185–213
10.1093/philmat/nkv022
Philosophia Mathematica
Advance Access publication August 17, 2015

Using Mathematics to Explain a Scientific Theory†

Michèle Friend* and Daniele Molinini**

*Philosophy Department, George Washington University, Washington, D.C. 20052, U.S.A.
E-mail: michele@gwu.edu
**Department of Philosophy, Communication and Visual Arts, University of Roma Tre, Rome, Italy
E-mail: dmolinini@gmail.com

ABSTRACT

We answer three questions: 1. Can we give a *wholly* mathematical explanation of a physical phenomenon? 2. Can we give a wholly mathematical explanation for a *whole physical theory*? 3. What is gained or lost in giving a wholly, or partially, mathematical explanation of a phenomenon or a scientific theory? To answer these questions we look at a project developed by Hajnal Andréka, Judit Madarász, István Németi and Gergely Székely. They, together with collaborators, present special relativity theory in a three-sorted first-order formal language.

1. INTRODUCTION

In this paper we discuss mathematical explanations in science (henceforth: MES). We answer three questions.

> *Question 1*: Can we give a *wholly* mathematical explanation of a physical phenomenon?

To answer this question we should define 'wholly mathematical' and 'explanation'.

†We should like to thank two anonymous reviewers for helpful suggestions for improvements, and the Hungarians who are working on this important project for their encouragement and friendship.

> *Question 2*: Can we give a wholly mathematical explanation for a *whole physical theory*?

Our answer is 'yes', and this implies a positive answer to the first question, for at least the phenomena of that theory. The interesting follow-on question is:

> *Question 3*: What is gained or lost in giving a wholly, or partially, mathematical explanation of a phenomenon or a scientific theory?

The three questions are new to the literature on mathematical explanations in science, but the last two are more radical. So in the next section, we shall write a little more about them.

In the 3rd section, we write about types of mathematical explanation in science, in particular we focus on (partly and wholly) mathematical explanations in science. In Section 4 we discuss an actual case of a mathematical explanation of a physical theory, namely special relativity theory. This will help us to answer the third question. But to give a thorough answer, we need more. A philosophical discussion of the case follows in Section 5. In the 6th section we consider some possible objections and in the last section we draw conclusions.[1]

2. THE NEW QUESTIONS

The use of mathematics in explanations in science is recognised in the literature [Steiner 1978; Baker 2005, Pincock 2007, Lyon and Colyvan 2008, Batterman 2010]. What is controversial is whether ultimately *all* such use can be dispensed with, without loss. We agree with others in the field that mathematical results are indispensable to *some* explanations of *some* physical phenomena, in the sense that if we nominalise the explanation, we have lost information, predictive power, precision, or we have a weaker understanding of the phenomena. We add nothing new debate directly, since it is possible, if difficult, to reject the project we consider in its entirety. Our contribution is indirect, and works through our answers to the second and third questions we raised in the introduction.

Our indirect contribution, nevertheless, strengthens the claim that is more common in the literature that mathematics is indispensable to some explanations in science. To situate our contribution in the literature, we think it is important to give sufficient background on MES to be able to indicate later what is new in our claims and to provide motivation for giving a new account of explanation in science. We give the background in Section 3.

Next, we start to develop our own account of MES for wholly mathematical explanations of physical phenomena and wholly mathematical explanations for a whole physical theory. We shall then be in a position to look at a particular

[1] Roughly, Molinini is most responsible for Section 3, and Friend is most responsible for Sections 4 and 6. We worked together on the introduction, Sections 2, 5 and on the conclusion.

case and propose it as an instance of a wholly mathematical explanation for a whole physical theory.

The particular case is a mathematical explanation for special relativity theory. We refer to the work of Hajnal Andréka, *et al.* [2002].[2] Henceforth, we shall refer to this and succeeding other work by the same authors and their collaborators as 'the Andréka-Németi project'.

In the Andréka-Németi project, axioms and definitions are written in a three-sorted first-order formal language,[3] and from these axioms we logically derive some of the standard 'laws of physics' and all of the typical phenomena and standard predications of special relativity theory. A similar project is undertaken by Thomas Benda in Taiwan [2008].[4] So, the Andréka-Németi group are not alone in developing such a project; in fact, it has a history.[5] Since the Andréka-Németi project is quite developed, and since the authors of this paper are more familiar with it than with Benda's project, we shall confine our attention to the Andréka-Németi project.

On the third question: the philosophical significance of introducing a distinction between mathematical explanations of whole scientific *theories* and mathematical explanations of particular empirical *phenomena* is that with a mathematical explanation of a scientific theory we engage very different and new questions about the theory and the phenomena. If our arguments for our answer to the second and third questions are convincing, then the standard epistemology of science is revised. Elaborating: in the *standard* view, explanations must

[2][Andréka *et al.*, 2002] only develops special relativity. However, in [Andréka *et al.*, 2012] the 2002 theory is extended to general relativity.

[3]We are picking up on the literature of *mathematical* explanations of science, where the distinction between logic and mathematics is not important. Thus, we continue to use the term 'mathematical language' since this fits better with the literature we are engaging. On the other hand, the Andréka-Németi group insist on the term 'logical foundations' for their relativity theory. The three-sorted language looks like a set theoretic language on the grounds that it contains 'membership'. Since some philosophers and mathematicians consider set theory to be mathematics and others consider it to be logic, we take the distinction as fuzzy (at best).

[4]An important difference between the projects is that Benda gives his axiomatic foundation in the language of Gödel-Bernays set theory, and, arguably, this presupposes more extravagant ontological commitments. But the argument about ontological commitment is tricky, since it depends on how much of the respective mathematical theories are presupposed, or drawn upon, by the axioms in the relativity theories.

[5]The idea of reconstructing the theory of special relativity through an axiomatic system is not a novelty. Work in this direction started in 1911 with Alfred Arthur Robb's book *Optical Geometry of Motion: A New View of the Theory of Relativity*. In his 1959 paper 'Axioms for relativistic kinematics with or without parity', P. Suppes issued the challenge to formalise special relativity in first-order logic. The challenge was taken up by Ax, Goldblatt and others. They successfully formalised the 'core' of special relativity. Ax presupposes Minkowski space-time. The Andréka-Németi group deepen the project by *deriving* Minkowski space-time (as did Goldblatt) and extend the project beyond the core to obtain the standard well-known results of special relativity, such as the result that clocks travelling fast appear to keep time more slowly when observed by a body travelling more slowly [Andréka *et al.*, 2007, Theorem 11.6, p. 631].

include observation or causal statements. In contrast, in the case we consider 'observation' is not empirical observation, but is replaced with 'co-ordinatising' (more later). 'Causation' is replaced with a formal representation of 'before and after' on the same trajectory of an inertial body and a calculation.[6] Thus, if this is to count as an explanation *at all*, then it is one that departs from the standard view since observation statements and causation are given a purely mathematical interpretation. Our account of a wholly mathematical explanation of a physical *theory* is one that answers a why question, is written in a formal mathematical language where we derive mathematical representations of what were previously thought of as 'laws of the science', and where the methodology comes from mathematical practice.

More interestingly, the philosophical significance of our conclusions reaches to the philosophy of mathematics, since it tells something of the nature of *mathematical* explanation, application, theory, model, and confirmation of mathematical theories and concepts. To draw this out, in Section 5 we shall further develop our account of mathematical explanation in science, to make it a pluralist account. Furthermore, and we shall return to this in the conclusions, the notion of a wholly mathematical explanation for a whole physical theory gives a new twist to the ontological dispute that is taking place around the enhanced indispensability argument for mathematical realism. This introduces the topics in our paper.

