1 - x^2) = 1 + x^2 + x^4 + x^6 + \dots$$

$$1 / (1 + x^2) = 1 - x^2 + x^4 - x^6 + \dots$$

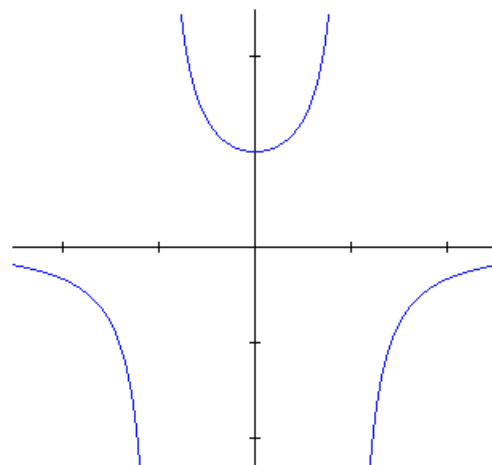
They have the same convergence behavior: For real x , each converges when $|x| < 1$ but diverges when $|x| > 1$. That is a remarkable similarity, and it stands out especially strongly against the obvious difference between the two functions regarding $x = 1$: As their graphs (below) show, although one function goes undefined at $x = 1$, the other function behaves quite soberly there. That the two functions' Taylor series have the same convergence behavior might therefore seem to be utterly

5. This example has been discussed by Steiner (1978, 18–19), Waisman (1982, 29–30), and Wilson (2006, 313–314).

coincidental, considering that the two functions appear to be quite dissimilar in other respects.



$$f(x) = 1 / (1 + x^2)$$



$$f(x) = 1 / (1 - x^2)$$

A *ground* for the fact that both series have this convergence behavior is that the first series exhibits this behavior and the second series does, too. A conjunction is grounded in its conjuncts.⁶ A proof of the conjunction that gives entirely separate proofs of its two conjuncts does nothing to suggest that it is no coincidence that the two functions share this convergence behavior.

However, it is in fact no coincidence, as we discover by looking at matters on the complex plane. The “radius of convergence theorem” says that a complex Taylor series will converge at all and only the points within a certain radius of the origin on the complex plane:

Radius of convergence theorem: For any power series $\sum a_n z^n$ (from $n = 0$ to ∞), either it converges for all complex numbers z , or it converges only for $z = 0$, or there is a number $R > 0$ such that it converges if $|z| < R$ and diverges if $|z| > R$.

The two Taylor series share the same convergence behavior because the two functions have another feature in common: They both go undefined at some point on the unit circle centered at the origin of the complex plane. This similarity explains why the two Taylor series behave alike. As mathematician Michael Spivak says, a proof of the convergence theorem “helps explain the behavior of certain Taylor series obtained for real functions” (Spivak 1980, 528), such as in the example I have mentioned. Thus, a proof that displays the *ground* of this mathematical result fails to explain why the result holds because such a proof fails to give the two functions’ common convergence behavior the same proof from another feature that the two functions share — when, in fact, they can be so proved.

With the Taylor series example and the tallying 7’s example, I have aimed (i) to argue that a mathematical fact’s grounds do not, simply by virtue of grounding it, thereby explain why that fact obtains and

6. As Correia and Schnieder (2012, 17) report, this fact about a conjunction’s grounds is widely acknowledged in the grounding literature (e.g., Fine 2012, 63; Schaffer 2016, 53). But shortly I will look briefly at an alternative view.

(ii) to suggest that these two examples are typical in that oftentimes, a mathematical proof that specifies a fact's grounds fails to explain why that fact obtains whereas any explanation of the fact does not specify its ground. None of these claims (for which I will argue further) suggests that no proof identifying some mathematical fact's grounds ever explains why that fact obtains. But I contend that even in such a case, what makes a given proof explanatory is *not* that it identifies the grounds of what it proves.

In section 2, I will give some further examples that distinguish proofs that explain from proofs that specify grounds, and I will give some further reasons why explaining and grounding tend to come apart. In particular, I will focus on the fact that explanatory proofs need not exhibit purity, tend not to be brute force, and often unify separate cases by identifying common reasons behind them even when those cases have distinct grounds. In section 3, I will sketch an account of what makes a proof explanatory. I will then use that account to defend the morals I have drawn from the examples I have already given. In section 4, I will use that account of explanation in mathematics to give two further reasons why proofs that specify grounds tend not to be proofs that explain. In particular, I will argue that for many theorems, there is no proof that explains why they obtain, even though these theorems have grounds. I will also argue that the role that context plays in making a proof explanatory ensures that explanatory power frequently diverges from grounding (since, unlike having a certain explanation, a theorem's having certain grounds is insensitive to context).

The question this paper aims to address is whether a proof's power to explain what it proves stands in any general relation to its specifying the ground of what it proves. Accordingly, this paper takes for granted that in mathematical practice, certain proofs differ from others in the degree to which they explain why a given theorem holds.⁷

7. For arguments that there is such a distinction (or matter of degree) of explanatoriness in mathematical practice, see Mancosu (2008), Lange (2016), and references therein. Pincock (2015, 11) gives a brief argument (distinct from

I will argue that a mathematical fact's grounds do not, by virtue of grounding it, explain why it obtains. We could nevertheless insist that grounds automatically explain in mathematics by making a bare stipulation: that grounds supply a *special kind* of explanation ("*grounding* explanations"). But this stipulation would not reveal that grounding is connected to explanation of the kind that figures in mathematical practice. That kind of explanation will be my exclusive focus throughout.

2. Purity, Brute Force, and Unification.

The proof that explains why the two Taylor series exhibit the same convergence behavior places the two Taylor series in a broader context by putting the two functions on the complex plane. By introducing imaginary numbers, the proof introduces concepts exogenous to the theorem being proved, since the original two Taylor series involve only real numbers. Thus, one lesson of the Taylor series example is that an *impure* proof can sometimes explain what it proves.

By a "pure" proof, I mean roughly a proof using only the concepts that are (in the sense I will clarify below) "intrinsic" (as mathematicians put it) to the theorem being proved so that "the resources of proof [are] restricted to those which determine [the theorem's] content" (Detlefsen 2008, 193).⁸ Presumably, a proof giving the *grounds* of the theorem being proved must be a pure proof. An impure proof is commonly said to introduce "extraneous" concepts — that is, presumably, alien to the

any argument I will give) that for the theorem that the quintic is unsolvable, an explanatory proof does not supply the theorem's ground.

