**The Metarepresentational Role of Mathematics in Scientific Explanations**

Forthcoming in *Philosophy of Science*

Colin McCullough-Benner

School of Philosophy, Religion, and History of Science, University of Leeds

C.M.McCullough-Benner@leeds.ac.uk

**Abstract:** Several philosophers have argued that to capture the generality of certain scientific explanations, we must count mathematical facts among their explanantia. I argue that we can better understand these explanations by adopting a more nuanced stance toward mathematical *representations*, recognizing the role of mathematical representation schemata in representing highly abstract features of physical systems. It is by picking out these abstract but non-mathematical features that explanations appealing to mathematics achieve a high degree of generality. The result is a rich conception of the role of mathematics in scientific explanations that does not require us to treat mathematical facts as explanantia.

**Acknowledgements:** I thank Steven French and Juha Saatsi for detailed feedback on multiple drafts, audiences at the University of Leeds and the 2021 APA Eastern Division Meeting for helpful discussion, and Audre Brokes for thoughtful comments delivered at the APA meeting.

# 1 Introduction

Mathematics clearly plays a role in many scientific explanations. Whenever we represent a physical domain mathematically, mathematics can be used to represent explanatory physical facts, the explanantia of explanations in that domain.[1] For instance, to explain why Jamie and Jon took six hours to drive from Detroit to Milwaukee, we might appeal to the facts that the route they traveled was approximately 360 miles long and that they traveled at approximately 60 miles per hour. While these facts are represented with numbers, they are not themselves facts about numbers. In such explanations, mathematics is said to play a *merely representational* role. According to *representationalists*, such cases exhaust the role of mathematics in scientific explanations (e.g., Melia 2000; Daly and Langford 2009; Saatsi 2011, 2016).

In contrast, *explanationists* have argued that mathematics sometimes plays a further, *distinctively explanatory* role. In these cases, mathematical facts would be among the explanantia of non-mathematical explananda. For instance, suppose Jamie and Jon brought five sandwiches to eat during their drive. We might take the purely mathematical fact that five isn't divisible by two to explain the fact that they didn't manage to split their sandwiches evenly (without cutting). This is what Saatsi (2016) calls a "thick" explanatory role: mathematical facts stand in an ontic relation of explanatory relevance to their explananda.[2] In such cases,

[1]Here and throughout this paper, I use 'explanans' and 'explanandum' to refer to the relata of the worldly dependence relations involved in successful explanations, rather than the communicative devices used to pick those relata out. For the purposes of this paper, I assume that such dependence relations must be involved in one way or another in the explanations I discuss.

[2]Note that this means that the influential modal account of mathematical explanations in

facts about the target system would have to stand in an objective dependence relation to mathematical facts, construed as genuine constituents of the world.

This role for mathematics has been defended in part on the grounds that some explanations have features that can only be explained if we take the mathematics involved to play a distinctively explanatory, rather than merely representational role in the explanation.[3] These explanations are thought to be more general than explanations in which mathematics plays a merely representational role, in that they carry more counterfactual information about their target systems (scope-generality) and in that they share an "explanatory core" with explanations of phenomena in other domains (topic-generality). For instance, parallel explanations in terms of divisibility explain why Jamie and Jon couldn't evenly split seven—or nine, or any odd number of—sandwiches (scope-generality) and why a philosopher can't evenly divide their five good ideas between two papers (topic-generality). In contrast, we couldn't similarly generalize the explanation of why it took six hours to travel from Detroit to Milwaukee, since the explanantia of that explanation are concrete features of the target system, rather than more general features shared by, say, systems in which different distances are traveled (scope-generality) or systems in entirely different domains (topic-generality). For much the same reason, we couldn't similarly generalize an explanation of why Jamie and Jon couldn't evenly split their sandwiches that appealed only to concrete properties of the physical system consisting of Jamie, Jon, and their sandwiches.

---

science proposed by Lange (2013) is not explanationist in the sense I have in mind here.

[3]For example, Colyvan (2002), Baker and Colyvan (2011), Lyon (2012), Plebani (2016), Baker (2017), and Baron (2020) all argue for a distinctively explanatory role for mathematics on the grounds of explanatory generality or cognate notions like robustness or unification.

The debate between representationalists and explanationists has primarily been carried out in the context of evaluating indispensability arguments for mathematical platonism. While the classic Quine-Putnam indispensability argument supports platonism on the grounds that mathematics is indispensable to our best scientific theories, more recent versions of the argument appeal to its *explanatory* indispensability. In short, the thought is that we should believe in mathematical objects for the same reason we believe in non-mathematical theoretical posits—due to their indispensable role as explanantia in our best scientific explanations.[4] While such arguments might fail for other reasons, their success depends in large part on whether mathematics plays the right sort of role in our best scientific explanations—in particular, the distinctively explanatory role described above, according to which physical explananda stand in ontic explanatory dependence relations to mathematical explanantia (Saatsi 2016). And it is on precisely this point that representationalists and explanationists disagree, with the former denying and the latter affirming that mathematics plays such a role.

But the highly abstract scientific explanations at the center of this debate are independently important. Explanations of this kind are crucial to the burgeoning literature on non-causal explanations in science (see, for example, the papers in Reutlinger and Saatsi 2018). Understanding how they work—and in particular how mathematics contributes to their explanatory generality—is crucial to understanding scientific explanation more generally. In this regard, existing explanationist and representationalist approaches *all* leave something to be desired.