3. TYPES OF MATHEMATICAL EXPLANATIONS IN SCIENCE

The philosophical analyses of MES, (*e.g.*, those proposed in [Baker, 2005; Batterman, 2010; Lyon and Colyvan, 2008; Pincock, 2007; Steiner, 1978]) focus on the notion of mathematical explanation of phenomena. Specifically, they concentrate on how mathematical posits yield explanatory power when used to account for an empirical phenomenon, be it physical, biological, or even social. For instance, Aidan Lyon and Mark Colyvan [2008] have provided an example in which the regular or chaotic behaviour of the Hénon-Heiles system, *i.e.*, a particle moving in a two-dimensional potential called the 'Hénon-Heiles potential', is explained using the mathematics of phase-space theory. According to them, it is only through the mathematical resources of phase-space theory that we get the reasons why high- (or low-) energy Hénon-Heiles systems exhibit chaotic (or regular) motion. By appealing to the mathematics of phase spaces, together with the mathematical tools provided by the Poincaré map and the theory of differential equations, we are giving a partially *mathematical* explanation. Indeed, this explanation essentially depends on the mathematics employed and it ignores the causal nature of the phenomenon being analysed. *Without* the mathematics we could only know *that* the Hénon-Heiles systems exhibit

[6]The notion of causation becomes more interesting in the development of general relativity, where it is replaced with curvature of lines or dimensions, and with the direction of the time dimension.

high- and low-energy or chaotic or regular motion but we would *lack* a precise mathematical description of that motion or the ability to show precisely when, or under what circumstances, the systems exhibit the different behaviours. Therefore, we conclude with Lyon and Colyvan that mathematics is indispensable to the explanation of this particular physical phenomenon, since if we dispense with the mathematics we incur a net loss.

Another example of MES has been discussed in [Baker, 2005] and relates to evolutionary biology. The specific biological phenomenon chosen by Baker concerns the prime-numbered life-cycle length of an insect called 'periodical cicada', henceforth just 'cicada'. It turns out that the emergence periods of the periodical cicada are exactly 13 or 17 years. The rest of the time they stay dormant underground. 13 and 17 are prime numbers. In order to explain why it is evolutionarily advantageous for the cicadas to have such a prime-numbered dormant period, biologists use a number-theoretical result.[7] The result tells us that the particular emergent periods of 13 and 17 minimise overlap with *lower* life-periods of predators and *nearby* life-periods of different subspecies (since mating between subspecies would produce offspring that would not be coordinated with either subspecies, thus reducing mating opportunities). Number theory, or more precisely a theorem in number theory, is therefore *essential*, or 'indispensable', for the general *explanation* provided by biologists, which, of course, also makes use of specific ecological facts and general biological laws. Again, we have a case where the explanation of a scientific fact, and more precisely of a biological phenomenon, depends essentially on a mathematical fact. The mathematical fact is essential in the sense of giving us a completely precise description and *explanation*. We maintain that the two examples above are (at least) *partly mathematical* explanations of physical phenomena.[8]

We make our first definition.

> A *partial* MES is one where appeal to mathematics is essential to, or indispensable to, the explanation.

That is, if we were to omit the mathematical part of the explanation, our explanation would be noticeably poorer; we would have a general idea, but it would lack sufficient precision to make predictions or satisfy us as an explanation.

To perform such assessment, *i.e.*, to remove the mathematical claim and see whether the explanation loses its explanatory power, amounts to performing what Christopher Pincock calls the 'replacement test':

[7] The number-theoretic result is actually a consequence of two lemmas. A detailed treatment of the mathematics involved in the cicada explanation is given in [Baker, 2005]; so we shall not repeat it here.

[8] Arguably, the first example is a wholly mathematical explanation, but we shall see this in more detail later.

> [W]e may wonder if any explanations make [essential] use of a mathematical claim ... I propose a simple replacement test to answer this sort of question. Starting with an explanation that involves a mathematical concept or theorem, the test involves considering what happens if this concept or theorem is removed. If the resulting argument is no longer an explanation or else has less explanatory power, then I conclude that the mathematical concept or theorem is making an explanatory contribution to the original explanation. [2012, p. 204]

Let us run the test for the cicadas. We focus on the mathematical concept of prime number, which is the mathematics used in the cicada explanation, and therefore we want to replace 'every 13 or 17 years, where these are prime numbers' with something 'non-mathematical', say 'not every' year, or 'infrequently'. Substituting for 'every 13 or 17 years', we are told that cicadas emerge *not every* year, and therefore, this decreases the probability of hybridizing with other subspecies. The substitution is woefully imprecise and misleading. There are two things to say. First, even the imprecise explanation contains mathematics: the idea of appearing 'not every' year (or 'infrequently'), already presupposes being able to individuate and count years, to subtract a year, from a set of years. It is not very exciting mathematics, and it is not very sophisticated mathematics, but it is still mathematics, for all that. The 'quantity' 'not every' is vague. It is more in the spirit of the logical quantifier 'not all'. Similarly, any notion of 'lessening the probability', however vague, is still a mathematical notion. So we have not eliminated mathematics from our explanation. This first retort depends on a stringent account of nominalising. There are, of course, different accounts of nominalising; so, with a sufficiently liberal notion of nominalisation, the re-phrasing is successfully nominalised. So our retort is weak. Nevertheless, there is more to say.

The second thing to say is that the imprecise explanation is poorer, since we cannot predict when each colony will emerge under the nominalisation. Moreover, while we understand that not mating between subspecies is explanatory, we do not have a sense of how it is that the cicada is *so very* successful in avoiding hybridization, or, looking at the rate of non-success, why it is that this is *so very* rare. Moreover, we might be misled into making other hypotheses to explain the phenomena. For example, it is consistent with our vague explanation to hypothesise that one colony sends out a signal to all the others which prevents or deters the others from emerging. This hypothesis was, in fact put forward. But this is false. If we miss the mathematically pre-determined pattern, we *cannot* explain why the signal sending is unsuccessful on very rare occasions. Please forgive the mathematics: to be precise, after coinciding they would first coincide again in 13×13, 13×17 or in 17×17 years. The imprecise yet partly mathematical, explanation is poorer and Pincock's replacement test is passed.[9]

---

[9]Note that some philosophers have proposed a nominalist interpretation of Baker's cicada case (*cf.*, [Daly and Langford, 2009; Saatsi, 2011]). The goal of these philosophers is

What about Lyon and Colyvan's example of MES? One of the main results of their 2008 paper is that even if phase-space theory can be nominalised (something which is, *prima facie*, feasible), the resulting nominalisation would not have the explanatory power of the mathematical explanation. Again, although a nominalistic treatment can be proposed, it lacks the explanatory power that the mathematical treatment offers. We conclude, then, that at the present state of play, there are some explanations of some physical phenomena for which appeal to some mathematics is essential.

This is sufficient background in the literature to start to answer our first question. To do so, we shall give more definitions.

> An *explanation* of a phenomenon is an answer to a why question concerning that phenomenon.

What is a *mathematical* explanation of an empirical phenomenon?

> A [*partly*] mathematical explanation of a phenomenon is an answer to a why question that includes essential appeal to mathematical facts [Mancosu, 2011].

This is not enough, since answers to why questions can be quite unsatisfactory, even when they do appeal to mathematics. Therefore, it will be useful to add a means of *rating* the answers to the why questions. Philosophers have proposed rating explanations in terms of particular virtues that explanations seem to have, such as elegance, unification, simplicity, and so on. We propose that there should be (at least) three elements as necessary preconditions:

(i) the why question should be recognized by some good (accepted) members of the scientific community;[10]

---

to argue against mathematical explanations of physical phenomena by showing that cases such as the cicada example can be explained nominalistically and the real explanation lies in the physical facts and the causal relationships involved in the phenomenon. Baker and Colyvan [2011] have replied to such attacks and defended the genuine explanatory character of the cicada example. Since the reply, there has been no further challenge for this case. Indeed, Davide Rizza, for instance, has provided the most elaborate nominalistic reconstruction of the cicada case [2011]. But Rizza's criticisms of Baker's example concerns the ontological commitment to abstract (mathematical) entities, not the explanatory role played by mathematics in the cicada case. Indeed, he recognises that mathematics plays an explanatory role in the cicada case: 'It seems to me that, simple as it is, the Magicicada case shows how mathematics can play a relevant explanatory role in a non-ontological way' [2011, p. 113]. Our concern in this paper is not with ontology directly but rather with the quality of the explanation.

[10]This highlights the social aspect of science. For those who prefer to think of science as absolute, or who prefer to ignore the socio-politico-economic aspect, our formulation is still acceptable since we do not insist that the recognition of the question occurs at the time when the question is asked. Thus, it is not constitutive or a quality, but an indicator of quality. We might recognise it only in considerable hindsight. Moreover, it should be

(ii) the explanation should be phenomenologically satisfying to those same members of the scientific community, so that they do not immediately feel the need to ask a further why question (the explanation has come to an end, for now);[11]

(iii) the explanation should recover all (or most)[12] of the phenomena with the complete precision that was found in the previous explanation.