8. For more on purity, see Detlefsen (2008), Detlefsen and Arana (2011), and references therein. How to make "purity" precise is a topic for another paper (though see the next paragraph). Different readings of "purity" could come apart in some cases, but I will argue that even under a liberal reading of "purity" that deems the appeal to imaginary numbers to be a pure proof of the fact about the two Taylor series, this proof's "purity" does not contribute to making it explanatory.

theorem and its grounds. Nevertheless, mathematicians recognize that an impure proof is in some cases more explanatory than a pure proof.⁹

It might be objected that a proof giving the grounds of the theorem being proved need not be a pure proof. A concept could fail to figure in the theorem, in the definitions of the concepts figuring in the theorem, in the definitions of the concepts figuring in those definitions, and so on, and nevertheless figure in the theorem's ground (just as one could have mastered the concepts of heat, temperature, and so forth without having any grasp of the concept of the kinetic energy of random molecular motion). However, it seems to me that as "purity" is understood in mathematics, if a concept figures in the ground of some mathematical fact, then a pure proof of that fact can appeal to that concept. A "pure" proof does not mean a proof using only the concepts that figure in the theorem being proved, or in the definitions of the concepts figuring in the theorem, or in the definitions of the concepts figuring in those definitions, and so on — or just the resources that one needs to master in order to understand the theorem being proved. Even if one can master the notion of a real number without having mastered the core concepts of set theory, a pure proof of a fact about real numbers can appeal to sets — if facts about real numbers are grounded in facts about sets. Indeed, the demands of purity are sometimes understood in terms of grounding, as when Bolzano is understood as holding that pure proofs of theorems in analysis (e.g., of the intermediate value theorem) cannot appeal to geometrical concepts because truths of analysis are not grounded in truths of geometry (Detlefsen and Arana 2011, 4–5; cf. Bigelow 1988, 31). A pure

9. Besides the Taylor series case, there are many other examples where impure proofs explain what pure proofs fail to explain. For example, Lange (2015a) has argued that mathematicians regard Desargues' theorem in two-dimensional Euclidean geometry as explained by a proof that not only introduces a third dimension that is exogenous to the theorem, but also puts the theorem in the context of projective geometry, which supplements Euclidean geometry with points at infinity (where parallels meet). This additional dimension and these points at infinity are extraneous to any theorem of Euclidean geometry. They are not involved in making it true. (However, the case is complicated; see Hallett 2008, 222–228.)

proof draws only on what is "intrinsic" to the theorem being proved, which automatically includes the theorem's grounds.

It might be objected that although the Taylor series theorem concerns two series composed entirely of real numbers, a proof that appeals to imaginary numbers should still be regarded as pure by virtue of giving information about the theorem's grounds — because facts about the reals and facts about the imaginaries and all of the other complex numbers have the same grounds: in facts about sets. However, even if facts about the reals and facts about the other complex numbers are alike grounded in set theory, those grounds are not given by the proof of the Taylor series theorem that appeals to the radius of convergence theorem. That proof does not proceed by expressing the two Taylor series in set-theoretic terms, nor is the radius of convergence theorem typically (or ever, as far as I know) proved in that way. Furthermore, we do not have to know that facts about the reals and facts about the other complex numbers are all grounded in set theory (or even simply that they have the same grounds) in order to appreciate the explanatory power of the proof of the Taylor series theorem that proceeds through the complex numbers. (In the next section, I will sketch a proposal regarding the source of a proof's explanatory power that accounts for our capacity to recognize this proof's explanatory power without our having to regard all facts about reals and other complex numbers as grounded in the same place.) That we can recognize the proof's explanatory power without recognizing that all complex numbers are alike grounded in set theory suggests that the proof's explanatory power does not depend on the existence of that common set-theoretic ground making the proof qualify as pure.¹⁰

That an impure proof can be explanatory (even where there is a pure proof of the same theorem) shows that at least some mathematical explanations do not work by supplying information about the grounds of the theorem being explained. Insofar as mathematical explanations

10. For a powerful argument that the reduction of arithmetic to set theory is not explanatory, see D'Alessandro (2018).

are often impure, mathematical explanations often fail to specify the grounds of the theorem being explained.¹¹

Arguably, a pure proof is roughly a proof that proceeds entirely from facts about the essences of the mathematical items figuring in the theorem being proved. For instance, the fact that $n!/k!(n-k)!$ is an integer (for natural numbers n and k where $n \geq k$) could be proved by showing that the formula gives the number of ways to choose a subset of k from a set of n . But because this proof does not proceed entirely in arithmetic terms, it is regarded by mathematicians as impure; the essences of all of the items figuring in the theorem are arithmetical. Some philosophers have held that if a fact follows from the essences of the items figuring in it, then that fact is grounded in facts about their essences. For instance, Rosen (2010, 119) suggests that the fact that every triangle has three angles is grounded in the fact that having three angles is part of what it is to be a triangle — is essential to triangularity. This view is an alternative (exclusively for those universal facts that are necessary) to the picture of the grounds of universal facts that I mentioned in the previous section, according to which universal facts are grounded in their instances.

Nevertheless, the explanatory proof of the Taylor series theorem does not proceed entirely from the essences of the items figuring in the theorem, because those essences do not involve imaginary numbers. That it is not unusual for explanatory proofs to be impure, even when

11. Although a proof's purity need not contribute to its explanatory power, purity is nevertheless a feature that mathematicians often seek and prize in proofs. Purity and explanatory power are both virtues in proofs, as are brevity, generalizability, simplicity, visualizability, theoretical fruitfulness, pedagogic value, and so forth. With so many (perhaps incommensurable) virtues, there may be no sense in which one proof is "better", all things considered, than another.

Nothing I have said has ruled out the possibility that in some cases, a proof's purity does contribute to its explanatory power — or, at least, that searching for a pure proof is a useful heuristic for finding an explanatory proof. (Detlefsen and Arana (2011, 8, fn. 31) give an example where a pure proof is sought and this search is cited as a possible means for finding an explanatory proof.) In the same way, the search for a visualizable proof may sometimes be a useful heuristic for finding an explanatory proof. Nevertheless, purity (and visualizability) would remain distinct virtues from explanatory power.

there are pure proofs of the same theorem, shows that even under the view that a fact can be grounded in essences, it often happens that a mathematical proof specifying a fact's grounds fails to explain why that fact obtains, whereas an explanation of the fact does not specify its ground.