If explanationists are right, we have to explain the very possibility of a mathematical fact's

[4]Perhaps the most influential formulation in terms of explanatory indispensability is in Baker (2005), which has spawned an enormous literature. For a survey, see Mancosu (2018, §3.2).

standing in an ontic relation of explanatory relevance to a physical fact. This is far from straightforward. Most realists about mathematics take mathematical and physical entities to belong to different ontological categories. But even nominalist interpretations and paraphrases of mathematical language—for instance, in terms of modal information about possible structures (modal structuralism) or facts about mathematical practice (some forms of fictionalism)—don't obviously pick out the right sort of entities to stand in such relations. We also would have to explain why increasing the *degree* of generality of an explanation eventually yields a different *kind* of explanation altogether—viz., one in which the explanans must be a mathematical, rather than physical, fact. But it is not obvious that such a difference in degree should yield a difference in kind .[5]

On the other hand, representationalist accounts, when they explicitly address the nature of mathematical representation, typically rely on an austere conception according to which mathematical facts simply "index" physical ones (Melia 2000; Daly and Langford 2009), which fails to do justice to the full range of contributions mathematics makes to science. Such accounts have little to say about how mathematics contributes to explanatory generality. Arguably the best move for the representationalist to make in the context of debates over the explanatory indispensability of mathematics is to deny that degrees of scope- or topic-generality beyond those straightforwardly available to the nominalist are explanatory *virtues* that would support the inference to the best explanation central to explanatory indispensability arguments (Knowles and Saatsi 2019). But regardless of whether these degrees of generality are virtuous—a question on which we can remain agnostic for the purposes of this paper—they are properties of certain explanations in which mathematics plays a prominent

[5]For an argument to that effect, see Jansson and Saatsi (2019).

role, and a good philosophical account of such explanations should be able to account for them.

In this paper, I argue that by looking more closely at the role of mathematics in scientific *representations* we can understand how scientific explanations can achieve such high degrees of scope- or topic-generality without including mathematical facts among their explanantia. In addition to representing particular target systems, mathematics allows us to represent properties of these representations that remain stable as the mathematics involved and its physical interpretation are allowed to vary. This is what I call the *metarepresentational role* of mathematics. In using the word 'metarepresentational' here, I do not mean to imply that in this role mathematics is just a "meta-level device" for reasoning about science from the outside, but rather to describe the use of mathematics within science to reason about features shared by collections of individual representations.[6] This metarepresentational contribution of mathematics allows us to reason about the very abstract features of physical target systems in virtue of which they are accurately represented by certain *kinds* of mathematical representation. These abstract features of target systems, rather than the mathematical facts relevant to representing them, are the explanantia of highly general explanations. While mathematics is at least practically necessary to pick out these features, this use of mathematics does not commit us to the truth of any purely mathematical (as opposed to physically interpreted) claim, and so does not support explanatory indispensability arguments for mathematical platonism. The result is a richer conception of the role of mathematics in scientific explanations that

[6]That said, if the reader balks at the use of the word 'metarepresentational' for any reason, what I call the "metarepresentational role" can without great loss be taken to be part of a significantly enriched account of the representational role of mathematics, rather than a distinct role.

significantly improves on existing representationalist and explanationist accounts.

In §2, I develop a more nuanced account of the role of mathematics in scientific representations, presenting an account of mathematical representation (§2.1) and extending it to explain the metarepresentational contributions of mathematics (§2.2). I then distinguish between scope- and topic-generality (§3) and present a scientific explanation, the famous number-theoretic explanation of the cycles of periodical cicadas, that has been taken to exhibit both (§4). In reference to this example, I show how we can account for both the scope-generality (§5) and the topic-generality (§6) of explanations appealing to mathematics in terms of the metarepresentational contributions of the mathematics without appealing to pure mathematical facts. Finally, I respond to the objection that explanations in which mathematics plays only a metarepresentational role have less explanatory depth than those in which mathematics plays a distinctively explanatory role (§7).

## 2 Representation and Metarepresentation

In its representational role, mathematics serves to construct individual representations of particular target systems. In its metarepresentational role, it serves to elucidate the properties shared by collections of individual representations, as well as very general properties of these representations' target systems that can be picked out only by reasoning about the features of collections of representations. The purpose of this section is to clarify and illustrate this central distinction before applying it to the case of scientific explanations appealing to mathematics in the rest of the paper.

## 2.1 Representation

In its representational role, mathematics helps scientists to construct representations of particular non-mathematical target systems. These representations support making inferences about their target systems on the basis of the relevant mathematics. To focus the discussion, I appeal to a particular account of how this works: the *Robustly Inferential Conception* (RIC) of mathematical scientific representations proposed by McCullough-Benner (2020, §4). I choose to work with RIC because I take it to represent the central components of mathematical scientific representations more perspicuously than the more standard mapping account in both its classic (Pincock 2012) and inferential (Bueno and Colyvan 2011) forms.[7] Nonetheless, with minor adjustments, the rest of this paper could be cast in terms of either version of the mapping account instead.[8]

According to RIC, mathematics places constraints on what the target system of a mathematical scientific representation must be like by specifying (physical) inferences about the target system that must preserve truth if all of the representation's informational content is true. Such representations have three ingredients:

**(RIC1)** a *physical interpretation* of the language of the mathematical theory sufficient to

[7]It also has the benefit of better accommodating episodes of scientific practice involving the application of inconsistent and otherwise unrigorous mathematics (McCullough-Benner 2020).

[8]Mapping accounts can be understood as special cases of RIC, with (RIC1) provided by the relevant structure and mapping, (RIC2) by a "structure-generating description" (Nguyen and Frigg 2017), and (RIC3) by those inferences that preserve truth when interpreted in terms of the relevant mathematical structure.

provide at least some expressions in this language with physical truth conditions,

**(RIC2)** an *initial description of the target system* in the language of the mathematical theory, under the physical interpretation RIC1,[9] and

**(RIC3)** a *collection of privileged inference patterns* from those licensed by the original mathematical theory.

Now consider the closure of RIC2 under the set of mathematical inference patterns in RIC3. The physical interpretation RIC1 assigns truth conditions to a subset of these mathematical expressions. It is precisely these expressions, under the physical interpretation provided by RIC1, that spell out the informational content of the representation.