Explanations are rated ordinally. To rate the quality of one explanation over another we compare them and see how they differ. To rate a new explanation higher than an old one it must:

(iv) contribute something new to the scientific theory. The explanation brings new questions that would not have been asked or new predictions that would not have been made under the old explanation.

What is our account of a *wholly mathematical* explanation of a physical phenomenon?

> A *wholly* mathematical explanation in science, for a phenomenon, is one that answers a why question, and is written (*i.e.*, fully expressed, or in principle can be expressed) in a mathematical language, where the constants can be interpreted in a physical theory, but equally, they could be interpreted mathematically, or left uninterpreted.

Our first example was the Hénon-Heiles potential. This is, *arguably*, such an explanation. We can leave the mathematics of phase spaces uninterpreted (by physical observations or entities). However, it will *not* count as purely mathematical if we think that the mathematics of phase spaces is *essentially explained in terms of* empirical observation or causation (even as the pre-conditions for causation). We leave aside this debate since we can do better.

---

obvious that we cannot demand that the questions be recognised by all members of the scientific community. Of course, the other limit case of the 'scientific community' is that there is one scientist who recognises the question, only briefly. When this is the case, if the question is a good one, it will take longer to be recognised by that same scientist or other scientists.

[11] A weaker version is that we are willing to suspend our further why questions about that phenomenon. It is worth being careful on this point. As scientists, we might be drawn to ask other, deeper questions, or meta-questions about the explanation we are given. This is a type of acceptance of the explanation, not a rejection, and indicates fruitfulness of the explanation.

[12] In general, the explanation should recover all of the phenomena, but as we shall see in the next section, there are occasions when we want to know also when it is that we lose a class of phenomena.

In this paper we shall consider a wholly mathematical explanation of several *phenomena* of science, since the case we are considering is a mathematical explanation of a *whole physical theory*, and therefore, of each of the phenomena of the theory.[13]

> Our account of a wholly mathematical explanation of a physical *theory* is one that answers a why question, is written in a formal mathematical language where we derive mathematical representations of what were previously thought of as 'laws of the science', and where the methodology comes from mathematical practice.[14]

In our case, the presentation begins with axioms written in a formal first-order language. From the axioms we can derive formal representations of what are more standardly thought of as 'laws of the scientific theory' *as theorems* of the formal axiomatic system. For example, we derive that observers cannot travel faster than the speed of light. Note that this is already more precise than the (disputed) law that 'nothing travels faster than light'. We then make the calculations and geometrical transformations to derive each of the phenomena of the theory; where the phenomena are tested against observation. What the formal system gives us is a means of calculating, given initial observation data substituting for the sorts in the language, what will be the case at a later time. For example, if we are interested in the trajectory of a particular inertial body, and how its clock appears to us, and we are given the data that the inertial body is travelling at a *particular* speed and on a trajectory which is at a *particular* angle with our own, then we can find out what the time will be for that inertial body at a *particular* later time for us. Thus the constants in the original logical language *can* be interpreted as observations. In fact this is the intention of the Andréka-Németi group.

However, it is also important to note that the constants, and, *mutatis mutandis* the 'prediction', can also be left uninterpreted, or be interpreted mathematically, as a mathematical theorem divorced from any intention to apply it to physics. We can make mathematical models of the theory. This is crucial for the explanation to count as *wholly* mathematical. It follows from giving a *wholly* mathematical explanation that we *could* just take the mathematical explanation as an interesting piece of mathematics in its own right, without paying heed to the initial motivation of explaining a physical theory. This last option

[13]Our claim is at two levels simultaneously. At the object level, each phenomenon receives a wholly mathematical explanation; at the meta-level the whole set of phenomena (and the previous laws of the 'theory') receive a wholly mathematical explanation.

[14]We are not certain that we need this last conjunct, but it does fit the particular example we have in mind. In our test case, it is the mathematical methodology that brings questions that would not have been asked under a different explanation.

is possible, but is strained in the case of the Andréka-Németi project, since it is not *mathematically* natural, obvious, or interesting, at least *ab initio*.[15]

4. AN ACTUAL CASE: THE RELATIVITY THEORIES EXPRESSED IN A THREE-SORTED FORMAL FIRST-ORDER LANGUAGE

Let us turn to our case study: the Andréka-Németi project. Insofar as this project is a convincing presentation of a mathematical explanation for a physical theory, we have a positive answer to the first and second questions. Moreover, if the mathematical explanation gives us new insights into the physical theory, which we would not have appreciated or thought of in the absence of the mathematical explanation, then we have a net epistemological gain in exploring the mathematical explanation. This anticipates our answer to the third question.

In the actual case we have in mind we are presented with the axioms:

Ax1 $G = \mathbf{Eucl}(n, \mathbf{F})$.
Ax2 $Obs \cup Ph \subseteq Ib$.
Ax3 $(\forall h \in Ib)(\forall m \in Obs)\Big(tr_m(h) \in G\Big)$.
Ax4 $(\forall m \in Obs)\Big(tr_m(m) = \bar{t}\Big)$.
Ax5 $(\forall m \in Obs)(\forall l \in G)\Big(ang^2(l) < 1 \Rightarrow (\exists k \in Obs)\, l = tr_m(k)$ and $ang^2(l) = 1 \Rightarrow (\exists ph \in Ph)\, l = tr_m(ph)\Big)$.
Ax6 $(\forall k, m \in Obs)\Big(Rng(w_m) = Rng(w_k)\Big)$.
AxE $(\forall m \in Obs)(\forall ph \in Ph)\, v_m(ph) = 1$.

The axioms are copied from [Andréka *et al.*, 2002, p. 52]. Let us explain how to read the axioms to disperse any suspicion that there are essentially physical notions included in them. G is a set of lines. Lines are part of geometry. Axiom 1 says that the set of lines are (straight) lines in a Euclidean n-dimensional vector space over a field $\mathbf{F}$. All of these are mathematically definable notions in a first-order language.

Intuitively, and for the intended interpretation, in the meta-language, 'Obs' stand for 'observers'. 'Ph' is the set of photons. 'Ib' is inertial bodies. 'Observing', here, means a co-ordinating system. Observers do not 'see' in our (causal) sense, and they do not need photons to 'see', then to pick out a point in space. Observers only pick out (give the co-ordinates for) a point, a set of points, or a line (made up of points) in an n-dimensional vector space. More carefully,

[15] That is, if we were to present the mathematical theory or the family of mathematical theories to mathematicians without telling them what the intended interpretation is, we doubt that the mathematicians would be interested in the theory, since it looks like an arbitrary (mathematically *ad hoc*) 3-sorted formal theory. Moreover it is fairly standard; so it does not look likely to provide new and interesting mathematical insights.

they do not perform an *action* of 'picking out', instead, they are just the source (initial data/premises of an argument/starting place for the calculation). Coordinates in Euclidean space are mathematical entities, n-tuples, where n is the number of dimensions. Lines can be thought of as trajectories of an inertial body. But they can also be thought of as simply straight lines in an n-dimensional vector space. (To avoid tedium, we shall not repeat all of these niceties, but will take them as read. We return to more intuitive and familiar language.) Axiom 2 tells us about the inertial bodies. Observers and photons are inertial bodies. Note that we have not ruled out the possibility of there being other sorts of bodies. We have also not ruled out the possibility that photons should be observers. This will be shown in Axiom 5. The upshot will be that they therefore travel on straight Euclidean lines. This is correct for special relativity. In general relativity we also have accelerating bodies.

The term 'tr' stands for 'trajectory'. So, the trajectory of h according to m, that is $tr_m(h)$, is a straight line in a vector space according to Axiom 3. In Axiom 4, we learn that observers can observe themselves. That is, coordinatising is reflexive. Their 'self-observation' takes them along their own time axis $\bar{t}$. This makes sense if we remember that inertial bodies travel in straight Euclidean lines, at a constant speed. All they can 'observe' about themselves is the time — the ticking of a clock. Notice that no supposition is made about time having a direction. It is simply a singled-out axis for an observer. If we choose one of several possible axioms of orthogonality, we can use it to give spatial dimensions relative to an observer's time axis. The observer then 'determines' the n-dimensional space in which it travels.