To suggest how common it is for mathematicians to regard an impure proof as explanatory, I would like to consider briefly another theorem in Euclidean geometry: that for any two circles, there is a line such that for any point on the line, the tangent from that point to one circle is equal to the tangent from that point to the other circle. This result holds whether the two circles intersect on the Euclidean plane in two points, in one point, or nowhere. These three cases must be proved separately if the proof is to be pure. Such a proof by cases makes it appear to be a coincidence that all three cases are alike in having such a line. But in fact, it is no coincidence. What makes it no coincidence is (roughly speaking) the existence of a *common* proof of all three cases. But that proof is impure: It involves points with imaginary coordinates, which fall nowhere on the Euclidean plane. The common proof is widely regarded as making the theorem no coincidence and as explaining why the theorem holds.¹² Nevertheless, since this proof uses imaginary coordinates, it fails to identify the theorem's grounds.

A proof that proceeds by dividing the theorem into three cases is usually carried out in coordinate geometry. It proceeds by assigning

12. The "common proof" proceeds by finding the line through the two points at which the circles intersect and then showing that for any point on this line, the tangent to one circle equals the tangent to the other. If the two circles intersect at no Euclidean points, then when we solve for the two points at which they intersect, we arrive at points with imaginary coordinates. For instance, if one circle is centered at $(0,20)$ with radius 16 and the other circle is centered at $(0,-15)$ with radius 9, then their equations are $x^2 + (y-20)^2 = 16^2$ and $x^2 + (y+15)^2 = 9^2$, which have common solutions $(12i, 0)$ and $(-12i, 0)$. Although these intersection points have imaginary x -coordinates, the proof proceeds in exactly the same way as it does when the intersection points are on the Euclidean plane: The two points lie on the line $y = 0$, which is on the Euclidean plane, and for any point on this line, the two tangents are equal (Sawyer 1955, 180–181). I agree with Sawyer (1943, 232) that complex numbers reveal "the reasons for results which had previously seemed quite accidental".

coordinates to various points and then showing by calculation that the tangent from a point on the given line to one circle is equal to the tangent from a point on the given line to the other circle. It proceeds by (what mathematicians call) “brute force”. Such examples¹³ suggest that oftentimes, a proof that identifies the theorem’s ground will be a brute-force proof and that oftentimes, brute-force proofs do not explain. They simply grind out the result from the essential properties of the set up. Their purity is not enough to make them explanatory. (In the next section, I will return to the tension between a proof’s being explanatory and its proceeding by brute force.)

I have argued that some mathematical proofs are not explanatory, despite identifying the grounds of the theorem being explained, because they proceed by brute force from those grounds and thereby fail to reveal the way that the theorem’s various cases are unified. Hence, the proof’s identifying the grounds of the theorem being explained gets in the way of the proof’s being explanatory. Another, indirect way to argue that this occurs in *mathematical* explanation is to show that the same thing happens sometimes in *scientific* explanation. One way for such an example to arise in science is for several derivative laws of nature to possess the same feature, where each of these derivative laws can be derived separately from its own grounds (that is, from the relevant fundamental laws) and where the various derivative laws differ in their grounds. The derivation thus treats the similarity among the derivative laws as a coincidence; it traces the derivative laws to no common source that accounts for their common feature. However, if the similarity among the various derivative laws is in fact no coincidence, then each of those laws exhibits the common feature for the same reason. A derivation from that source fails to identify the grounds of the derivative laws, but it explains why they are similar.

For example, gravitational and electrostatic forces are similar in that both conserve energy. The fact that gravitational forces conserve energy follows from the gravitational force law (and the laws of

motion). A separate derivation from other fundamental laws shows that electrostatic forces conserve energy. These separate derivations identify the more fundamental laws in virtue of which it is true that each of these two interactions conserves energy. These more fundamental laws are the grounds of the fact that the two forces conserve energy. But these two separate derivations treat it as a coincidence that these two forces share this feature. In fact, many physicists have maintained that it is no coincidence: The two forces both conserve energy because every force has *got* to conserve energy (Lange 2016). Energy conservation arises from a fundamental spacetime symmetry meta-law that constrains the fundamental force laws (requiring them all to be invariant under arbitrary time translations) so that (within a Lagrangian framework) every fundamental force must conserve energy. Because gravitational and electrostatic forces both conserve energy for the same reason, it is no coincidence that they both do. The separate grounds of the two derivative laws fail to explain why they are alike in both conserving energy.

Scientific examples like this motivate giving a similar interpretation to various mathematical examples. Consider a result¹⁴ regarding exponentiation (the binomial theorem) and a result regarding differentiation (the general Leibniz rule):

A result regarding powers: If f and g are numbers and n is a natural number, then the binomial theorem says that

$$(f + g)^n = f^n + \binom{n}{1} f^{n-1} g + \binom{n}{2} f^{n-2} g^2 + \dots + \binom{n}{n-1} f g^{n-1} + g^n$$

$$= \sum_{k=0}^n \binom{n}{k} f^{n-k} g^k$$

where $\binom{n}{k} = n! / k!(n-k)!$

13. Coordinate geometry can likewise prove Desargues’ theorem in two-dimensional Euclidean geometry (see note 9) by “brute force”.

14. I originally discussed this result in Lange 2015b, though not in connection with grounding’s relation to mathematical explanation.

A result regarding derivatives: If $f(x)$ and $g(x)$ are n -times differentiable functions of real numbers x , and if $f^{(n)} = (d^n f)/(dx^n)$ is the n^{th} derivative of f (and the 0^{th} derivative of f is f), then a generalization of the product rule (“general Leibniz rule”) says that

$$(fg)^n(x) = g \frac{d^n f}{dx^n} + n \frac{dg}{dx} \frac{d^{n-1} f}{dx^{n-1}} + \frac{n(n-1)}{(1)(2)} \frac{d^2 g}{dx^2} \frac{d^{n-2} f}{dx^{n-2}} + \dots$$

$$= \sum_{k=0}^n \binom{n}{k} f^{(n-k)}(x) g^{(k)}(x).$$

Leibniz noticed the striking analogy between these two results as early as 1695; he even argued in a 1697 letter to Wallis that his notation was better than Newton’s because it makes this analogy more salient (Koppelman 1971, 157–158). As Koppelman (1971, 171) puts it, the similarity “certainly calls for an explanation”. Johann Bernoulli wrote to Leibniz in 1695: “Nothing is more elegant than the agreement you have observed ... doubtless there is some underlying secret” (Leibniz 2004, 398).