Consider a model of a weight suspended from a spring as a damped harmonic oscillator. Its behavior is represented by the differential equation $m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = 0$ along with an initial condition and perhaps equations determining values for the constants $m$, $c$, and $k$. Together (and under the physical interpretation RIC1), these equations provide component RIC2, the initial description of the target system. In this case, $x$ is interpreted as the vertical displacement of the object suspended from the spring (and its derivatives with respect to $t$ the vertical components of velocity and acceleration); $t$ is interpreted as time elapsed; $m$ is interpreted as the mass of the object suspended from the spring; $c$ represents the effect of damping (related to

[9]RIC2 is an "initial" description in the sense that it must be in place before the representation can be used to make inferences about its target system. It should not be confused with "initial conditions", which may or may not be part of RIC2 and in no case exhaust it. Mapping accounts similarly must posit a "structure-generating description" to specify the structure of the target system in a way that allows an appropriate mapping to be constructed (Nguyen and Frigg 2017).

the force $F_f$ due to friction by the equation $F_f = -c\frac{dx}{dt}$); and $k$ represents the effect of the tension of the spring (related to the force $F$ required to extend the spring by length $x$ by the equation $F = -kx$). This suffices to provide numerical solutions of this and related equations with physical truth conditions, and so it suffices for RIC1. Finally, RIC3 is simply the set of inference patterns licensed by real analysis.[10]

Now, consider the closure of the (mathematically interpreted) set of equations in RIC2 under the inferences licensed by real analysis. A subset of the expressions in this set are assigned physical truth conditions by the interpretation RIC1. The expressions in this subset, under that physical interpretation, constitute the informational content of the representation. This then allows us to see how such a representation can support surrogative reasoning about the target system via mathematical reasoning about the real numbers. Whatever results we can derive purely mathematically in real analysis by means of the equations in RIC2 can be brought to bear on the physical target system by interpreting them according to the interpretation rules in RIC1, provided that RIC1 supplies those results with physical truth conditions. Since such physically interpreted claims are by definition part of the informational content of the representation, such inferences preserve truth on the condition that all of the representation's informational content is true, and are therefore licensed by the representation.[11]

---

[10]Restriction of the inference patterns in RIC3 is only required when the mathematics is inconsistent or otherwise unrigorous (McCullough-Benner 2020).

[11]Of course, most—if not all—scientific representations have some informational content that is untrue. What matters to scientists is instead whether their representations are accurate in the appropriate respects. In light of this, I might seem to be presupposing an implausibly close relationship between truth-preservation and accuracy. But I need not make any such

## 2.2 Metarepresentation

A *representation schema* is a collection of mathematical representations that share common features but vary with respect to some aspect of the representation. In particular, this means that one may vary the physical interpretation assigned to the mathematical vocabulary (RIC1), the specification of the target system (RIC2), or the mathematical apparatus itself (RIC3).

To see how this works, consider how we might extend the representation of a weight suspended from a spring to a representation schema. Since a number of very different physical systems can be usefully modeled as harmonic oscillators, we might be interested in schemata that allow either the physical interpretation of the mathematical language (RIC1) or the initial description of the target system (RIC2)—or both—to vary in certain ways.

For instance, we might be interested in the features the representation from the previous section shares with representations that specify other values for the constants $m$, $c$, and $k$. In that case, we could reason about the schema consisting of representations that share the same physical interpretation (RIC1) and underlying mathematical framework (RIC3) but differ with respect to the initial description of the target system (RIC2) in that (at most) different values are given to the constants $m$, $c$, and $k$.

We might also be interested in features that this representation shares with those of other kinds of systems that can be represented as damped harmonic oscillators, such as pendulums or certain electrical circuits. In that case, we might start with the representation schema

---

presupposition. If a representation is accurate in some respects but has some false informational content, there is no guarantee that the inferences licensed by the representation are actually truth- or accuracy- preserving. Whether such an inference should be made depends also on whether the representation is accurate in the right respects.

considered above, which allows the values of the relevant constants to vary, and further allow the physical interpretation (RIC1) to vary in a limited way (so that the relevant mathematics can be interpreted in terms of pendulums and electrical oscillators but not other sorts of target system). For instance, to allow representations of pendulums to be instances of our schema, we must allow $x$ to be angular (rather than vertical) displacement, $m$ to be rotational inertia (rather than mass), and so on. If we also wish to include representations that do not specify these constants directly, but derive them from other features of the target system, we may need to allow more substantial changes to the initial specification of the target system (RIC2)—namely by allowing it to incorporate the (physically interpreted) equations used to derive the values of these constants.

Reasoning with representation schemata allows us to do two things we cannot do with individual representations alone. First, the mathematics common to the instances of the schema (or a mathematical framework into which these instances are embedded) can be used to shed light on features shared by the individual representations in the schema. Second and closely related, reasoning with the schema allows for reasoning about general features shared by the target systems represented by individual representations in the schema, which are not captured (in their full generality) by these individual representations. As might be expected, capturing these general features, both of representations and their target systems, is central to achieving the kind of generality possessed by scientific representations in which mathematics seems to play an explanatory role, to which I turn in the next section.

# 3 Explanatory Generality

Some explanationists have claimed that the degree and kind of generality possessed by certain scientific explanations requires them to include mathematical facts among their explanantia. Baker (2017) gives perhaps the clearest statement of this idea by distinguishing two kinds of generality that mathematics might help us achieve: scope-generality and topic-generality.

## 3.1 Scope-Generality

An explanation is more *scope-general* the wider the range of counterfactual situations to which it applies in which the explanans is varied (but involving the same sort of target system). Consider the explanation of why Jamie and Jon couldn't evenly split five sandwiches that appeals to the mathematical fact that five is not divisible by two. This explanation is scope-general in that a parallel explanation can be used to determine what happens whatever number of sandwiches or passengers there are: one can divide $n$ sandwiches evenly among $m$ passengers if and only if $n$ is divisible by $m$. And so we can explain by exactly the same means why it's not possible to evenly split nine sandwiches between two people, five sandwiches among three people, and so on. A less scope-general explanation would explain why Jamie and Jon cannot evenly split five sandwiches without telling us anything about what happens when the number of sandwiches or passengers is different. And an explanation with an intermediate degree of scope-generality might tell us what would happen if there were $n$ sandwiches and $m$ passengers, provided $n$ and $m$ don't exceed some finite upper bound.

Some explanationists have argued that, for some explananda, explanations in which mathematics plays a distinctively explanatory role can achieve a degree of scope-generality that cannot be matched by explanations in which mathematics plays a merely representational role.

For instance, Baker and Colyvan (2011, 331) argue that explanations in which mathematics plays a merely representational role must be "less general and less robust" (i.e., less scope-general) than those in which the mathematics is explanatory. Lyon (2012, 567) similarly claims that explanations in which mathematics plays an explanatory role are more "robust"[12] with respect to causal-historical details. Plebani (2016, 553) criticizes certain nominalist-friendly explanations on the grounds that they operate "at the wrong level of generality" for the similar reason that a nominalistic explanation's explanans would have to include too many concrete details.