It is Axiom 4 which is important for deducing the strange clock effects of special-relativity theory. Axiom 5 distinguishes inertial observers from photons by their velocity. The Andréka-Németi group consider the angle between l and the time axis, for $l \in \mathbf{Eucl}(n, \mathbf{F})$. They use the square of the tangent (instead of the tangent itself) of the angle, namely $ang^2(l)$, in order not to have to presuppose that the angles have real values. Similarly, they do not want to presuppose that points on a line must have real values, or that lines are best represented by the real line (of points). Points are n-tuples, where n is the number of dimensions. Thus, the field might only have a rational number of possible points. They leave open the possibility that lines and points could also have real values, and when they need this, they introduce an axiom allowing for values of all roots. More explicitly, for $l = \{r + a \cdot s : a \in \mathbf{F}\} \in \mathbf{Eucl}(n, \mathbf{F})$, they define $ang^2(l)$ in the following way:

$$
\begin{aligned}
ang^2(l) &= \frac{s_1^2 + s_2^2 + \cdots + s_{n-1}^2}{s_0^2} && \text{if } s_0 \neq 0, \text{ and} \\
ang^2(l) &= \infty && \text{if } s_0 = 0,
\end{aligned}
$$

where s_0 is the vertical (or *time*) component of vector s, and $s_1, s_2, \ldots, s_{n-1}$ its horizontal (or *space*) components.

Take the first conjunct of Axiom 5. For all observers, if the angle squared of a line, l, is less than 1 with respect to the trajectory of the observer, then the trajectory of the inertial body making the line is that of an(other) observer. The second conjunct is about photons. If the angle squared of the trajectory of the inertial body is equal to 1, then the body travelling on that line is defined to be a photon.[16] We do not even have a mysterious constant number featuring as the speed of light. Anything travelling slower than light is simply labelled 'observer' by definition. 'Photons travel at the speed of light' is wholly determined by means of the fixed square of the angle with respect to the time axis for an observer. So we distinguish observers from photons using a formal mathematical stipulation in an axiom of the theory.

It turns out that Axiom 6 does a lot of work in the derivations of some of the paradigmatic phenomena in special relativity theory. Because it is appealed to so frequently the Andréka-Németi group deploy a mathematical methodology: later in the 2002 text, they look at weaker versions of Axiom 6. Note this, because we shall return to it in the next section, since they do not treat mathematical axioms as laws in science. Regardless, this version of Axiom 6 tells us that for any two observers k and m, the range 'Rng' of the worldview of m is identical to that of k. That is, the two observers co-ordinate the same sets of points and lines. They might, of course, have different relations bearing between said points, since they might be travelling at different speeds from each other, or in different directions. The worldview of an observer is simply a set of points that they observe as they travel. So they are just sets of n-tuples, where n is the number of dimensions.

Axiom E is so called because it is 'Einstein's axiom'. We learn from it that the speed of photons, *i.e.*, of light, is fixed. It is fixed at 'the angle squared of the line of the photon' $= 1$ for all observers. All photons are observed as travelling at the fixed speed of light. Indeed, they form a light cone through their angle with the observer (who observes himself as travelling in time only). Essentially, all that happens in the derivations is that we derive what is observed relative to an observer. The observation is not a 'seeing'. It is simply a fixing of a set of co-ordinates which forms a straight line in a vector space. We can then immediately derive the fact that photons are not observers. What observers 'observe' (pick out the co-ordinates) of each other is calculated by transformations. The transformations are affine transformations for the most part. The others are Poincaré transformations, Lorentz transformations, or Galilean transformations, and these are just calculated in the usual and *purely mathematical* way.

The reason we worked through the axioms was to convince the reader that all that has been said by the axioms is mathematical. It can be interpreted, as we have suggested, in order to explain special relativity theory and to derive the phenomena of special relativity. Indeed this is the intention of the Andréka-Németi group.

[16] Axiom 5 'states that we have the tools for (performing) thought experiments: on any appropriate straight line [in the n-dimensional vector space over $\mathbf{F}$] we can assume there is an observer; and the same for photons' [Andréka *et al.*, 2002, p. 50].

5. PHILOSOPHICAL REMARKS ABOUT THE ACTUAL CASE

As we mentioned at the end of Section 3, it is in no way necessary to make the intended interpretation, since the explanation is wholly mathematical. Without the interpretation in the meta-language, we could happily follow the proofs without a physical interpretation, or with a quite different one.

The new theory of special relativity looks quite interesting in this sparse mathematical description. There are no accelerating bodies (no curvature in the trajectories), no energy, and no mass. Or more carefully, we are free to ignore this interpretation of the mathematics if we so choose, and we can still carry out the calculations and predictions of special relativity theory with complete precision. In the mathematical, or logical, theory, (what can be interpreted as) bodies do not collide, in the sense of changing the path of another body. If two inertial bodies meet they do no more than intersect each other in a Euclidean plane. All they 'do' is intersect. Thus, all we have is lines in fields in a Euclidean geometry. The Andréka-Németi group later add a new axiom to give a unique direction to time to make up Minkowski space. They only do this when they 'need' it to explain some phenomena 'better' in a 'less static' way.

We have mentioned 'adding axioms' three times: adding an axiom to give real values, modifications to Axiom 6, and now, adding an axiom to give a direction to time. We now address what this means. The list of axioms we gave above forms the basic system of axioms. The Andréka-Németi group call it BASAX. It is sufficient to derive all of the phenomena of special relativity theory. But the Andréka-Németi group working on this project are using a mathematical methodology. The axiom system given above is *not* complete for the *entire project* even if it meets the present needs of special relativity theory. The axioms above are treated as *hypotheses*. They are not thought to be literally true or physically revealing.

The axiomatisation of physical theories is not new and is not restricted to the relativity theories. Hilbert suggested axiomatising, especially mechanics, in 1900 [Truesdell, 1991, p. 6, n. 1]. However, the Andréka-Németi project is unlike previous axiomatisations of physical theories. Here, let us just point out two differences. One is that we are not presented with one set of axioms, but with many different non-equivalent sets of axioms, what the Andréka-Németi group call 'variants of special relativity' [Andréka *et al.*, 2012, p. 8]. The other is that the (sets of) axioms can stand alone, and can constitute the physical theory. In contrast, in previous so-called axiomatisations of physical theories we are presented with one set of axioms that is meant to give us the basis for calculations in the physical theory. The axioms are (maybe mistakenly presented as being) indispensably accompanied by a less formal theory. In the Andréka-Németi project, different sets of axioms are (maybe mistakenly) presented as being strictly dispensable. In their words: 'we want to obtain a formalized theory which contains its own "interpretation" ' [Andréka *et al.*, 2012, p. 8]. That is, ultimately, the mathematics can be used to interpret itself.

There are two ways of thinking about this, depending on how we want to individuate 'theories'. If we individuate theories by a set of axioms and rules of inference, then, the methodology of the entire Andréka-Németi project includes several theories, and there is not one theory of special relativity; there are several. For example, we can add axioms to BASAX or we can take some away. We can weaken or strengthen axioms. This is interesting to do, in order to extend our explanation. We shall give more details later.

In contrast, if we insist on the definite article when we say '*the* theory of special relativity', then we can say that *the* Andréka-Németi theory of special relativity is a family of mathematical 'systems', where the *family* is the theory, and the systems are individuated by sets of axioms together with rules of inference, a notion of model, and so on.

Regardless, what we are left with are still *mathematical* explanations either in the form of theories or in the form of families of mathematical systems. This is what we wanted to emphasise in this section. And these (this) explanation(s) invite(s) mathematical questions. We shall return to this when we discuss the methodology more philosophically.

Summarising in the language of the first way of thinking of the Andréka-Németi project: it gives a series of mathematical theories individuated by different combinations of axioms. The explanations (if they are really explanations) are mathematical explanations (as opposed to physical explanations). If the Andréka-Németi group are giving explanations, then we have answered questions one and two in the positive. It is then possible (since actual) to give wholly mathematical explanations for physical theories and, as a special case, wholly mathematical explanations for physical phenomena.

The more thorny question is: are they really *explanations*? According to our characterisation of explanation they are. They answer the why questions about the laws and the phenomena of special relativity using mathematical language. Remember our criteria for explanation.

(i) the why question should be recognised by some good (accepted) members of the scientific community.