The conjunction of these two expansion theorems is grounded in its conjuncts. We could derive the binomial theorem from its ground, involving what exponentiation is. We could likewise derive the general Leibniz rule from its ground, including what differentiation is. But Bernoulli recognized that those derivations do not supply the underlying “secret” he wondered about. These are two separate derivations; what a derivative amounts to plays no role in the ground of the binomial theorem. So these two separate derivations treat the analogy between these two results as a coincidence. But in fact, as Bernoulli suspected, the analogy has a common origin, and so the analogy is no coincidence. Its explanation in mathematics does not work by tracing the similarity’s grounds.

As Gregory (1841, iv) put it, the similarity is not “founded on accidental analogy”. Just as there is a common reason why both gravitation and electrostatic forces conserve energy, so there is a common reason why exponentiation and differentiation are alike in the respect exhibited by the expansion theorem. The two operations are alike in *this* respect because they are alike in obeying (what Gregory called) the same three laws of combination:

	Regarding exponentiation	Regarding differentiation
	a, f, and g are numbers	$f(x,y)$ and $g(x)$ are functions
commutative law:	$fg = gf$	$\frac{\partial}{\partial x} \frac{\partial}{\partial y} f = \frac{\partial}{\partial y} \frac{\partial}{\partial x} f$
distributive law:	$a(f+g) = af + ag$	$\frac{d}{dx} (f+g) = \frac{d}{dx} f + \frac{d}{dx} g$
law of repetition:	$f^m f^n = f^{m+n}$	$\frac{d^n}{dx^n} \frac{d^m}{dx^m} f = \frac{d^{n+m}}{dx^{n+m}} f$

From the fact that exponentiation obeys these three laws, the binomial theorem follows, and in the very same way, the product rule follows from the fact that differentiation obeys these three laws. Thus, as Gregory (1841, 237) put it, both of these results “depend only on the laws of combination to which the symbols are subject, and are therefore true of all symbols, whatever their nature may be, which are subject to the same laws of combination”.¹⁵ I interpret Gregory, in referring to “dependence”, to be talking about mathematical explanation. The grounds of the two separate theorems do not combine to mathematically explain why their conjunction holds.¹⁶

15. Today we would put the point in terms of the expansion theorem holding of any commutative ring. (Presumably, Gregory is best understood as not literally meaning “symbols” here, since the same result could be expressed in many symbols, and in each case it would obtain for the same reason. Shortly I quote Boole, who does better in referring to the explanation as lying in common features of the two operations, exponentiation and differentiation, not of their symbols.)

16. Regarding the “dependence” of the expansion theorem on the three “laws of

Other commentators have made the same observations as Gregory. For example, François-Joseph Servois (1814–15, 142) said regarding the analogy, “It is necessary to find the cause, and everything is very happily explained”. That the two operations obey the same laws of combination is, he said, “la véritable origine” (151) of the analogy between the two results. Separate, unrelated derivations of the two results would prove them and display their grounds, but would not explain why their conjunction holds. As Boole (1841, 119) said of the analogy deployed to solve a linear differential equation by solving an algebraic equation and then exchanging powers for derivatives:

The analogy ... is very remarkable, and unless we employed a method of solution common to both problems, it would not be easy to see the reason for so close a resemblance in the solution of two different kinds of equations. But the process which I have here exhibited shows, that the form of the solution depends solely on ... processes which are common to the two operations under considerations, being founded only on the common laws of the combination of the symbols.

combination”, one might initially be tempted to interpret Gregory and Boole (below) as talking about grounding. However, notice that Gregory and Boole emphasize that the expansion theorem depends *only* on the two operations’ having those three “laws of combination” in common. The tacit contrast they are drawing is with the expansion theorem depending also on what those operations are – on (as Gregory puts it) the “nature” of exponentiation and the “nature” of differentiation. Now presumably, the fact that exponentiation, for instance, obeys a given law of combination is itself grounded in what exponentiation is. So if the “dependence” of the expansion theorem on the laws of combination were also a matter of grounding, then by the transitivity of grounding, the expansion theorem would also “depend” on the natures of exponentiation and differentiation, losing the tacit contrast that Gregory and Boole are drawing. (Virtually all accounts of grounding regard grounding as transitive (Correia and Schnieder 2012, 32), as long as any implicit contrasts properly line up (Schaffer 2016, 55) — that is, as long as the fact that G rather than G’ grounds the fact that H rather than H’, and the fact that H rather than H’ grounds the fact that J rather than J’.)

Once again, I take the kind of “dependence” that Boole is talking about, where the binomial and general Leibniz rules “depend” only on the laws of combination, to be a kind of explanation that is not supplied by the grounds. Exponentiation and differentiation obey the same laws of expansion because they obey the same three laws of combination.

This example illustrates how grounds often fail to explain in mathematics for something like the same reason as grounds often fail to explain in science. As I mentioned at the end of section 1, we could simply stipulate that grounds, in virtue of being grounds, give a special kind of explanation: *grounding* explanation. But as we have seen, this is not explanation as it figures in mathematical practice.

It might be objected that exponentiation’s obeying the three laws of combination grounds exponentiation’s expansion theorem and that differentiation’s obeying the three laws of combination grounds differentiation’s expansion theorem. On this view, the fact that exponentiation and differentiation are alike with respect to obeying the three laws of combination both grounds and explains their being alike with respect to obeying the expansion theorem.

However, I see no reason (beyond the fact that the laws of combination entail the expansion theorem) to say that an operation’s obeying the three laws of combination grounds its expansion theorem. I see no reason to resist the thought that just as each of the three combination laws for the operation is grounded separately in the operation’s essence, so the expansion theorem is also grounded separately in the operation’s essence — with none of these four helping to ground another. This characterization is supported by a comparison to the scientific case. A spacetime symmetry principle does not help to ground the fact that the gravitational force conserves energy. Rather, the gravitational force’s essence (specified by the fundamental gravitational force law) grounds its conserving energy, and the force law neither helps to ground nor is partly grounded by a spacetime symmetry principle. That the gravitational force exhibits a certain spacetime symmetry (namely, that its force law is invariant

under arbitrary time translations) and that it conserves energy are separately grounded in what gravitation is.

Furthermore, set the above argument aside and suppose exponentiation's three laws of combination were grounding exponentiation's expansion theorem and were grounded, in turn, by what exponentiation essentially is — and similarly for differentiation. Then it would be puzzling (if explanation in mathematics works by supplying information about how a mathematical fact arises from its grounds) why the fact that the two operations are alike with respect to obeying the expansion theorem is not mathematically explained by what exponentiation essentially is and what differentiation essentially is. But I have argued that this is no explanation in mathematical practice.