After presenting a more sophisticated example of a scope-general explanation in §4, I will argue in §5 that these arguments miss the mark. The scope-generality of explanations in which mathematics plays a distinctively explanatory role can be matched by that of explanations in which it plays only a metarepresentational role. Such explanations are formulated in terms of a representation schema that carries more counterfactual information than individual representations of the relevant target system.

## 3.2 Topic-Generality

An explanation is *topic-general* to the extent that it does not depend on the concrete features of a particular target system, but has an explanatory core that can be used to formulate parallel explanations about target systems of radically different types. Consider again the mathematical explanation of why Jamie and Jon cannot evenly split five sandwiches. The core of this

[12]The sense of 'robust' in both Baker and Colyvan (2011) and Lyon (2012) should be distinguished from its meaning in more general discussions of modeling. Baker, Colyvan, and Lyon have something narrower in mind—viz., a kind of explanatory scope-generality.

explanation has nothing to do with Jamie, Jon, or sandwiches. A parallel explanation can be formulated for any finite system of discrete objects that one might want to partition into a certain number of equinumerous groups. The explanatory core is something like

1. It is possible to divide $n$ $F$s into $m$ even groups if and only if $n$ is divisible by $m$.

2. Five is not divisible by two.

3. So, it is not possible to divide five $F$s into two even groups.

We might think of this as an explanation schema à la Kitcher (1989) that can be filled in by replacing '$F$' with a term referring to some kind of discrete object to produce a concrete explanation.[13] In that case, we get our toy explanation by substituting 'sandwiches' for '$F$'. But we get just as good an explanation by replacing '$F$' with any other expression picking out a group of discrete objects. For instance, we can produce in this way an explanation of why it is not possible to divide five students into two even discussion groups or of why it is not possible

[13]Central to Kitcher's unificationist account of explanation is the notion of an argument pattern, which consists of a deduction with the non-logical vocabulary replaced with schematic letters, a set of "filling instructions" specifying how those letters can be filled in to produce a legal instance of the schema, and a classification of the schematic argument. According to the unificationist account, an ideal explanation is an instance of an argument pattern belonging to the set of such patterns that best unifies the set of beliefs accepted by scientists at a particular time. But we need not accept the entirety of this unificationist account to represent explanations in terms of argument patterns in this way. A similar strategy is pursued by Baron (2020), who proposes a hybrid unificationist-counterfactual account of mathematical explanations in science.

to distribute one's five good ideas evenly across two papers (assuming, of course, that ideas are discrete objects).

Alternatively, we might think of this as a set of propositions that is itself topic-general in that it doesn't appeal to concrete features of any type of non-mathematical target system in particular. Rather, if we take (1) and (3) to implicitly quantify over all types $F$ of discrete objects and (1) to implicitly quantify over all $n, m \in \mathbb{N}$, it appeals only to a mathematical fact about natural numbers (2) and a proposition connecting mathematical facts with very general facts about dividing discrete objects into groups (1). Particular explanations are produced by supplementing this explanatory core with propositions about the particular domain of the explanation. We get our toy explanation by adding the proposition that sandwiches are discrete objects and concluding from this and (3) that five sandwiches cannot be divided into two even groups. But we might add other propositions to produce parallel explanations—for instance, that students are discrete objects or that ideas are.

Explanationists have argued that topic-generality is a distinguishing feature of explanations in which mathematics plays an explanatory role. In several places, Colyvan (2002, 72; 2013, 1042) claims that mathematics is genuinely explanatory due to its unifying power—i.e., its ability to produce topic-general explanations. The central claim of the paper in which Baker (2017) articulates the distinction between scope- and topic-generality is that topic-generality requires mathematics to play a genuinely explanatory role even if scope-generality does not. Along similar lines, Baron (2020) develops a view according to which topic-generality (alongside a few other conditions) distinguishes genuinely mathematical explanations from explanations in which the mathematics is merely representational.

After considering a more sophisticated example (§4) and the case of scope-generality (§5), I will argue in §6 that we can explain the topic-generality of such scientific explanations without treating the mathematics as playing a distinctively explanatory role. Again the mathematics should be understood as playing a metarepresentational role. In this case, the core of an explanation can be formulated in terms of a representation schema that allows the interpretation of the mathematics to vary. The very abstract features of physical target systems in virtue of which they are correctly represented by some instance of this schema then serve as explanantia in the topic-general explanation.

# 4 Example: The Prime Cycles of Periodical Cicadas

A classic example of an explanation that is both scope- and topic-general is the number-theoretic explanation of the prime periods of periodical cicadas, first discussed in the philosophical literature by Baker (2005). Periodical cicadas spend most of their lives in a larval stage underground, emerging as adults for a single season, during which they reproduce and die. In extant species, this life-cycle lasts either thirteen or seventeen years. Why are these period lengths adaptive?[14]

According to Goles, Schulz, and Markus (2001), it is because prime periods optimize evasion of periodical predators with shorter life-cycles. Under certain assumptions, where $x$ is

[14]Most presentations of this case, including Baker's original paper, treat the *existence* of 13- and 17-year cycles as the explanandum, and what I say in the rest of the paper can naturally be adapted to such an explanation. However, there is good reason to doubt that the optimality of prime cycles actually played a role in their selection (Wakil and Justus 2017). As Wakil and Justus conclude, we can avoid this problem by instead treating the *adaptiveness* of 13- and 17-year cycles as the explanandum of the number-theoretic explanation.

the length of the cicadas' life-cycle and $y$ the length of the predators' life-cycle, the cicadas' average fitness over a period of $xy$ years is a decreasing function of $\gcd(x, y)$, while the predators' average fitness over the same period is an increasing function of $\gcd(x, y)$. By definition, all and only prime numbers minimize gcd with all smaller positive integers. Assuming the predators must have a shorter period than the cicadas, this means that it is only when cicadas have a prime period that neither the cicadas nor their predators could increase their fitness by changing their period. To reflect biological constraints, Goles et al. constrain cicada periods to be between 12 and 18 years, resulting in 13- and 17-year periods.[15]

This explanation is both highly scope general and highly topic general, and these features will be the focus of the next two sections. I will argue that these two kinds of generality do not give us reason to think the mathematics is playing a distinctively explanatory role. Instead, the

[15]My presentation of this example is simplified in several ways in order to streamline the rest of the paper. For one thing, the explanation I present is not the only number-theoretic explanation of this phenomenon. Another discussed by Baker (2005) is the suggestion that prime periods are advantageous because they minimize hybridization with other periodical cicada species. What I say in the rest of the paper can straightforwardly be adapted to this explanation. For another, I ignore several important biological details, such as the pressures of nymphal crowding and deviations from 13- and 17-year periods (most often by four years) sometimes observed in existing cicada species. (See, for example, Wakil and Justus 2017.) While these details are essential to explaining why cicadas actually developed these periods, one can explain why those periods are *adaptive* without them. Since I only discuss the latter sort of explanation, and since my aim is not to give a complete account of this particular case, I take these omissions to be harmless.

move from representation to metarepresentation allows us to pick out features of the phenomenon that are sufficiently general to ground explanations with the same degree of scope-generality (§5) and topic-generality (§6) as explanations in which mathematics plays a distinctively explanatory role without appealing to pure mathematical facts.