In our case, the Andréka-Németi group form a community. They publish in top journals in the field. They have contacts and academics who follow their work with interest in Canada, the U.S.A., Brazil, England, the Netherlands, Belgium, Germany, Romania, Bulgaria, and Taiwan, amongst others.

(ii) the explanation should be phenomenologically satisfying to those same members of the scientific community, so that they do not immediately feel the need to ask a further why question (the explanation has come to an end, for now).

Adopting a strategy that is generally used in the debate on mathematical explanation (*cf.* [Baker, 2009; Mancosu, 2011]), we note that *scientific practice* provides evidence for the genuine character of the explanation. In other words, the testimony (part of the practice) of the intuitions of the practising scientists provides evidence that we have a genuine MES. In the Andréka-Németi project, the explicitly stated motivations are to find the (logical) reasons for the laws and phenomena in the relativity theories. They want to analyse the logical structure of relativity theories and find which most basic (*i.e.*, mathematical/logical) axioms are responsible for a certain theorem:

> In our approach, axiomatization is not the end of the story, but rather the beginning. Namely: axiomatizations of relativity are not ends in themselves (goals), instead, they are only tools. Our goals are to obtain simple, transparent, easy-to-communicate insights into the nature of relativity, to get a deeper understanding of relativity, to simplify it, to provide a foundation for it. Another aim is to make relativity theory accessible for many people (as fully as possible). Further, we intend to analyze the logical structure of the theory: which assumptions are responsible for which predictions; what happens if we weaken/fine-tune the assumptions, [we explore] what we could have done differently. We seek insights, a deeper understanding. We could call this approach 'reverse relativity' in analogy with 'reverse mathematics'. [Andréka *et al.*, 2007, p. 608][17]

In communicating their results, the Andréka-Németi group regard their explanations as really explanatory. For instance, they consider that their logical proof of why fast-moving clocks appear to slow down from the point of view of a slower-moving observer provides genuine (logical) explanations of this phenomenon [Székely, 2011]. Therefore, it is easy to see how in this case the testimonies of scientists support the claim that we are confronted with a genuine case of MES (unless we want to dismiss the members of the Andréka-Németi group as 'not scientists' but then what are they?).[18] Additionally, it should be noted that these scientists even consider these explanations to be more natural than the explanations provided within the standard formulation of relativity theories, and this is because of their logical character, and logic is considered to be a more basic tool of communication and understanding than carefully trained and developed (physical) intuitions.

(iii) The explanation should recover all (or most) of the phenomena with the complete precision that was found in the previous explanation.

Throughout, the Andréka-Németi group have been careful to be explaining the phenomena of special relativity. In some cases they reason counter-factually with respect to the phenomena, and thereby discover the strengths and limitations of some axioms. For example, they might find out that by weakening

[17] *Cf.* for more similar claims in the Introduction chapters of [Andréka *et al.*, 2002] and [Székely, 2009].

[18] The Andréka-Németi group think of themselves as logicians, but unlike many logicians they are not interested in developing logical systems for their own sake; they are interested in scientific theories and the relationship between logic and the scientific phenomena.

an axiom too much we lose some phenomena but not other phenomena. Sometimes this is their motivation for changing the axioms! They thereby discover axiomatic or mathematical limitations. Other reasons for changing axioms include: elegance, simplicity (sufficient axioms), naturalness (easier axioms), or maximising the explanation — saying exactly and only what is needed in order to capture some, or all, of the phenomena. This is a type of 'reverse-mathematics' motivation included in their methodology.

Furthermore, for the purposes of rating the explanation, we look to see if:

(iv) the new explanation contributes something new to the scientific theory, if the explanation brings new questions that would not have been asked or new predictions that would not have been made under the old explanation.

The traditional 'explanations' are poorer.[19] To illustrate this, consider the following example. In a standard textbook on space-time physics we have the following 'explanation' for the idea that no particle travels faster than light: 'No particle has ever been observed to travel faster than light. Therefore, a particle will always travel less than one meter of distance in one meter of light-travel time' [Taylor and Wheeler, 1966, p. 32]. Apart from the logical fallacy of inferring from lack of observation of x to x's never happening, we have a mere stipulation ('supported' by poor induction) that there are no particles traveling faster than light. Such an 'explanation' invites the question 'why?' and invites inquiry into the nature of 'observation', 'particle', 'travel', and 'light'. What does it mean to say that 'no *particle* has been *observed* to *travel faster than light*?' In contrast to this 'explanation', the Andréka-Németi group derive a theorem which tells us that there are no faster-than-light *observers* [Andréka *et al.*, 2012, p. 2]. They derive it from Axiom 5 very quickly, but that does not rule out the possibility that there should be faster-than-light particles. So this is only the beginning of their explanation. For, already in [2002] they show us a model for faster-than-light *particles* in a 2-dimensional setting! This gives us a type of mathematical limitation result. In another work, they demonstrate the

[19]It might be objected that our view of mathematical explanations is entirely in line with Hempel's account of scientific explanation and his idea of explanations as deductions. This is false. The differences between the explanations we are considering here and Hempel's are: (1) the Andréka-Németi group are using a series of mathematical theories, not just one; so they offer both proofs within a particular axiomatic theory and proofs of the connections between the theories; (2) few of the proofs are formal — but they can be made so; in fact there are formal proofs that check all of the informal proofs; (3) some proofs are semantic, coming from model theory, which is also not hypothetico-*deductive* at all; (4) the axioms are not hypotheses, in Hempel's sense. They are not black-box hypotheses waiting to be elucidated; they are written in the formal language of the theory or family of theories. They are not premises to a deductive argument. The explanations given by the Andréka-Németi group have (temporarily) stopped with the axioms, but are completed with the other sets of axioms. So, there is a difference, albeit a subtle one, between Hempel's hypotheses and the axioms of the Andréka-Németi group.

consistency of faster-than-light particles with a set of axioms which is powerful enough to derive all of the (other) paradigmatic and observed effects of relativity theory [Székely, 2012]. In this way they explore exactly what it would take logically to allow faster-than-light particles. Because of these further enquiries and their attending explanations, it seems fair to say that the Andréka-Németi explanation is much richer than the standard explanation. The Andréka-Németi project passes the Pincock replacement test.

In our quoted standard explanation, are we making a straw-man argument? No. Taylor and Wheeler are from MIT and Princeton, respectively. While the text we refer to for a traditional explanation is meant as a textbook for students, not for professional scientists in the field; their explanations end with physical observations and 'physical' constants (such as 'observation', 'particle', 'travel', and 'light'), not mathematical constants. And there is not much *further* explanation of these in the standard professional literature.

Our claim here is that when there is no further explanation we are left with the following reactions:

(i) we give up further search since we cannot 'understand' (any better);
(ii) we develop an 'intuition' which corresponds to the constants; or
(iii) we seek further explanation in the form of concepts we already understand.

Many students and less formally educated people fall in to (i). Most professional physicists fall into (ii). The Andréka-Németi group fall into (iii). In the absence of *already* having the required intuitions, when we are confronted with these possibilities we experience what we shall call a '*malaise*'. It is this malaise which motivates the Andréka-Németi project.

Let us illustrate the *malaise* with a story told by Németi. In a class on relativity theory attended by Németi, the professor explained the twin paradox to the students. The students were puzzled, wondered at this 'paradox', and generally experienced a sense of *malaise*. This is all we mean by '*malaise*' here. Németi then asked the professor for a better explanation. Instead of an explanation, he was told the following: continue with your courses on relativity theory. Write a Ph.D. thesis in relativity theory. Become a professor teaching relativity theory. Then if you are very fortunate, after a few years, you will understand the twin paradox.[20]

We do not think that the story is unrepresentative of relativity theory (as it is usually presented and taught).[21] We saw an example of a standard explanation for particles not travelling faster than light earlier. We interpret the story in the following way. The professor himself was unable to give a better explanation.

[20] The story was told in a conversation among Németi, Andréka, and Friend.

[21] We do not want to make a survey of the literature and text books. This is too tedious. But even Einstein states his principle in English, not in a formal language: that the laws of physics should behave in the same way for all particles and everywhere in the universe.

But he had 'gotten used to it'. He had followed (ii) in the above methodology which is standard in the practice of physics. Or, he observed in his colleagues that they had 'gotten used to it' and was waiting for the day when he would 'get used to it'. 'Getting used to' something is a matter of time and exposure, and either is a type of explanation or *substitutes for* explanation!

