1 Introduction

Some scientific explanations involve mathematics. Within mathematics, some proofs are said to explain. Do these practices tell us anything about the nature of explanation or mathematics? In this Element this daunting topic is divided into four parts. First, can any traditional theory of scientific explanation make sense of the place of mathematics in explanation? Each traditional theory that is discussed is a *monist* theory because it supposes that what makes something a legitimate explanation is always the same (Section 2). Second, if traditional monist theories are inadequate, is there some way to develop a more flexible but still monist approach that will clarify how mathematics can help to explain (Section 3)? After a consideration of the limitations of some recent flexible monist accounts, the options for a pluralist approach are examined. What sort of pluralism about explanation is best equipped to clarify how mathematics can help to explain in science and in mathematics itself? While a pluralist can allow that different sorts of explanations work differently, it still remains important to clarify the value of explanations (Section 4). Finally, how can the mathematical elements of an explanation be integrated into the physical world? Some of the evidence for a novel scientific posit may be traced to the explanatory power that this posit would afford, were it to exist. Can a similar kind of explanatory evidence be provided for the existence of mathematical objects, and if not, why not? (Section 5).

In his 2001 paper "Mathematical Explanation: Problems and Prospects" Paolo Mancosu argues that "mathematical explanations can be used to test theories of scientific explanation and that an account of mathematical explanation might have important consequences for the philosophy of science" (Mancosu 2001, p. 102). This Element builds on this point by considering how various approaches to scientific explanation can make sense of both (i) explanatory proofs in pure mathematics and (ii) scientific explanations that turn essentially on mathematical resources. In Sections 2–4 I argue that the best option for clarifying how these explanations work is pluralism about explanation. This means that different explanations employ different explanatory relevance relations when they indicate why some target is the way that it is. In Section 5 I consider the significance of mathematical explanation for the interpretation of pure mathematics. I argue that the existence of genuine mathematical explanations does not support the existence of mathematical objects through the use of inference to the best explanation.

My own interest in mathematical explanation can be traced directly to the pioneering work of Paolo Mancosu (see especially Mancosu 2000, 2001, 2008, 2018).[1] I was lucky enough to have Mancosu as the advisor for my 2002

[1] A new version of (Mancosu 2018) is currently in preparation.

dissertation on questions related to the applicability of mathematics. This Element attempts to follow Mancosu's call to attend carefully to mathematical and scientific practice in philosophical work. I believe that the pluralism about explanation that I argue for is consistent with Mancosu's views, but he may not agree with the account of what all explanations have in common that I offer here (Section 4.3).

A draft of this Element benefitted enormously from comments by Sam Baron, Andre Curtis-Trudel, Marc Lange, and Paolo Mancosu. I have unfortunately not been able to address all of their helpful suggestions here, and they are, of course, not responsible for any remaining errors or oversimplifications of the issues discussed. I am also grateful to two anonymous referees for their insightful reactions to the penultimate version of this Element. I hope this Element will help to introduce new readers to the wonders of mathematical explanation, and also to inspire new work on this complex topic.

2 The Challenge to Traditional Theories of Scientific Explanation

This section starts by introducing five principles that are used to test competing accounts of explanation, and illustrates how these tests work by developing some standard objections to accounts that emphasize derivation and unification (Section 2.1). This section also considers three causal accounts of explanation and argues that they are unable to make sense of the contrast between explanations that merely employ mathematics to represent something else and explanations whose explanatory power is tied more directly to the mathematics employed (Section 2.2).

2.1 Derivation and Unification

Philosophical investigations of a topic like explanation typically take for granted some principles about that topic that make it possible to test competing accounts. This Element takes for granted five principles. It supposes that an account of explanation aims to cover all explanations, including those found in science and mathematics. As I will argue, many accounts fail to respect the principles articulated in this section. We can thus use these principles to identify the problems with various accounts of explanation that have been proposed, especially when one considers how mathematics figures into explanations. Of course, one may avoid these problems by rejecting one or more of the principles that are assumed here.

All participants in these debates agree that a scientific explanation provides a reason why something is the case. The target of an explanation may be something particular, like a specific event or state. The target may also be something general, like a recurring pattern or phenomenon. A legitimate explanation of a target indicates why that target is the way it is. This motivates our first principle for accounts of explanation:

1. There is an important distinction between a description of some target of explanation and an explanation of that target.