3. What Makes a Proof Explanatory?

I have now given several reasons why it is typical for proofs identifying a mathematical result's grounds not to explain why that result holds, whereas an explanatory proof of the same mathematical result fails to identify its grounds. The reasons I have given for this divergence between grounding and mathematically explaining are that proofs identifying the result's grounds can incorrectly depict the result as coincidental, must be pure, and often proceed by brute force. I have also argued that the unification that mathematical explanations frequently reveal, precisely by failing to specify the result's grounds, is similar to the unification that scientific explanations frequently achieve by failing to identify the result's grounds.

In this section, I will approach all of these matters from a different direction: by asking more generally what makes one proof but not another explanatory in mathematics. An account of this difference should account for the differences in explanatory power between the various proofs we have just looked at. A general account of mathematical explanation should also account for whatever tension there may be between a proof's explanatory power and its purity, its proceeding by brute force, or its treating the result it proves as

coincidental rather than unifying that result's various components in the way that certain scientific explanations unify. Of course, a full account of explanation in mathematics goes well beyond the scope of this paper. But a sketch of such an account would support my arguments in the previous section. In addition, the account I will now sketch suggests two further arguments that mathematical theorems are not automatically explained by their grounds. I will give those arguments in section 4.

To motivate my proposed account of mathematical explanation, consider this theorem first proved by d'Alembert:¹⁷

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

Why do the solutions of a polynomial equation with real coefficients come in these complex-conjugate pairs? It is easy to prove that they do. First, we can use straightforward calculation to show that we get the same result whether we multiply two complex numbers and then take their complex conjugate or whether we go in the opposite order by taking the numbers' complex conjugates and then multiplying them:

Show that $\bar{z}\bar{w} = \overline{zw}$:

$$\begin{aligned} \text{Let } z &= a + bi \text{ and } w = c + di. \text{ Then } \bar{z}\bar{w} = (a - bi)(c - di) = \\ &ac - bd + i(-bc - ad) = ac - bd - i(bc + ad) = \overline{ac - bd + i(bc + ad)} \\ &= \overline{(a + bi)(c + di)} = \overline{zw}. \end{aligned}$$

Hence, $\bar{z}^2 = \bar{z}\bar{z} = \overline{z^2} = \overline{z^2}$, and likewise for all other powers. Therefore, $\bar{z}^n + a_{n-1}\bar{z}^{n-1} + \dots + a_0 = \overline{z^n + a_{n-1}z^{n-1} + \dots + a_0}$. We can likewise use

17. I have given this example, along with a more extended defense of the proposal that I am about to make, in Lange (2016).

direct calculation to show that summation and complex conjugation can be taken in either order:

Show that $\bar{z} + \bar{w} = \overline{z + w}$:

Let $z = a + bi$ and $w = c + di$. Then $\bar{z} + \bar{w} = (a - bi) + (c - di) =$

$$a + c + i(-b - d) = a + c - i(b + d) = \overline{a + bi + c + di} = \overline{z + w}.$$

Thus, $\overline{z^n + a_{n-1}z^{n-1} + \dots + a_0} = \overline{z^n + a_{n-1}z^{n-1} + \dots + a_0}$, which equals $\overline{0}$ and hence 0 if z is a solution to the original equation.

Although we have just proved that the complex conjugate of one solution is another solution, this proof takes a brute-force approach. The theorem's striking feature is that the equation's nonreal solutions 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 leave us with another solution? The proof we just saw depicts this symmetry as just turning out that way when we plug and chug. A derivation of the theorem from a similar symmetry in the original problem is what would *explain why* the theorem holds. That is, an explanation of the theorem would be a proof that exploits the invariance of the set up under the replacement of i with $-i$.

The sought-after explanation¹⁸ is that $-i$ could play exactly the same roles in the axioms of complex arithmetic as i plays. Each has 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. Of course, i and $-i$ are not equal. 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 arises from the same symmetry in the axioms of complex arithmetic.

18. See Feynman et al. (1963, 22–7).

In this example, a proof is explanatory because it exploits a symmetry in the problem — a symmetry of the same kind as initially struck us in the fact being explained. I propose to generalize this diagnosis of why the two proofs of d'Alembert's theorem differ in their explanatory power. Although the symmetry between i and $-i$ is the feature of d'Alembert's theorem that initially jumped out at us, a result could have some other sort of salient feature. That feature of the result prompts a "why" question that would be answered by a proof deriving that result from a feature of the set up that is similar to the result's salient feature. On this proposal, to ask *why* the result holds is to ask for a proof that exploits a certain kind of feature in the set up: exactly the same kind of feature that stands out in the result. The distinction between proofs that explain and proofs that merely prove exists only when some feature of the result is salient. That feature's salience can privilege some proof as explanatory.¹⁹

This proposal accounts for the difference in explanatory power between the two proofs (in section 1) of the fact that 7 appears 50,000 times between 1 and 99,999. The striking thing about this fact, as we saw, is that the number of 7's is one-half of 100,000, which is just one more than 99,999. The first proof simply tallied all of the appearances of 7's, and so although it gave the theorem's grounds, it did not derive the theorem by exploiting a feature of the problem ("How many times does 7 appear...?") that is similar to the salient feature of the result. By contrast, the second proof construes the set up so that it is solved by multiplying 100,000 by 5 digits per number divided by 10 options for each digit — that is, by one-half. This proof explains, then, by virtue of tracing the one-half of 100,000 that is the result's striking feature to a similar feature of the set up.

19. Obviously, I cannot hope to defend this proposal properly here; see Lange (2016). Even if this proposal is not fully satisfactory as it stands, I hope it shows how a more general account of mathematical explanation could be used to support the previous section's argument that mathematical explanation and grounding typically diverge — and can suggest additional arguments for this conclusion (such as those given in the next section).

The proof identifying this theorem's grounds proceeds by brute force. On this proposal, no brute-force proof is explanatory when the salient feature of the theorem being explained is its symmetry (as in d'Alembert's theorem) or some other such feature that a brute-force proof fails to exploit. A brute-force approach is not selective in its focus; it simply plugs everything in and calculates everything out. By contrast, an explanation must (on this proposal) pick out particular features of the set up that are similar to the result's salient features, tracing the result's salient feature back to them.