# 5 Metarepresentation and Scope-Generality

The number-theoretic explanation is highly scope-general. It does not just explain why the actual periods of actual cicadas are advantageous, but also why prime cycles would be advantageous if cicadas had different biological constraints on their life-cycles. Supposing cicada periods were biologically constrained to another range (say, 12 to 25 years), we have a parallel explanation of why prime periods within that range would be advantageous. The number-theoretic explanation indeed seems to be more scope-general than explanations that appeal only to nominalistically acceptable properties, which would seem to be limited to cases in which the cicadas' periods cannot exceed some finite bound.

The best existing responses to this line of thought concede for the sake of argument that explanations in which mathematics plays a distinctively explanatory role enjoy a higher degree of scope-generality than other explanations, but downplay the importance of such high degrees of generality. For instance, Knowles and Saatsi (2019) argue persuasively that scientists have no reason to prefer an explanation that holds for all ranges of possible cicada periods to one that holds only for periods up to some suitably high finite bound. If, say, the nominalistic explanation works for cicada period lengths up to the age of the universe, the additional counterfactual information provided by the explanation in which mathematics plays a distinctively explanatory role will be of dubious scientific value indeed!

But regardless of whether the high degree of scope generality enjoyed by the mathematical explanation is indeed a virtue, it does seem to be a genuine feature of this and other scientific explanations in which mathematics plays a prominent role. A complete philosophical understanding of the role of mathematics in scientific explanations therefore still requires an account of what makes this greater degree of generality possible (or why, despite appearances, it is not possible). In the rest of this section, I provide such an account in terms of the metarepresentational role of mathematics. If this account is correct, nominalists need not settle for explanations with limited scope-generality like those defended by Knowles and Saatsi.

The scope-generality of the explanation can naturally be understood in terms of a representation schema in which only the range of biologically feasible life-cycles for the cicadas and their predators is allowed to vary. Instances of this schema are those that can be produced from Goles et al.'s initial representation by varying the initial description of the target system (RIC2) with respect to the range of possible cicada periods. For the explanation to go through for all instances of the schema, we must also require that the maximum predator period is shorter than the minimum cicada period, that the range of predator periods is sufficiently broad to make all non-prime cicada periods unstable, and that the range of cicada periods contains at least one prime.

Instances of the resulting schema represent particular counterfactual situations in which the cicadas and predators are biologically constrained to have life cycles within a particular range. For each of these instances, we can run a parallel explanation of why the prime cycles within that range are optimal for predator avoidance. Where $p_1, \ldots, p_n$ are the prime numbers in the range of numbers representing biologically feasible cicada periods and $x$ and $y$ represent the cicadas' and predators' periods in years, respectively, a parallel series of applications of the inference patterns in (RIC3) and purely physical inferences yields the conclusion that periods of

$p_1$ or ... or $p_n$ years are optimal for predator avoidance:

1. Via the inference patterns licensed by number theory, infer that $p_1, \ldots, p_n$ are prime.

2. From 1, infer $\gcd(p_i, y) = 1$ for $1 \leq i \leq n$ and $y_{\min} \leq y \leq y_{\max}$.

3. From 2, infer that there is no $y'$ such that $\gcd(p_i, y') < \gcd(p_i, y)$ when $1 \leq i \leq n$, $y_{\min} \leq y \leq y_{\max}$, and $y_{\min} \leq y' \leq y_{\max}$.

4. From 2, infer that there is no $x'$ such that $\gcd(x', y) < \gcd(p_i, y))$ when $1 \leq i \leq n$, $x_{\min} \leq x' \leq x_{\max}$, and $y_{\min} \leq y \leq y_{\max}$.

5. From the physical interpretation of 3, infer that there is no way for the predators to increase their average fitness[16] by changing their period, provided the cicadas' period is $p_1$ or ... or $p_n$ years.

6. From the physical interpretation of 4, infer that there is no way for the cicadas to increase their average fitness by changing their period, provided that period is $p_1$ or ... or $p_n$ years.

7. From 5 and 6, infer that periods of $p_1, \ldots,$ and $p_n$ years are optimal with respect to periodical predator avoidance.

8. By a similar chain of reasoning, infer that, when the cicadas' period is not $p_1$ or ... or $p_n$ years, either the predators or the cicadas can increase their fitness by changing their

[16]This is how Goles et al.'s (2001) model works, anyway. Really, we should take their predator and cicada fitness functions to represent only the ways interactions between the predator and cicada periods contribute to overall fitness.

period, and so those periods are not optimal with respect to periodical predator avoidance.

9. Conclude from 7 and 8 that the cicadas' optimal periods with respect to periodical predator avoidance are exactly $p_1, \ldots,$ and $p_n$ years.

Recognizing that this is the case for each instance of the schema allows us to formulate a more general explanation of why periodical cicadas have prime cycles: no matter which instance of the representation schema picks out the right ranges of biologically feasible periods for the cicadas and predators, the chain of inferences above goes through. This explanation goes beyond the explanations made available by the individual instances of the representation schema by incorporating all of the counterfactual information expressed by the schema; it tells us not just what happens given that cicada lifecycles are constrained to be (say) between 12 and 18 years, but also what would happen if that range were different. This is made possible by reasoning not about the particular physical facts represented by a given instance of the schema, but the features shared by the systems represented by all these instances, which are the explanantia of the scope-general explanation. And this is achieved by reasoning about the schema as a whole. The result is scope-generality: the explanation generalizes over all relevant ranges of biologically feasible predator and cicada periods because it generalizes over all instances of the representation schema.