This principle does not definitively refute any account, as a defender of any account is liable to interpret "an important distinction" in their own self-serving way. However, I will appeal to this principle to help to clarify my reasons for questioning this or that proposal.[2]

A closely related principle involves the distinction between the evidence that some phenomenon has some character and an explanation of that aspect of the phenomenon. For example, careful paleontological investigations may determine that the rate of the Earth's rotation on its axis is decreasing. But additional accounts of the gravitational interactions between the Earth and the Moon are needed to explain this change. Our second principle is thus:

2. There is an important distinction between the evidence for some fact and an explanation of that fact.

The third principle that I will deploy assumes that there is an order to explanation. If one says that B is the case because of A, then A provides a reason for B being the way that it is. In some respect, then, A must be more basic or fundamental than B. In causal explanation, A is partly responsible for the occurrence of B, and so in this sense is also more basic. If this is right, then it would be illegitimate to appeal to B when explaining A. One way to summarize this point is to say that an explanatory relevance relation is asymmetrical: when A stands in that relation to B, then B does not stand in that very relation to A. However, the issue is complicated by the fact that an explanation may have parts. To allow for explanations with parts, our third principle has the following formulation:

3. (Priority) If A is part of an explanation of B, then B is not part of an explanation of A.

For example, the mass of the Moon is part of the explanation for why the rate of rotation of the Earth on its axis is decreasing over time. Our third principle thus requires that the decreasing rate of rotation of the Earth on its axis is not part of an explanation of the mass of the Moon.

[2] For a recent discussion of this issue, see Taylor (forthcoming).

Our fourth and fifth principles help to identify the subject matter of this Element. This is the special character of some explanations that involve mathematics. As we will see later in this section, many philosophers maintain that some explanations that involve mathematics use the mathematics in a special way that renders the explanation genuinely or distinctively mathematical. Following Baker and Baron, I will call such explanations "genuine mathematical explanations."[3] Of course, not everyone agrees that there are genuine mathematical explanations. But our fourth principle takes for granted that genuine mathematical explanations exist and requires that an account clarify their character:

4. There is a special way that mathematics may appear in a scientific explanation that makes it a genuine mathematical explanation.

Our fifth (and final) principle relates to pure mathematics. One goal of mathematical activity is to obtain a proof of a theorem. Mathematicians sometimes praise or criticize a proof based on its explanatory power. In certain contexts, it is thought valuable to explain why a theorem is the case even after it has been given a proof that is otherwise adequate. Our fifth principle assumes that this feature of mathematical practice is legitimate:

5. Some proofs of a theorem explain why that theorem is the case, while other proofs do not explain why that theorem is the case.[4]

Combining our fourth and fifth principles will turn out to be a powerful tool to criticize some proposed accounts of explanation. Many proposals will fail the fourth or fifth test because they do not allow for genuine mathematical explanations or they rule out explanatory proofs. As with the other principles, this does not provide a definitive refutation of these proposals, but it does clarify their limitations and also why some may reject those proposals.

Much of our discussion will turn on cases where mathematics appears in an explanation. Our first case is an answer to the question, "Why is the shadow cast by the Ohio Stadium flagpole 49 m in length at 3 p.m.?" An explanation of this state may appeal to the position of the sun at the time and to the height of the flagpole. But this information does not seem sufficient to explain the length of the shadow, as there is a deductive gap between the statements that provide this information and the statement characterizing the target of the explanation:

[3] See especially Baker (2005), Lange (2013), and Baron (2019).
[4] Contrary to the claims of D'Alessandro (2020), nobody assumes that all explanations in pure mathematics are proofs. See especially Mancosu (2001) and Lange (2018b). I restrict my focus here to proofs to make the discussion tractable.

1. At 3 p.m., the light rays from the sun hit the top of the flagpole at an angle of 45°.
2. The height of the flagpole is 49 m.

   Therefore, the length of the shadow is 49 m.

To close this deductive gap, we need to add a statement from geometric optics that involves trigonometry:

3. The length x of the shadow cast by any pole of height y m when the light hits at an angle of 45° satisfies the following equation: $\tan 45° = x \text{ m}/y \text{ m}$.

As $\tan 45° = 1$ and $y = 49$, it follows that $x = 49$ (Figure 1). So for this type of case, at least, the role for the mathematics in the explanation is to permit the deduction of a statement characterizing the explanatory target.

Although the idea has a long history, Hempel is the philosopher who did the most to argue that a necessary condition on an important kind of explanation is that the explanation provide a deduction of a statement describing the explanatory target. Hempel called such explanations "deductive-nomological" (D-N) explanations. The term "nomological" indicates an additional necessary condition on such deductions: they must deduce their target statement through the essential use of a scientific law. Our statement 3 would be the law for this D-N explanation. This is Hempel's way of distinguishing an explanation from a description.