This proposal can also account for the difference in explanatory power between the two proofs that the two Taylor series exhibit the same convergence behavior. What strikes us as remarkable about the theorem is that it identifies a feature that is common to the two Taylor series, despite the obvious differences between the two functions. In this context, the point of asking for an explanation is to ask for a proof that treats the two series alike, deriving their convergence behavior in the same way from some other feature that the two functions share. A proof that treats the two series separately, even if it is pure and reveals the theorem's grounds, fails to explain because it fails to derive the two series' convergence behavior in the same way from another feature that the two functions share. By contrast, the proof using the radius of convergence theorem treats the two functions together, deriving their common convergence behavior from another feature that they share — namely, that they go undefined somewhere on the unit circle centered at the origin of the complex plane. This proof's impurity allows it to explain why the theorem holds.²⁰

20. The case of Desargues' theorem (see notes 9 and 13) is similar (following Lange 2015a). A brute-force proof using homogeneous coordinates just plugs everything in and turns the meat grinder. By contrast, a striking feature of Desargues' theorem is that it identifies something common to all three of the intersections of the corresponding sides of the two triangles in perspective — namely, that each of those intersections falls on the same line. The proof of Desargues's theorem that exits to the third dimension allows the two coplanar triangles in perspective to be the projections of triangles on different planes. The proof derives the similarity among the three pairs of corresponding sides from another such similarity: that in each case, the two

The result's salient feature is that it identifies a property shared by the two series (namely, their convergence behavior). The proof appealing to the radius of convergence theorem is explanatory because in appealing to each series' possessing a singularity on the complex plane's unit circle, this proof exploits a property that is shared by the two series and is thus similar to the result's salient feature.

Likewise, the striking thing about the theorem concerning exponentiation and differentiation, as Leibniz and Bernoulli remarked, is that it identifies a respect in which the two operations are similar. Therefore, proofs that treat the two operations separately fail to explain why exponentiation and differentiation are alike in this respect. By contrast, the theorem is explained when the similarity it displays is derived uniformly from another similarity between exponentiation and differentiation — namely, that they obey the same three "rules of combination".

In this way, the proposal I have sketched concerning the source of a proof's explanatory power accounts for the way that a proof's having explanatory power may fail to require its purity, may be precluded by its taking a brute-force approach, and may be undercut by its treating the result it proves as coincidental (instead of unifying that result's various components in the same manner as certain scientific explanations do). This proposal thus accounts for the divergence that we have seen between grounding and mathematically explaining.

On this account, a proof deriving the theorem from its grounds may explain, but its specifying the theorem's grounds would then be incidental to its explanatory power. It would derive its explanatory power not by virtue of tracing the theorem to its grounds, but rather by virtue of tracing the theorem to a feature of the set up that is like the theorem's salient feature.

corresponding sides are projections of lines on the *same* two planes — which must intersect in a line. Although the proof exiting to the third dimension is impure, it explains by showing how the theorem's salient feature arises from a similar salient feature in the set up.

4. Two Further Arguments that Grounds do not Explain.

The account of mathematical explanation I have just sketched suggests two further arguments for the view that mathematical theorems are not automatically explained by their grounds.

According to the account I have just sketched, if a mathematical result exhibits no striking feature at all, then there is nothing that its explanation over and above its proof would amount to. In other words, there is nothing that it would mean to ask why the result holds, over and above asking for a proof of the result. Likewise, the account entails that a mathematical result exhibiting some striking feature may have no proof exploiting a similar feature of the setup, in which case the result has no mathematical explanation.

These consequences of the proposed account seem correct. Mathematicians who are trying to explain why a given theorem holds, prompted by the theorem's exhibiting some striking feature, routinely acknowledge that the theorem may in fact have no proof that explains why it holds. For example, after asking why the two Taylor series I discussed earlier have the same convergence behavior, despite the stark differences between the two functions, 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!" (Of course, he goes on to point out, "In this case there does happen to be an explanation ..." — namely, involving the impure proof appealing to the radius of convergence theorem.)²¹

However, if a mathematical theorem were automatically explained by its grounds, then every theorem, no matter how pedestrian, would have an explanation. This seems contrary to mathematical practice. It would make Spivak incorrect in acknowledging that a given theorem having a salient feature might turn out to have no explanation.

I will conclude this paper by describing how the previous section's account of explanatory proofs supplies yet another argument for the

view that a theorem's grounds do not generally explain why it obtains. This argument emphasizes the role that the proposed account assigns to the *salience* of some feature possessed by the given theorem.

On this account, whether a given proof qualifies as explanatory (indeed, whether there is any distinction between explanatory and non-explanatory proofs of some theorem) depends on some feature(s) of the theorem being salient. For example, when we encountered d'Alembert's theorem (in section 3), its salient feature was the symmetry between i and $-i$ that it identifies in the solutions to polynomial equations. The salient feature of the theorem concerning exponentiation and differentiation (in section 2) was its revealing a respect in which the two operations are alike. The salient feature of the theorem (in section 1) that 7 appears 50,000 times in a list of the numbers from 1 to 99,999 is that 50,000 is half of 100,000, which is nearly 99,999.

Which feature of a theorem is salient (and whether it has any salient feature at all) depends on the context. Indeed, as the conversational context shifts, a feature that had been salient can retreat into the background as a new feature becomes salient. My account predicts that when that happens, there may be a corresponding shift in which proof (if any) qualifies as explanatory.

This prediction is borne out in mathematical practice. Let's look at an example.²²

Ask your friends to select any two numbers and to insert one in the first row and the other in the second row of the following table:

21. Gale (1991, 41) also acknowledges that a theorem, although proved, need not have an *explanatory* proof.

22. I discussed this example in Lange 2016 to make a different point.

1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	
total	

Then ask your friends to complete the table by inserting the sum of those two rows in row 3, the sum of rows 2 and 3 in row 4, and so forth through row 10 — and, finally, to compute the grand total by summing all of the numbers in rows 1 through 10. While your friends are furiously completing the table, you look over their shoulders until they fill in row 7. Then you simply take that entry, multiply it by 11 in your head, and boldly announce the grand total long before they reach it. Here, for example, is the completed table when the two initial numbers are 4 and 7.