At this point, one might worry about the use of mathematical vocabulary in setting out the explanation above. In particular, (1) appears to be nothing but a pure mathematical fact, since 'prime' is not assigned a physical interpretation by the RIC1 component of the representations in the schema. And (2) and (3) also appear to state pure mathematical facts, albeit ones that can also be given a physical interpretation. And if this is the case, it seems wrong to say that mathematical facts are not among the explanantia of this explanation.

But this rests on a misunderstanding the role of (1)–(8) above. (1)–(4) result solely from the application of the inference patterns in the RIC3 component—the collection of privileged mathematical inference patterns from those licensed by the original mathematical theory—of the instances of the representation schema. The role of the inference patterns in RIC3 is solely to determine which inferences from physically interpreted claims in the language of the mathematical theory to physically interpreted claims in that language preserve truth according to the representation. For this purpose, whether (1)–(4) are true under their standard mathematical interpretation is irrelevant. Even supposing a form of mathematical error theory were true, so that the mathematical interpretations of (1)–(4) were false, (5)–(8) would still be consequences of each representation in the schema, since the inference patterns yielding (1)–(4) would still be in RIC3, and each of these representations has as part of its informational content the physically interpreted versions of any claims derivable from the initial specification of the target system RIC2 via the inference patterns in RIC3. So, while (1)–(4) have the surface appearance of statements of mathematical facts, in this context they merely express the inferences licensed by the instances of the representation schema. As a result, the mathematics involved in their expression does not play a distinctively explanatory role.

# 6 Metarepresentation and Topic-Generality

The cicada explanation is also topic-general. According to Baker (2017) (cf. Baron 2020), it has an explanatory core that is not specific to periodical cicadas but that applies generally to phenomena involving unit cycles with certain features.

Following Baker (2017, 201f), but with adjustments to fit the description of the example here, this explanatory core consists of the propositions:

**($M_1$)** The gcd of numbers $m, n$ is minimal if and only if $m$ and $n$ are coprime. (pure mathematical fact)

**($UC_1$)** The number of co-occurrences of the same pair of cycle elements of two unit cycles of periods $m$ and $n$ in an interval of length $mn$ is equal to $\gcd(m, n)$. (fact about unit cycles)

**($UC_2$)** So, any pair of of unit cycles with periods $m$ and $n$ minimizes the number of co-occurrences of the same pair of cycle elements over an interval of length $mn$ if and only if $m$ and $n$ are coprime. (fact about unit cycles, from $M_1$, $UC_1$)

**($M_2$)** All and only prime numbers are coprime with all smaller numbers. (pure mathematical fact)

**($UC_3$)** So, given a unit cycle $p_m$ of length $m$ and a range of unit cycles $q_i$ with lengths $i < m$, $p_m$ minimizes the number of co-occurrences of the same pair of cycle elements over an interval of length $mi$ for all $q_i$ if and only if $m$ is prime (fact about unit cycles, from $M_2$, $UC_2$)

From $UC_3$, thus derived, and facts about periodical cicadas and their predators in particular, we can then derive the conclusion that 13- and 17-year periods are optimal with respect to predator avoidance, reasoning in much the same way as above.

This explanatory core is then (at least in part) shared by explanations of phenomena in other domains. For instance, we might give a parallel explanation of why the number of teeth on the front and rear gears of brakeless fixed-gear bicycles are optimal when they are coprime (Baker 2017, 203f). In this case, we need only $M_1$, $UC_1$ and $UC_2$, since appeal to primeness is unnecessary. We then can add particular facts about fixed-gear bikes to derive the explanandum. Stopping such a bike involves locking the pedals at a particular point (so that the front gear is always in the same position), thus locking the rear tire, causing the bike to skid to a stop. This causes wear on the tire where it skids. To maximize the life of the bike's tires, one should choose a gear ratio that minimizes the frequency with which the same part of the tire skids when the pedals are in braking position. Since we can treat the positions of the front and rear gears of the bike as unit cycles, $UC_2$ tells us that this happens just when the numbers of teeth on the two gears are coprime.

The thought is then that topic-generality of this kind can only be achieved when mathematical facts are included as part of the explanatory core. The explanatory core of an explanation in which mathematics doesn't play an explanatory role must appeal to particular facts about the domain of the explanation that prevent it from generalizing to other domains in this way. If we explain why 13- or 17-year cicada periods are advantageous in terms of properties of temporal intervals, as Saatsi (2011) does, this rules out using the same explanantia to explain the optimal gear configurations of fixed-gear bicycles.

In the rest of this section, I argue that topic-generality is possible because mathematics plays a metarepresentational role in the explanation. The explanatory core doesn't include pure mathematical facts like $M_1$ and $M_2$ above, but only very general, mathematically represented physical facts. And representation of these very general facts is made possible by the move from individual mathematical representations to a representation schema. Recall the pendulum example from section 2. Moving to a representation schema in which the physical interpretation RIC1 was allowed to vary made it possible to reason about the features such a pendulum shares with radically different systems, like electrical oscillators. A similar move to a schema in which RIC1 is allowed to vary allows us to capture features shared by periodical cicada populations and fixed-gear bicycles.

All that was needed for the scope-general explanation above to go through was for there to be two entities with cycles of length $x$ and $y$ in some unit such that $x_{\min} \leq x \leq x_{\max}$ and $y_{\min} \leq y \leq y_{\max}$ for $x_{\min}, x_{\max}, y_{\min}, y_{\max} \in \mathbb{N}$, with a guarantee that $x > y$, and that it is optimal to minimize the intersection between these cycles over an $xy$-year period. These features constitute the "explanatory core" of the scope-general explanation in the previous subsection. The basic predator-prey dynamics represented in Goles et al.'s model explain how these features are realized, but the explanatory core is independent of those details. In this respect, the explanatory core I present here is analogous to the explanatory core proposed by Baker (2017). The crucial difference is that Baker's explanatory core includes pure mathematical facts ($M_1$ and $M_2$) relating prime numbers, gcd, and coprimeness, while mine includes only very general, mathematically represented physical facts. But the explanatory core I present here allows for the formulation of explanations just as topic-general as those built on the explanatory core presented by Baker.