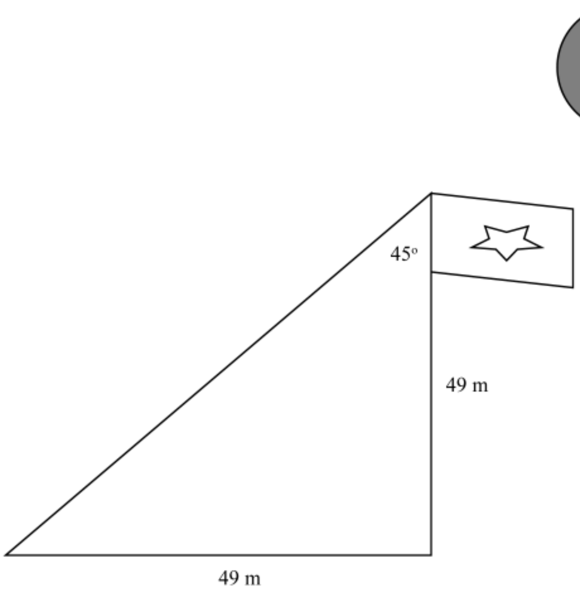

Figure 1 The flagpole and the shadow.

Hempel offered different motivations for the need for laws in explanations. In his famous *Aspects of Scientific Explanation*, for example, Hempel (1965) says "[(i)] the argument shows that, given the particular circumstances and the laws in question, the occurrence of the phenomenon *was to be expected*; and [(ii)] it is in this sense that the explanation enables us to *understand why* the phenomenon occurred." But in addition "[(iii)] it is in virtue of such laws that the particular facts cited in the explanans possess explanatory relevance to the explanandum phenomenon" (p. 337).[5] The relationship between (i), (ii), and (iii) is far from clear. One interpretation of Hempel is that what makes something explanatorily relevant is that this fact could have been used to lawfully predict that state in advance. It is this that constitutes our understanding of that state.

The most influential objection to Hempel's D-N account takes for granted that some laws permit deductions with a troubling sort of symmetry. Our flagpole case was in fact introduced into these debates to illustrate one such troubling case.[6] For in addition to the presumably acceptable explanation just given, the following deductive argument also seems to meet all of Hempel's necessary conditions on D-N explanations:

1. At 3 p.m., the light rays from the sun hit the top of the flagpole at an angle of 45°.
2'. The length of the shadow is 49 m.
3'. The height y of any pole that casts a shadow of length x m when the light hits at an angle of 45° satisfies the equation: $\tan 45° = x \text{ m}/y \text{ m}$.
Therefore, the height of the flagpole is 49 m.

If (3) is a law, then (3') is also a law. As both arguments are deductively valid, Hempel seems to lack any principled reason to exclude this explanation. But if both explanations are granted, then we violate our third principle concerning the order of explanatory priority. This principle says that if A is part of an explanation for B, then B cannot be part of an explanation for A. But here we have the height of the flagpole being part of an explanation of the length of the shadow, and also the length of the shadow being part of an explanation of the height of the flagpole. Thus we face a choice between agreeing with Hempel and maintaining an order of explanatory priority. Applying our third principle requires rejecting Hempel's approach to explanation.

Most philosophers of science have accepted explanatory priority and thus rejected Hempel's D-N account as inadequate. By far the most popular

[5] The explanandum is the target of the explanation or what is being explained. The explanans is the explanation itself or what provides the explanation.
[6] See Salmon (1989) for extensive discussion of this and other objections to Hempel.

approach adds a causal condition on explanation, which we consider in Section 2.2. However, Kitcher offered a different diagnosis of the failings of Hempel's approach. He argued that Hempel failed because he tried to assess explanations individually. The alternative approach that Kitcher pursued is to evaluate explanations globally based on how well they help to unify or systematize a collection of accepted statements: "Science advances our understanding of nature by showing us how to derive descriptions of many phenomena, using the same derivation again and again, and, in demonstrating this, it teaches us how to reduce the number of types of facts we have to accept as ultimate (or brute)" (Kitcher 1989, p. 432). At any given time in the history of science, there will be some set of accepted statements K. The "explanatory store" over K will specify a set of argument patterns that permit some members of K to be derived from other members of K. An explanation (with respect to this K) will then be an instance of such an argument pattern. Kitcher adds that an explanation is legitimate when it appears "in the explanatory store in the limit of the rational development of scientific practice" (Kitcher 1989, p. 498).