1.	4
2.	7
3.	11
4.	18
5.	29
6.	47
7.	76
8.	123
9.	199
10.	322
total	836

To see if your success was just a fluke, your friends may ask you to perform your feat beginning with different numbers (and even with fractions or negative numbers). Eventually, you will divulge your secret: that for any initial two numbers, the grand total equals 11 times the entry in row 7. Your audience will presumably then want to know *why* this theorem holds. In this sort of context, the theorem's salient feature is obviously that it allows your trick to work for any two initial numbers. In asking why the trick works, your audience wants to know whether this similarity among the theorem's cases is a coincidence or not. In other words, your audience is interested in a proof that deduces this feature that is common to all of the cases from some other feature that is common to them all. That is the point in asking *why* the theorem holds.

In asking why the theorem holds, the members of your audience are demanding the kind of proof that books of mathematical "magic" standardly present as explaining why the trick works. The following explanation comes from Gardner's *Mathematical Circus* (1979, 101–104, 167–168). The trick's explanation is that in any table, the two initial numbers x and y generate a "generalized Fibonacci sequence" on the subsequent rows: $x, y, x + y, x + 2y, 2x + 3y, \dots$. Gardner displays this sequence in the following table:

1.	x
2.	y
3.	$x + y$
4.	$x + 2y$
5.	$2x + 3y$
6.	$3x + 5y$
7.	$5x + 8y$
8.	$8x + 13y$
9.	$13x + 21y$
10.	$21x + 34y$
total	$55x + 88y$

As Gardner's table proves, the sum of the first 10 members in any such sequence is $55x + 88y$, which is 11 times the 7th member of the sequence ($5x + 8y$). This "common proof" covering all cases alike makes it no mathematical coincidence that the trick works in any possible case. The proof unifies all the cases falling under the theorem.

Now for the shift in context. Looking at Gardner's table, we think of the x 's as forming one sequence and the y 's as forming another, separate sequence. We recognize that the coefficients of the x terms in lines 3 through 10 are the first eight members of the Fibonacci sequence and the coefficients of the y terms in lines 2 through 10 are the first nine members. So Gardner's explanation reveals that the result's holding on the x -side consists of the sum of the first 8 Fibonacci numbers plus 1 equaling 11 times the fifth Fibonacci number (line 7's x -coefficient). The result's holding on the y -side consists of the sum of the first nine Fibonacci numbers equaling 11 times the sixth Fibonacci number (line 7's y -coefficient). In other words (where F_i is the i^{th} Fibonacci number):

$$F_1 + F_2 + \dots + F_8 + 1 = 11 F_5$$

$$F_1 + F_2 + \dots + F_9 = 11 F_6$$

In the wake of Gardner's explanation, the context has shifted; a different feature of the theorem has become salient. Having proved the theorem by using the table filled in with x 's and y 's so as to give a common proof for every instance of the theorem, we now find ourselves considering the original result as having two components: an x -result and a y -result. The salient feature is now that the x -sum works out so that the coefficient in the grand total is 11 times the coefficient on the seventh line and that the y -sum works out in the very same way. The result's salient feature is now that the x - and y -sums are alike in this respect. This feature could not *originally* have been the result's salient feature, since we could not originally have decomposed the theorem into x - and y -results; we could do that only after having Gardner's explanation.

In this new context, a proof explaining why the theorem holds has to treat the result's x - and y -components alike, deducing each in the same way from a feature they share. Such a proof must show that the above two equations concerning the Fibonacci sequence are no coincidence — have a common proof. Gardner's original explanation cannot do that, since it treats the x -sum separately from the y -sum. So it does not explain in the new context.

This shift is visible in the mathematics literature. For instance, one textbook (Benjamin and Quinn 2003, 30–31) begins by presenting the trick under the title "A Magic Trick", indicating the kind of proof that would be explanatory (namely, one revealing it to be no coincidence that the trick succeeds for any x and y). The book then gives Gardner's proof, commenting that "[t]he explanation of this trick involves nothing more than high school algebra". However, once the table decomposes the result into the x - and y -components, a previously unrecognized feature of the result becomes salient. The textbook records that the given explanation cannot help raising "why" questions it cannot answer, remarking: "... the total of Rows 1 through 10 will sum to $55x + 88y$. As luck would have it, (actually by the next identity), the number in Row 7 is $5x + 8y$ " (Benjamin and Quinn 2003, 30–31). The result's x - and y -components are the above two equations regarding the first eight and first nine Fibonacci numbers, and the proof depicts these equations as if their jointly holding were a matter of "luck," having no common proof: a mathematical coincidence. But it is no coincidence — as the textbook foreshadowed by saying "actually, by the next identity".

That "next identity" concerns Fibonacci numbers. To explain why the two equations hold, think of the Fibonacci sequence as *doubly* infinite. With $F_1 = 1$ and $F_2 = 1$, it follows that $F_0 = 0$, $F_{-1} = 1$, etc. Note that F_{-1} and F_0 are the x -coefficients in lines 1 and 2 of Gardner's table. So the following equations capture the x - and y -sides, respectively:

$$F_{-1} + F_0 + \dots + F_8 = 11 F_5$$

$$F_0 + F_1 + \dots + F_9 = 11 F_6$$

This pair is no coincidence. The two equations can be given the same proof — a proof that for any 10 consecutive members of the doubly-infinite Fibonacci sequence, their sum equals 11 times the seventh member in the sum. (I relegate the boring proof to an endnote.²³ It gives the same treatment to every 10-term segment of the doubly-infinite Fibonacci sequence.)

In virtue of this common proof, it is no coincidence that the x - and y -sides in Gardner's table both make the trick work. The trick's x - and y -components both work because of another feature they share: Each involves the sum of ten consecutive members of the doubly-infinite Fibonacci sequence. So (on section 3's account) this proof explains why the trick works, when this "why" question is asked in a context where the trick's salient feature is that the x - and y -sides have something in common. The proof explains by tracing that similarity to another feature that the two sides share.

This example illustrates how a shift in context can alter what it takes for a proof to be explanatory. A proof's explanatory power depends on the salience of particular features of the theorem being explained, and their salience is context-dependent. By contrast, a theorem's ground is not context dependent.²⁴ A theorem's truth-makers remain its truth-

23. For any integer n ,

$$\begin{aligned} & F_{n+1} + F_{n+2} + F_{n+3} + F_{n+4} + F_{n+5} + F_{n+6} + F_{n+7} + F_{n+8} + F_{n+9} + F_{n+10} \\ &= 2F_{n+3} + 2F_{n+4} + 2F_{n+5} + 0F_{n+6} + F_{n+7} + 2F_{n+8} + 2F_{n+9} \\ &= 4F_{n+5} + 3F_{n+7} + 4F_{n+8} \\ &= 4F_{n+5} + 4F_{n+6} + 7F_{n+7} \\ &= 11F_{n+7} \end{aligned}$$

24. As I mentioned in note 16, some grounding theorists (e.g., Schaffer 2016) regard grounding as contrastive: the fact that G rather than G' grounds the fact that H rather than H' . Such contrastive grounding facts are not context-dependent, but the relevant contrast to H may differ in different contexts and remain implicit. So (as a referee has suggested) the fact that H grounds G may be context-dependent: true for one implicit, context-dependent contrast and false for another.