We can achieve this degree of topic-generality by moving to a yet more general representation schema that allows the mathematical representation to vary, as long as the three features above (or potentially a subset thereof) are preserved. Clearly, it is necessary to allow the physical interpretation of the mathematical language (RIC1) to vary, so that $x$ and $y$ can come to represent, say, spatial magnitudes, or numbers of teeth on interlocking gears, rather than temporal magnitudes. In addition, it will be necessary to allow the initial description of the target system (RIC2) to vary with respect to its mathematical formulation in order to capture the different ways in which the three conditions above might be realized.

Parallel explanations can then be given for phenomena represented by instances of this schema that represent phenomena in radically different domains. Consider again the explanation of why it is optimal to have coprime numbers of teeth on the gears of brakeless fixed-gear bicycles, as discussed by Baker (2017).[17] In this case, $x$ and $y$ come to represent the period of the front and rear gears, respectively, where the appropriate unit is teeth. Since the numbers of teeth on the gears do not tend to be prime (but only coprime), it is not necessary to restrict the representation schema to cases in which $x > y$ (though this happens to hold in realistic cases). There is some finite possible range of values for $x$ and $y$ determined by various constraints, such as that the gear ratios should be within a reasonable range for the purposes of actually riding the bike and that the gears should be possible to manufacture for a reasonable price. Configurations that minimize the co-occurrence of the same pair of cycle elements (i.e., gear positions) are optimal in this case because they maximize the wear on the rear tire maximally even, thus maximizing its useful life. So we have a subset of the explanatory core of the cicada explanation, together with an explanation of how the target system realizes the abstract conditions in the explanatory core.[18]

[17]Baron (2020) presents a similar explanation of why interlocking gears on machines in general maximize the life of the machine by ensuring even wear. What I say about Baker's example can naturally be adapted to Baron's.

[18]Rather than formulate this explanation in terms of the explanatory core of the cicada explanation, construed as a set of propositions, we can equivalently formulate an explanation schema à la Kitcher (1989) or Baron (2020) that unifies the explanations of the cycles of periodical cicadas and of the gear configurations of fixed gear bikes. Whenever we have a target system that is accurately represented by an instance of our representation schema, a formally identical argument explains the parallel explanandum for that target system. The difference between the sort of explanation schema I propose and the one Baron (2020) proposes then parallels the difference between my explanatory core and the one proposed by Baker (2017). Baron's schema

This gives us the essential ingredients for an explanation of the optimal gear configurations of fixed-gear bikes that runs parallel to the explanation of the optimal periods of periodical cicadas: for each instance of this schema physically interpreted in terms of fixed-gear bikes, we can run an explanation of why certain gear configurations are optimal by showing that these gear configurations minimize the co-occurrence of the same pair of gear positions by checking each of these gear configurations individually. And if we wish to produce a corresponding scope-general explanation, we can do so by appealing to the coprimeness of the numbers of teeth on the gears in the same deflationary way as the appeal to primeness in the scope-general cicada explanation: no matter which instance of the representation schema is accurate, coprime periods maximize the life of the rear tire by minimizing the frequency with which the rear gear (and so the rear tire) is configured in the same way when the front gear is in braking position. As in the case of the scope-general cicada explanation, this use of 'coprime' does not involve an appeal to mathematical facts, since it serves only to pick out certain patterns of inference (RIC3) common to instances of the representation schema, and these patterns of inference would be licensed by the representations regardless of whether the relevant statements including mathematical vocabulary like 'coprime' were true when given their usual mathematical interpretation. And so we can understand the topic-generality of the cicada explanation without ascribing a distinctively explanatory role to the relevant mathematics.

---

appeals to pure mathematical facts (as opposed to merely mathematically represented physical facts) and in particular to subjunctive conditionals with purely mathematical antecedents. In contrast, my approach allows us to formulate an explanation schema that appeals only to very general, mathematically represented physical facts.

# 7 Explanatory Depth

Baker anticipates something similar to the account I provide in the previous section and objects: "[A]lthough a schema about unit cycles has more topic-generality than [a] schema about life-cycle periods, instantiations of the schema will still end up being treated as disjoint facts, and thus the overall explanation will be less unified and will have less explanatory depth than the full [mathematical explanation of the cicadas' periods]" (2017, 12). So even if the representation schema I discuss at the end of the last section can support topic-general explanations, the resulting explanations lack the *depth* of explanations in which mathematics plays a distinctively explanatory role. The thought would seem to be that each instance of that schema would support explanations that rely on concrete facts that are themselves in need of explanation. If the schema incorporated mathematical facts, as Baker's does, then these could be used to explain these more concrete facts. But, at first glance, I seem to have no resources to similarly explain the explanatory features of these more particular explanations in terms of "deeper" features of reality that transcend the various domains represented by instances of my representation schema.

But I think this is too hasty. The features in virtue of which a target system is accurately represented by an instance of the representation schema are themselves extremely general, due to the wide range of target systems covered by the schema. In the case of the schema that generalizes the cicada representation, this requires only (1) two entities in the target systems with cycles of $x$ and $y$ in some unit, (2) finite upper and lower bounds on both $x$ and $y$, (3) that minimizing the frequency of the intersection of these cycles is optimal, and (4) that the mathematical inference patterns used in the explanation (a subset of RIC3) preserve truth under the physical interpretation RIC1.

The first three of these conditions are quite abstract, but distinctly physical; the explanatory cores and explanation schemata put forward by Baron and Baker incorporate similar conditions (for instance, Baker's very general claims about unit cycles). The final condition plays a role closer to that of the mathematical facts in Baker's and Baron's explanation schemata, but it too picks out very high level physical features shared by the relevant target systems. It requires each target system $T$ to be such that certain patterns of inference, which could in principle be spelled out in terms of purely physically interpreted mathematical language, preserve truth when interpreted in terms of $T$.

So it is not that the explanation schema that I put forward picks out only a very disjoint collection of topic-specific physical properties to serve as explanantia. Rather, it picks out extremely general properties common to the full range of target systems covered by the relevant representation schema, including target systems in very different domains. Arguably such properties are *better* placed to do the right explanatory work than bare mathematical facts, as they are straightforwardly instantiated in each target system accurately represented by an instance of the schema.