To appreciate the explanatory role for mathematics that Kitcher's approach creates, it will be useful to introduce a case from pure mathematics.[7] As assumed in our fifth principle, only some proofs in mathematics are judged to explain the theorem proven. In Euclid's *Elements* the solution to a problem involves constructing a geometric figure using the limited resources licensed by his postulates, for example to connect any two points by a line or to draw a circle around a point with the radius of some line. One such problem is to bisect an angle BAC (Book I, proposition 9; see Figure 2).

The first step to solve this problem is to pick some point on line BA. Call this point D. A circle centered on A and of radius AD can then be drawn to cross line CA at a new point E. Radius AD is equal in length to AE. An earlier construction in Euclid shows how to construct an equilateral triangle on any given line. Construct such a triangle on line DE, with a third corner F. Finally, connect F to A. The triangles ADF and AEF are congruent, as they have three sides of the same length (AD = AE, DF = EF, AF = AF). As congruent triangles have corresponding angles of the same size, angle DAF = angle EAF, and the bisection is completed. Let us suppose that this construction not only proves that every angle can be bisected, but also that it explains why every angle can be bisected.

A construction procedure like the bisection of an angle can be iterated, and so it is clearly possible to divide an angle into n equal parts when n satisfies the equation $n = 2^m$, for some m ($n = 2, 4, 8, 16, \ldots$). Discussing a case like this,

[7] For Kitcher's own examples from pure mathematics, see Kitcher (1989, p. 424).

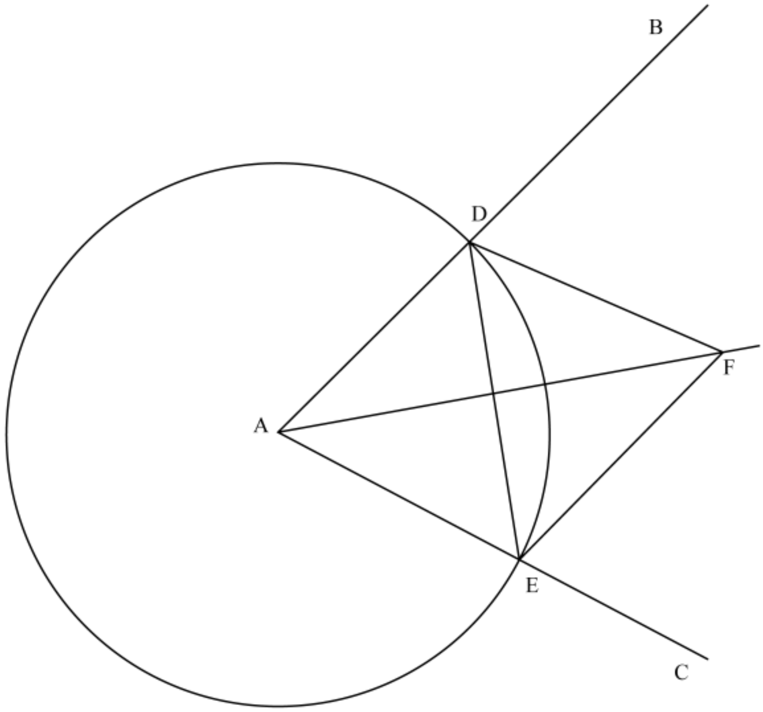

Figure 2 Bisecting an angle.

Kitcher notes that "[e]ven when we are interested in explaining a particular event or state, the explanation we desire may well be one that would also explain something quite general, and any attention to the local details may be misguided and explanatorily inadequate" (Kitcher 1989, p. 427). In this case, one may ask why an angle can be divided into 64 equal parts. While one proposed explanation would lay out all the steps of the construction, Kitcher's view is that a better explanation would connect the 64-part case to all the cases that are amenable to a unified construction procedure. The goal of unification is often prominent in pure mathematics as well as in many scientific cases. In Kitcher's terms, this would lead the explanatory store over a K that includes Euclidean geometry to contain a single argument pattern that covers all divisions of angles into n parts, where n is a power of 2. The instances of this pattern would then count as explaining their respective theorems.

More generally, the explanatory store over K is arrived at by identifying the smallest number of stringent argument patterns that permit the most members of K to be conclusions of some instance of some such pattern.[8] It is not clear if

[8] A stringent argument pattern places substantial conditions on how its instances can be generated. These conditions are needed to avoid making unifications too easy to achieve.