Take the first disjunct. If intuition, or a sense of familiarity *is* a type of explanation, then with the intuition, the *explanation* has come to an end, maybe a temporary end. The twin paradox is after all called a 'paradox'. It is supposed to be puzzling. It invites why questions. That is, it invites further explanation; so at best it is an incomplete explanation. It is interesting to observe that labelling it a 'paradox' seems to license an authority *not* to have a further explanation! This little story is about a *lack* of explanation in a perfectly robust scientific theory. So at least the Andréka-Németi explanation is more complete than the standard one, since it does not rely on physical intuitions, but rather, on logic and mathematics. Maybe these are intuitions too, but they are *arguably* more fundamental or more basic.

After all, where could we look for a better explanation than the one given in the story? Not to the laws of relativity theory, since they promptly lead us to the paradox, and leave some physical constants without further explanation, except implicitly through the other laws. Instead, we have to question the physical laws themselves, and ask for explanations of those. How can we do this? The answer turns on what we think is 'more primitive' or 'more basic' than a scientific law. The answer we (and many scientists) give is: mathematics and logic are more primitive.[22] But ultimately, of course, this can be disputed. In this case we have a draw.

Nevertheless it remains that the Andréka-Németi group are so satisfied with their explanations that they are now adding refinements, and suggesting further experiments, and discovering new results (some of which are not easy to test). That is, the explanations are fruitful. Moreover they are sufficiently satisfied that they extended the theory (or family of theories) of special relativity to that of general relativity. They are presently working on representing Newtonian kinematics and are just starting to look into using their methodology to explain quantum theory. (But they are cautious about this, since they recognise that it will be very difficult. The project might take several generations of scientists and logicians.)

There is one more argument supporting the claim that the Andréka-Németi project gives us genuine examples of MES. Apart from the practice, the genuine character of the explanations given within the Andréka-Németi project is supported by a pluralist account of mathematical explanation in science, as

[22] The evidence that we are not alone, or first, in thinking of mathematics and logic as more primitive is that the mathematical foundation of the relativity theories has a history. Axiomatisations of special relativity have been studied, among others, by [Ax, 1978; Benda, 2008; Goldblatt, 1987; Suppes, 1959] and the Andréka-Németi project.

proposed in [Molinini, 2011].[23] Molinini proposes to investigate MES using two different notions: *intellectual tools* and *conceptual resources*.[24] An intellectual tool is an ability to reason that is used in the practice of explaining a scientific fact, while a conceptual resource is a concept that permits the use of our intellectual tools in a particular situation [Molinini, 2011, p. 352]. There is a plurality of explanations in science. They differ from one another in what intellectual tools are used and in what conceptual resources are used. In the Andréka-Németi project, conceptual resources are provided by axioms (and different sets of axioms) written in the formal first-order language and give us mathematical concepts that allow us to analyse or see a physical situation in a certain way. We then use logical reasoning as a tool to reason over that situation.[25] Reasoning[26] in the standard practice of physics is not as strict as mathematical or logical reasoning.

Intellectual tools and conceptual resources vary from one community of investigators to the next, and the factors that influence these are: the subjective preferences and aptitudes of individual members and the historical context of the community. In Molinini's account, the explanations provided in the Andréka-Németi project are genuine MES. However, note that our argument here does not depend on our adopting Molinini's account of MES. It is only strengthened by, or fits best with, such an account.[27] If we have another pluralist account of explanation in science, then the Andréka-Németi project might well count as an explanation amongst others. If we required a single account of explanation in science, it is possible that the Andréka-Németi explanation would not count! But then we shift the burden of proof. The philosopher holding to such a single

[23] After all, as Juha Saatsi has observed, it might be thought that scientific practice does not provide evidence for genuine MES. Even if some scientists claim that we have a genuine MES, 'surely it is down to the philosopher of explanation to scrutinize their claims and set them right' [Saatsi, 2011, p. 153]. In the context of our example, however, it is worth noting that the intuitions coming from the group are not simply intuitions of working scientists, but also of logicians and philosophers, sometimes all embodied in the same person.

[24] Obviously, it would take more space to illustrate here how the account proposed by Molinini fits the case of the relativity-project theory. Nevertheless, such an analysis is provided in [Friend, 2014]; so we shall not reproduce it here but we shall offer only the general strategy.

[25] In passing, note that the mathematics used by the Andréka-Németi group is a perfect tool for reasoning. Indeed, it makes the relativity-theory project objective in the sense that it is subject to logical correction [Friend, 2014].

[26] 'Reasoning' should be thought of in contrast to making calculations, since these are rigorous and precise in physics.

[27] Although [Friend, 2014] gives the only available analysis of the explanations provided in the Andréka-Németi project according to Molinini's account, other accounts might be thought to accommodate these explanations as well. For instance, our impression is that the Andréka-Németi explanations could be accounted for in terms of the *inferential conception of the applicability of mathematics* advocated by [Bueno and Colyvan, 2011; Bueno and French, 2012]. However interesting, we shall not pursue this issue here, and we leave it for future work.

account of scientific explanation would have to say what the Andréka-Németi group are doing if it is not explaining, especially since they think of their project as exactly explaining an area of science.[28]

6. FURTHER DOUBTS: ARE THESE 'GENUINE' EXPLANATIONS?

We have slipped in the word 'genuine' without due precision. There could be another type of doubt. Some philosophers of science draw a very natural, Quinean sort of distinction between a *description* or *account* and an *explanation*.[29] The idea behind the distinction is to emphasise that a description or account is not 'genuine', even if one concedes that it could count as a poor explanation.

> A *description* falls short of a genuine explanation, since it does not tell us why the phenomenon is happening to this sort of physical object, and not another.

In other words, it does not fall strictly within the confines of the phenomena investigated by that science. The description might be *consistent with* the observations of that science, but it will be applicable to other phenomena in another science. In contrast,

> A genuine *explanation* for a physical phenomenon (or set of phenomena) uniquely focuses on those physical objects, and *cannot* be re-applied to some other phenomenon (a phenomenon supervening on another type of object, or set of objects, altogether).

Given this distinction, the argument goes, mathematics by itself cannot *explain* physical phenomena, it can only *describe* them since the mathematical theory can be re-applied to other phenomena with quite different objects (by re-interpreting the constants). Therefore, there can be no genuine explanations of physical phenomena which are wholly mathematical. Call this argument A. If you are not swayed by this argument then skip the rest of this section.

There are at least three counter-arguments. We start with the weakest one. Focus on the 'uniqueness of the description'. We use 'description' in order not to beg any questions, 'uniqueness' applies to the set of data, or the phenomena being explained. In the case of the Andréka-Németi project uniqueness is guaranteed when we have *both* the mathematical theory and the meta-theory which *interprets* the mathematics, *i.e.*, gives the application of the mathematics to

[28] These are interesting questions but not to make this treatment too long, we leave them aside for a future project.

[29] Private conversation between Bueno and Friend, June 2012, São Paulo (Brazil). We do not claim that this doubt originated with Bueno. Rather, it is a very natural doubt, and it was Bueno who drew Friend's attention to it.

the physical phenomena. See, for example, how we interpreted the constants of the axioms in Section 4. Otherwise we simply have a mathematical theory. The meta-language then gives us one of several possible applications of the theory. Thus, without the meta-language part of the explanation — without this particular interpretation in mind — there is no uniqueness, and it would be unlikely for someone exposed only to the mathematical theory to guess at the intended application (which was the impetus for developing the particular mathematical theory).

Under this counter-argument, it is the package: mathematical theory + meta-theoretical intuitive explanation which counts, then, as a *partly* mathematical explanation of the physical phenomena we observe in the relativity theories. In other words, it is because the mathematical theory is checked against the particular 'data', *i.e.*, the predictions of the original physical theory, that we know that the mathematical theory applies to the physics. And we note, *only as a corollary*, that the meta-theoretical intuitive explanation is strictly dispensable for understanding the *mathematics*. However, it is *in*dispensable for understanding the *physics*. This answer is a start, but it is not quite right, and misses a lot of subtleties.