But at this point, Baker, Baron, and others might object that I have begged the question against them by insisting that the very general features picked out by mathematical representation schemata of the kind I have considered are not mathematical. Now, it is true that these features often do not seem to be expressible in purely non-mathematical language, and mathematical language has an important role to play in the account I've presented here. Any mathematical representation schema requires that certain patterns of inference, spelled out in terms of physically interpreted mathematical language,[19] preserve truth when interpreted in

[19] These are those inferences from physically interpreted premises to physically interpreted

terms of the target system of any instance of the schema. But the physical interpretation of this language plays a crucial role, and the result is something quite different from a bare mathematical fact (such as that all and only primes $p$ minimize $\gcd(p, q)$ for all $q < p$), which is given no such interpretation.

The general features I appeal to are highly abstract, mathematically represented features of physical systems, while mathematical facts concern features of abstract objects (or whatever we take the subject matter of pure mathematics to be). These features may coincide if we *already accept* certain versions of mathematical realism—namely, ante rem structuralism, Aristotelian realism, and some versions of neo-logicism—and it is not my aim to show that these views are false. Indeed, if one already accepts one of these views, there is still much to be gained by embracing my account of the metarepresentational role of mathematics, as it then provides a nuanced account of how mathematics contributes to explanatory generality an, in particular, how mathematical objects fit into such explanations.

But if we don't accept one of these views at the outset, then there seems to be no good reason to claim that the general features I appeal to are themselves mathematical. Both scope- and topic-generality are a matter of degree. And, all else equal, the more scope- or topic-generality an explanation possesses, the more abstract the properties in virtue of which a given system is correctly represented by some instance of the representation schema needed to articulate the explanation. To conclude that the general features I appeal to are themselves mathematical, we would need a further reason to think that these differences in degree should, after some threshold, yield a difference in kind.

---

conclusions that can be arrived at by applying only the inference patterns in (RIC3).

# 8 Conclusion

I have argued that we can capture the high degree of generality and depth possessed by explanations in which mathematics figures prominently without treating the mathematics as playing a distinctively explanatory role—that is, without treating mathematical facts as themselves among the explanantia of the explanation. But doing so requires us to move to a more nuanced account of the role of mathematics in scientific representations, one that recognizes a new, metarepresentational role for mathematics in exploring properties shared by collections of mathematical representations.

The result is an account of the role of mathematics in scientific explanations that takes a middle path between the explanationist approaches of the likes of Baker, Baron, and Colyvan and the representationalist approaches of the likes of Melia and Saatsi, incorporating the best aspects of both approaches. Like existing representationalist accounts, it takes the role of mathematics in scientific explanations to be representational at bottom. But it takes this role to be significantly richer than simply "indexing" physical facts, allowing mathematics to do some (metarepresentational) heavy lifting in scientific explanations. And while the account does not support a distinctively explanatory role for mathematics, as explanationists would have it, and so in particular does not support explanation-based indispensability arguments for mathematical platonism, it fills an important gap in both platonist and nominalist accounts of mathematical explanations in science by explaining how the use of mathematics contributes to explanatory generality.

# References

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

———. 2017. "Mathematics and Explanatory Generality." *Philosophia Mathematica* 25:194–209.

Baker, Alan, and Mark Colyvan. 2011. "Indexing and Mathematical Explanation." *Philosophia Mathematica* 19:323–34.

Baron, Sam. 2020. "Counterfactual Scheming." *Mind* 129:535–62.

Bueno, Otávio, and Mark Colyvan. 2011. "An Inferential Conception of the Application of Mathematics." *Noûs* 45:345–74.

Colyvan, Mark. 2002. "Mathematics and Aesthetic Considerations in Science." *Mind* 111:69–74.

———. 2013. "Road Work Ahead: Heavy Machinery on the Easy Road." *Mind* 121:1031–46.

Daly, Chris, and Simon Langford. 2009. "Mathematical Explanation and Indispensability Arguments." *Philosophical Quarterly* 59:641–58.

Goles, Eric, Oliver Schulz, and Mario Markus. 2001. "Prime Number Selection of Cycles in a Predator-Prey Model." *Complexity* 6:33–38.

Jansson, Lina, and Juha Saatsi. 2019. "Explanatory Abstractions." *British Journal for the Philosophy of Science* 70:817–44.

Kitcher, Philip. 1989. "Explanatory Unification and the Causal Structure of the World." In *Scientific Explanation,* edited by Kitcher Philip and Wesley C. Salmon, 410–505. Minneapolis: University of Minnesota Press.

Knowles, Robert, and Juha Saatsi. 2019. "Mathematics and Explanatory Generality: Nothing But Cognitive Salience." *Erkenntnis.* doi:https://doi.org/10.1007/s10670-019-00146-x.

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

Lyon, Aidan. 2012. "Mathematical Explanations of Empirical Facts, and Mathematical Realism." *Australasian Journal of Philosophy* 90:559–78.

Mancosu, Paolo. 2018. "Explanation in Mathematics." In *The Stanford Encyclopedia of Philosophy,* Summer 2018, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2018/entries/mathematics-explanation/.

McCullough-Benner, Colin. 2020. "Representing the World with Inconsistent Mathematics." *British Journal for the Philosophy of Science* 71:1331–58.

Melia, Joseph. 2000. "Weaseling Away the Indispensability Argument." *Mind* 109:458–79.

Nguyen, James, and Roman Frigg. 2017. "Mathematics is not the only language in the book of nature." *Synthese.* doi:https://doi.org/10.1007/s11229-017-1526-5.

Pincock, Christopher. 2012. *Mathematics and Scientific Representation.* Oxford: Oxford University Press.

Plebani, Matteo. 2016. “Nominalistic Content, Grounding, and Covering Generalizations: Reply to ‘Grounding and the Indispensability Argument’.” *Synthese* 193:549–58.

Reutlinger, Alexander, and Juha Saatsi, eds. 2018. *Explanation Beyond Causation: Philosophical Perspectives on Non-Causal Explanations.* Oxford: Oxford University Press.

Saatsi, Juha. 2016. “On the ‘Indispensable Explanatory Role’ of Mathematics.” *Mind* 125:1045–70.

———. 2011. “The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science.” *British Journal for the Philosophy of Science* 62:143–54.

Wakil, Samantha, and James Justus. 2017. “Mathematical Explanation and the Biological Optimization Fallacy.” *Philosophy of Science* 84:916–30.