Kitcher's procedure for picking out the explanatory store over K is well defined, and there are difficult issues associated with how this procedure is meant to accommodate scientific change. Kitcher does offer an intriguing place for mathematics in explanation, though. For it does seem that mathematical theorems are well suited to unify large numbers of mathematical and nonmathematical claims. This is clear even from our two cases so far. For both the flagpole case and the bisection case, the generalizability of the mathematical result shows how many similar cases can be treated in a uniform fashion. All the members of K that involve a shadow being of a certain length can be handled using a single argument pattern whose crucial premise generalizes (3): The length x meters of the shadow cast by any object of height y meters when the light hits at an angle of z degrees satisfies the equation tan z degrees = x meters / y meters. Similarly, for any number of parts $n = 2^m$, the complete instructions for how to bisect any angle into that many parts can be given using $(n - 1)$ iterations of the angle bisection construction.[9]

Kitcher also proposed an ingenious way to preserve explanatory priority using his unificationist approach to explanation (Kitcher 1989, p. 484). His general strategy was to argue that any purported explanatory store E(K) over K that allowed for troubling symmetrical pairs of derivations would have redundant argument patterns. So, E(K) could be replaced by a different explanatory store E'(K) that would provide a better unification of K by disqualifying one of the proposed explanations. In the flagpole case, Kitcher allows for an argument pattern that derives the height of flagpoles from lengths of shadows using our generalization of (3). But he insists that there will be another argument pattern that derives the dimensions of ordinary objects like flagpoles and towers in terms of their origin and development or, as we might put it, their constitution. The height of the flagpole can be derived by summarizing how it was constructed, so that its parts combine to yield an object of this height. If E(K) has this argument pattern and also has an argument pattern H that derives the height of the flagpole based on the length of the shadow it casts, then argument pattern H only derives statements that can also be derived in some other way. Thus, an explanatory store E'(K) that drops pattern H and retains the constitution pattern would mark an improvement.

The powerful role of mathematics in unifying derivations turns out to be a problem for Kitcher's approach. We can see this by recalling our first two principles for an account of explanation: a description is not an explanation, and evidence for a target is not an explanation of that target. Kitcher is preoccupied

[9] As each bisection increases the number of parts by 1, dividing an angle into $n = 2^m$ parts requires $n - 1$ bisections.

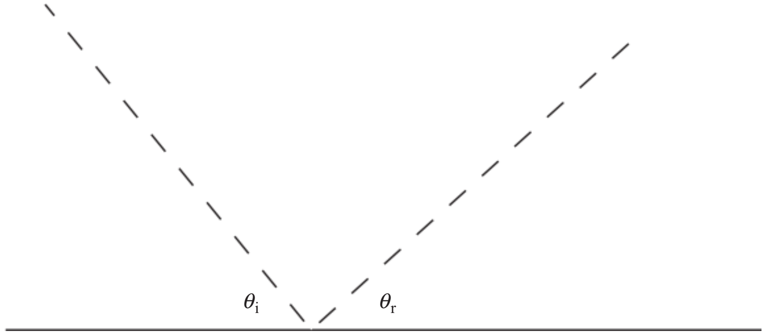

Figure 3 Law of reflection.

with the number of stringent argument patterns needed to derive a given set of claims. Here is a case from geometric optics that Kitcher's analysis gets wrong.[10] The law of reflection says that when a ray of light hits a reflecting surface like a mirror, the angle of incidence θ_i equals the angle of reflection θ_r (Figure 3).

Another law of optics is Snell's law, which concerns refraction: when a light ray goes from one medium (like air) to another (like water), the direction of the ray will change or refract. If the new medium is more dense than the old medium, the angle that the ray makes to the normal axis will decrease so that $\theta_1 > \theta_2$ (Figure 4).

Snell's law connects the ratio of the sines of these angles to the ratio between the so-called refractive indexes n_1, n_2 of these media:

$$n_1/n_2 = \sin \theta_2/\sin \theta_1.$$

The measured refractive indexes of different media were found to increase with the density of the media, but no further explanation for why this law obtained was apparent.[11]

If we follow Kitcher, then we should adopt an argument pattern as explanatory if it permits one to treat reflection and refraction together.[12] One such argument pattern deploys Fermat's principle that a light ray will travel between two points on the path that minimizes the time of the trip. Both the law of refraction and Snell's law can be derived in a uniform fashion from Fermat's principle. The derivations take the endpoints of the light ray's path to be fixed and vary the point O at which

10 More involved examples of the same kind are used to criticize Kitcher in Morrison (2000). See also Hafner and Mancosu (2008) for criticisms of Kitcher's proposal for explanatory proofs.
11 See Nahin (2004) for more on this case.
12 This assumes a set of statements K where there is no way to obtain an explanatory store over K that lacks this pattern and that scores better on Kitcher's criteria. Arguably, such a K was present during Fermat's time, prior to our current understanding of the nature of light.