The second counter-argument is less conciliatory, and attacks the distinction between description and explanation. In the case of physical sciences, we suspect that the distinction relies on the notion of causation supported by appropriate physical laws, and possibly particular views of causation. That is, an explanation is such, only if it has an indispensable, and irreducible, causal element (making the observations unique in the sense of the origin of each of the causes). The distinction between a description and an explanation begs the question against the very idea of a purely mathematical *explanation* of physical phenomena. This is because *any* mathematical theory can be re-applied elsewhere. Therefore, *a priori*, *i.e.*, in light of the distinction, there can be no (wholly) mathematical explanations of physical phenomena.

When confronted with a question-begging position, one way out is to offer an equally question-begging counter-position which comes to the opposite, or at least a different, conclusion. We present such a position for argumentative-strategic reasons: we want to show that there are two question-begging theories of explanation which come to quite different conclusions. We do not think that one is right and the other wrong. We think both are wrong. Here is the temporary strategic move to present a counter-question-begging position.

Purported 'explanations' of physical phenomena which stop at, or have as primitives, physical constants, intuitions, or purportedly irreducibly physical ideas expressed as physical laws are *never* proper explanations of why the phenomena occur. They only tell us *that* they occur.

Argument: an *explanation* of why something is the case has to reach deeper than just to point to the physical causal laws, and then derive from those laws the phenomena of the theory. The reason we have to reach deeper is that such a purported explanation still leaves us dissatisfied, or it relies on our having correct intuitions. It does not answer *why*; it answers *that*. A deeper and more satisfying explanation can be had only by looking at the underlying mathematics or logic of the theory. That is, the explanation should explain the physical laws too, by treating elements in the laws as logical constants, and it is only by looking at the underlying mathematics and logic that we can do this. Therefore, any 'explanation' worthy of being so called has to be mathematical or logical. Call this argument B.

Not only does argument B raise the standards very high for explanations in science, but, it is circular and begs the question; but so does A when we presuppose that explanations in science require causation. What should we conclude? It is a matter of taste and training as to which explanations we find more satisfying. For those who develop the required intuitions, reaction (ii), they will be satisfied with the existent explanations. Others will adopt reaction (i) or (iii). Nevertheless, the 'actual therefore possible' argument stands — especially if it is a matter of taste or ability to develop intuitions! That is, given our example, it is possible (if only recognised by those of similar tastes in explanation) for there to be a purely mathematical explanation for a physical theory and for physical phenomena.

From a question-begging argument, such as A or B, we should draw no direct conclusion. However, we can draw an indirect conclusion: some philosophers *insist* that 'explanations' of physical phenomena must include direct inescapable reference (guaranteeing uniqueness) via causation and laws to the physical phenomena, hence A. Other philosophers do not recognise the above distinction, and therefore admit the possibility of purely mathematical explanations for physical phenomena. For our case here, the latter philosophers will not require the conciliatory first argument, but will happily accept that our actual case counts as a purely mathematical genuine explanation of physical phenomena.

For the former philosophers that subscribe to argument A, are they warranted in their insistence? The situation is more subtle than we suggested in the first counter-argument (about the package mathematical theory + intuitive/causal/observation-laden meta-language interpretation that together ensure uniqueness of the description, *i.e.*, an explanation). We address the missed subtleties.

First, in some applications, the fit between the mathematical theory and the data might be unhappy — as it sometimes is in physics. It is usual for there to be some massaging of the raw mathematical theory to fit the data; at the very least this takes the form of corrections to the mathematical idealisations and at worst we have gerrymandering which makes no *mathematical* sense, such as with re-normalisation. (On re-normalisation techniques see [Steiner, 1992] and [Maddy, 1997, Ch. 6].) Thus, applying mathematics is not straightforward. Especially if it makes no mathematical sense, the massaged mathematical theory will not be re-applied elsewhere, since it will not count as a proper mathematical theory! So we have uniqueness, but for mathematically perverse reasons. This is not the case with the Andréka-Németi project, since the mathematical theories are all mathematically acceptable (internally consistent).

Second, in other cases where the mathematics is coherent, the adjustment, or gerrymandering of the theory will make it mathematically 'unnatural' or 'odd' or 'non-standard'. In all these cases, the re-application will be delayed, if not postponed indefinitely — and this for reasons to do with mathematical psychology — what mathematicians count as 'natural' or 'normal' or 'standard'. And the latter is influenced by mathematicians' *mathematics* and *logic* background, not based on their encounters with applications in physics. Thus, for reasons of mathematical 'unnaturalness' we will not find the mathematical theory re-applied in wholly other areas.

Prima facie, the Andréka-Németi project is like this. The mathematical theories are standard, but mathematically uninteresting or *ad hoc*. In particular, Axiom 5, distinguishing observers from photons by their velocity, or square of the angle with the dimensional axes, is mathematically *ad hoc*, even if it is physically elegant.[30] Therefore, it is unlikely that the mathematical theories will be re-applied to other areas of mathematics or to other scientific theories, at least not in the foreseeable future. But there is a more subtle point to add.

Third, our case is not of one mathematical theory but of several, whose interrelations with the physical theory are made explicit by specifying which axioms are needed for which phenomena. Moreover, the interrelations between the theories is also made explicit by saying which sets of axioms strictly imply which other sets, or which sets of axioms are equivalent to which other sets. Thus, this presentation of special relativity theory prompts some purely mathematical enquiry about limitative results of the theory and individual axioms. These are a logician's or a mathematician's questions, not a physicist's (although, of course an individual physicist is partly also a mathematician, and so might well ask these questions too, but he does so as a mathematician). In fact, the distinction between a mathematician and a physicist is not so easy to maintain in the Andréka-Németi project, and the blurring of the distinction plays well for our position. This concerns the methodology of the Andréka-Németi project. The methodology of tweaking axioms and proving the logical meta-relations between theories is not driven by purely mathematical concerns (since the mathematics is not *prima facie* mathematically interesting), but by the combination of the mathematics with the intended interpretation.

Fourth, the mathematical theories in our actual case are mathematical; and we have claimed that it would be unnatural to re-apply them elsewhere — either to other theories of science or to other theories of mathematics, at least as they stand. But this is oversimplified. If we think mathematical theories can be used to explain science, the next obvious steps are to use mathematical theories also to explain: cosmology theory, Newtonian mechanics, or quantum theory. These projects are on the agenda. Note, however, that it is highly unlikely that it will be the same suite of mathematical theories (individuated by sets of axioms and rules of inference and construction of models) which explain the other physical theories; cosmology theory requires non-standard interpretations of space and time (and therefore, different ones from that of the relativity theories), quantum

[30] Interestingly, because it is mathematically *ad hoc*, it invites mathematicians and logicians to change the axiom in some way — thereby deepening the explanation!

theory seems to require its own logic and some subtle means of distinguishing particle from wave.

Prima facie, we have several sciences, *i.e.*, special relativity, general relativity, electro-magnetic theory, Newtonian mechanics, quantum theory, various branches of chemistry and biology, and so on. The individuating and partitioning of these theories has a history. The history is driven by an interplay between institutions on the one hand and theory making, observations, and tests on the other hand. The relativity theories help us to explain phenomena that are far away and very fast. The observations focused on in Newtonian mechanics, which confirm Newtonian mechanics, concern 'medium-sized dry goods'. They are enough to predict the impact of two cars, to calculate the speed with which a medium-sized dry object will meet the ground when dropped from a particular height, *etc.*

It is not clear where the boundary is between the theories. In fact, this is one of the motivations for thinking that Newtonian mechanics is strictly false and that we should do everything in what was considered to be the realm of Newtonian mechanics by appeal to the relativity theories. But even if we do this, then consider the borderline between quantum theory and the relativity theories. Argument A pre-supposes that we have a definite set of observations we want to explain with our theory. The set of observations is not fixed. There are vague, or fuzzy, boundaries. So we will not explain *all and only* the phenomena of that theory with the theory. For example, quantum 'entanglements' might be 'explained' by appeal to superliminal particles (particles that travel faster than light). This is just a tenuous hypothesis, but it indicates a possible link between the theories.