However, although the explanandum in a mathematical explanation may involve a (tacit) contrast, the context-dependence of this contrast is insufficient to account for the context-dependence of mathematical explanation: The (context-dependent) salience of some feature of the explanandum is

makers however the context may shift.²⁵ This suggests that grounding and explaining in mathematics are fundamentally distinct.

References

- Benjamin, Arthur T., and Jennifer J. Quinn 2003. *Proofs that Really Count*. Washington DC: Mathematical Association of America.
- Bigelow, John 1988. *The Reality of Numbers*. Oxford: Clarendon.
- Boole, George 1841. On the integration of linear differential equations with constant coefficients. *Cambridge Mathematical Journal* 2: 114–119.
- Correia, Fabrice and Benjamin Schnieder 2012. Grounding: An opinionated introduction. In Correia and Schnieder (eds.), *Metaphysical Grounding*. Cambridge: Cambridge University Press, pp. 1–36.
- D'Alessandro, William 2018. Arithmetic, set theory, reduction and explanation. *Synthese* 195 (11): 5059–5089.

something over and above any (context-dependent) contrast to the explanandum. Although in some cases, a contrast renders a feature of the explanandum salient, we can introduce a contrast even to some pedestrian mathematical theorem without rendering any of its features salient. There then remains nothing that it would be to explain why that theorem holds. Furthermore, even under a fixed contrast to the explanandum, a shift in context can shift which of its features is salient. For instance, in asking why Gardner's mathematical trick works, we presumably mean to ask why it works for *all* initial numbers x and y rather than for only *some* (or for *none*). This contrast is not enough to pick out whether we seek a proof that treats every x, y pair alike or a proof that treats the x - and y -sums alike. Even a proof that does neither would rule out the contrasting alternative to the trick's always working, despite failing to explain (in any of these contexts) why the trick works.

25. The ground of the fact that the x - and y -sides of Gardner's table both work out may be that for any 10 consecutive members of the doubly-infinite Fibonacci sequence, their sum equals 11 times the seventh member in the sum. In that case, perhaps one of the explanations of the mathematic trick supplies the grounds of the trick's success. (An account of mathematical facts' grounds would have to say whether the regularity regarding every 10-member Fibonacci sequence is grounded in facts about individual 10-member sequences.) Nevertheless, that proof does not acquire its explanatory power from supplying the grounds of the trick's success — since in a different context, the ground remains but the proof lacks explanatory power.

- Detlefsen, Michael 2008. Purity as an ideal of proof. In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 179–197.
- Detlefsen, Michael, and Andrew Arana 2011. Purity of methods. *Philosophers' Imprint* 11 (2): 1–20.
- Dreyfus, Tommy, and Theodore Eisenberg 1986. On the aesthetics of mathematical thought. *For the Learning of Mathematics* 6(1): 2–10.
- Feynman, Richard, Robert Leighton, and Matthew Sands 1963. *The Feynman Lectures on Physics*, vol. 1. Reading, MA: Addison-Wesley.
- Fine, Kit 2001. The question of realism. *Philosophers' Imprint* 1 (2): 1–30.
- Fine, Kit 2012. Guide to ground. In Correia and Schnieder (eds.), *Metaphysical Grounding*. Cambridge: Cambridge University Press, pp. 37–80.
- Gale, David 1991. Mathematical entertainments. *Mathematical Intelligencer* 13 (1): 40–42.
- Gardner, Martin 1979. *Mathematical Circus*. New York: Knopf.
- Gregory, Duncan 1841. *Examples of the Processes of the Differential and Integral Calculus*. London: Deighton.
- Hallett, Michael 2008. Reflections on the purity of method in Hilbert's *Grundlagen der Geometrie*. In Paolo Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 198–255.
- Koppelman, Elaine 1971. The calculus of operations and the rise of abstract algebra. *Archive for History of Exact Sciences* 8 (3): 155–242.
- Lange, Marc 2010. What are mathematical coincidences (and why does it matter)? *Mind* 119 (474): 307–340.
- Lange, Marc 2015a. Explanation, existence, and natural properties in mathematics: a case study – Desargues' theorem. *Dialectica* 69 (4): 435–472.
- Lange, Marc 2015b. Depth and explanation in mathematics. *Philosophia Mathematica* 23 (2): 196–214.
- Lange, Marc 2016. *Because Without Cause*. New York: Oxford University Press.
- Leibniz, Gottfried Wilhelm 2004. *Sämtliche Schriften und Briefe*. Deutsche Akademie der Wissenschaften, ed. Reihe III, Band VI Part A. Darmstadt, Leipzig, and Berlin: Akademie-Verlag.
- Mancosu, Paolo 2008. Mathematical explanation: why it matters. In Mancosu (ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 134–150.
- Pincock, Christopher 2015. The unsolvability of the quintic: A case study in abstract mathematical explanation. *Philosophers' Imprint* 15 (3): 1–19.
- Rosen, Gideon 2010. Metaphysical dependence: Grounding and reduction. In Bob Hale and Aviv Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press, pp. 109–136.
- Sawyer, W.W. 1943. *Mathematician's Delight*. Harmondsworth: Penguin.
- Sawyer, W.W. 1955. *Prelude to Mathematics*. Harmondsworth: Penguin.
- Schaffer, Jonathan 2016. Grounding in the image of causation. *Philosophical Studies* 173 (1): 49–100.
- Servois, François-Joseph 1814–15. Réflexions sur les divers systèmes d'exposition des principes du calcul différentiel, et en particulier, sur la doctrine des infiniment petits. *Annales des Mathématiques Pures et Appliquées* 5: 141–170.
- Spivak, Michael 1980. *Calculus*, 2nd ed. Berkeley: Publish or Perish.
- Steiner, Mark 1978. Mathematics, explanation, and scientific knowledge. *Nous* 12: 17–28.
- Waismann, Friedrich 1982. *Lectures on the Philosophy of Mathematics*. Amsterdam: Rodopi.
- Wilson, Mark 2006. *Wandering Significance*. Oxford: Clarendon.