Worse: we issue a challenge. For any set of phenomena, and for any theory that explains those phenomena and is intended to pick out only those phenomena, and not other phenomena, we believe it is possible to find another set of phenomena (extensionally different from the first set) that is also explained by the same theory. For example, neoclassical economic theory uses Newtonian mechanics as the central science of the theory — maybe metaphorically, but they are quite convinced by it, nevertheless. At the very least, if we allow elements of a mathematical model to count as objects and relations between those elements to count as phenomena, then insofar as we have a mathematical explanation, we can often make other mathematical models; so the 'theory' can be 'interpreted' by different models. This might be abuse of the language. Regardless, there is another way. We can relax the actual individuation of the 'sciences' and allow some quirky objects to enter the science, or introduce some Cambridge objects.

While we can play these argumentative games, there is an underlying serious point. We should not be too naïve about our observations, especially in relativity theory. Observation statements are interpreted. Phenomena are interpreted. Moreover, they can be interpreted in different (but still consistent) ways. So, all that the observations do is to fix some conditions for the interpretation of the constants or primitive concepts in the theory. There is no stand-alone set of phenomena, or observations that can only be interpreted in one way. Instead, there are packages consisting in observations + interpretations, and these are accompanied by a logic or mathematics. Therefore, with any explanatory theory, we start with sets of data, phenomena, or observations; we interpret them and build a theory; we might then build other theories and revisit the observations in the light of the different theories. We then notice that the 'observations' we started with appear a little different, depending on our interpretations and theories. For example, an 'event' in the Andréka-Németi rendition of special relativity is just a point in time/space. It is the intersection of two trajectories.

Given these four counter-arguments to argument A, we should draw some more general conclusions concerning the idea of a scientific explanation, especially when it takes the form of a mathematical theory, or a suite of such theories. We turn to more general comments in the conclusion.

7. CONCLUSION

We have drawn a new and important distinction between two different types of MES: wholly mathematical explanation of *phenomena* and mathematical explanation of scientific *theories*. The motivation for drawing this distinction comes from the fact that all the literature on MES focuses on individual phenomena; but in our analysis of an actual case, *i.e.*, the Andréka-Németi project, we find an explanation of a theory, or more carefully, of a set of phenomena that was (more traditionally) explained by the laws of special relativity. The explanation is written in a three-sorted first-order mathematical language. This can be observed from the axioms and definitions given in the 4th section. For further scrutiny of the claim, the reader is referred to the actual texts [Andréka *et al.*, 2002; Székely, 2009].

If we have a full first-order mathematical theory from which we can derive representations of all of the purported laws of special relativity as theorems of the new mathematical axiomatic theories, and we can also derive the phenomena of the theory, then we have answered in the positive the first and second questions with which we began. The more interesting question is the third: what are the advantages of giving a wholly mathematical explanation of a physical theory?

We now enumerate the gains. We (1) learn something more about explanations in science. Some explanations of scientific phenomena are mathematical because they follow from a mathematical explanation of the whole set of phenomena, but this is not the general case in science, in fact it is so far unique to the actual case we look at.[31] The cicada example proposed by Baker most clearly shows that we might have partly mathematical explanations of phenomena that do not *come from* a mathematical theory of that area of science. We do not have a mathematical formulation of evolutionary biology (and perhaps

[31] Note also that the work of Ax and Goldblatt, for example, only gave the 'core' of special relativity theory. They stopped short of deriving the typical phenomena. Moreover, as a result, since their mathematical basis is different from that used by the Andréka-Németi group, it is still not clear that they had all that they would have needed to derive all the phenomena.

it is inconceivable at present to give such a formulation without significant loss of information). Furthermore, evolutionary theory does not predict the prime-numbered emergence of cicadas, and this is why biologists *appeal to* number theory to *complete* their explanation. This example tells us about the taxonomy of MES. There is no clear relationship between, on the one hand, partly mathematical explanations of phenomena and wholly mathematical explanations of phenomena and, on the other hand, the relationship between explanations of theories and explanations of phenomena.

It might be objected that in the Andréka-Németi project we are not faced with *genuine mathematical* explanations. In the 5th and 6th Sections we have addressed this doubt, and we have provided some arguments to defend the genuine character of our example. In light of the several considerations we have offered, the explanation given within the Andréka-Németi project should be considered to be a genuine case of MES.

In general, the other epistemological gains are that we can derive new results about the physical and the mathematical theories. With this sort of explanation we are prompted to ask questions *we would not have asked* if we only had a 'physical' explanation. Mathematical explanations prompt mathematical and logical questions — questions about consistency of phenomena with axioms of the theory (this exploration tells us very explicitly what axioms explain what phenomena), and this, in turn, prompts questions about the interdependence of phenomena. And these questions are further explored by looking to the interdependence of the mathematical axioms or theories.

To enumerate the further gains, consider the structure of the explanations. The Andréka-Németi group start with a particular phenomenon; they explain it using a particular set of axioms and attending proofs of the phenomenon. These axioms are also sufficient to derive representations of what *were previously* considered to be laws of the physical theory, *as theorems* of the new mathematical theory. Or, we derive phenomena as mathematical theorems that can be understood (grasped) in terms of the 'laws' of the standard physical theory or in terms of the formal axiomatic theory. This is (2) a 'deepening' of our understanding. Furthermore, it is (3) more precise. We can then carry on the derivation to make general predictions. We can feed in particular initial data (from observations, or hypothetical observations) to make 'predictions' that we can test by measurement and observation. This is a type of interpolation within the theory; we confirm what we already know in the science. So far, the Andréka-Németi project has recovered all of the observed current data of special and general relativity theory. Even better: (4) in exploring and developing the explanations we also learn about our mathematical concepts. The gain is not only in 'deepening' our understanding of the relativity theories. We also (5) make the theory more accessible, since it is developed in terms of simpler or more primitive (logical) notions, but also (6) in making predictions in the scientific theory — which we might eventually test. If we did not think that faster-than-light objects or 'particles' are consistent with our theory, then we would never think to look for their effects, or take seriously the data that suggest that they do exist. The advantage of bringing the mathematical methodology to bear on the physics is in (7) counter-factual reasoning. For example, we might learn that an axiom is stronger than it needs to be to recover the data. Or, we might learn that if we subtract an axiom from a set of axioms we lose some data or observations. So the axiom is *necessary* for the completeness of the theory (completeness with respect to capturing all the phenomena associated with the science). Again, this is a logical or mathematical question, and it is suggested by this methodology very clearly, since we are free to make these 'mathematical experiments' with the axioms. Axioms are thought of as hypotheses not as physical laws. We also (8) explore what is consistent with the theories that explain all the phenomena. For example we learn that it is consistent with most of the theories (and with the known phenomena) that some objects travel faster than light.[32] We think that the Andréka-Németi project makes a significant contribution to physics.

Finally (9), the notion of wholly mathematical explanation for a whole physical theory gives a new twist to the ontological dispute that is taking place around the enhanced indispensability argument for mathematical realism.[33] So the project reaches into disputes in metaphysics. Here, we sketch at least two potential new issues that emerge from our analysis and that have a direct impact on the dispute between platonists and nominalists. Until now, the platonists endorsed the enhanced (or explanation-based) version of the indispensability argument to support their realism about mathematical entities. For instance, Baker and Colyvan have focused on the indispensable explanatory role of some mathematical object or of a piece of mathematics like a theorem (*cf.*, [Baker, 2009; Lyon and Colyvan, 2008]).

(I): Take one of our main results that it is a mathematical theory (individuated by sets of axioms) and a mathematical methodology (used to navigate between the sets of axioms, such as model theory) that together are playing an explanatory role in science. Philosophers who accept this result and who are platonists (about mathematical entities), can recast the indispensability argument and argue for the explanatory indispensability of a lot of mathematics. The first-order theories used by the Andréka-Németi group have links and ties to arguably most of mathematics; so under a suitable 'rounding out' of the mathematics we recover most of mathematics, maybe most immediately via Zermelo-Fraenkel set theory.[34] Moreover the Andréka-Németi group use several sets of axioms; so their methodology invites extending their theories to others in mathematics

[32] The Andréka-Németi group have recently worked out this result for 3 and 4 dimensions [Andréka *et al.*, 2014].

[33] In this article we are interested in explanations that essentially rely on mathematical results, not on the existence of mathematical entities. Nevertheless, because of the importance of the debate on mathematical explanation for the platonist/nominalist debate *via* the enhanced indispensability argument, and because of the potential contribution that our study might have in this area, we discuss this issue briefly here.

[34] The parts not included are the parts that have not (yet) been linked to the rest of mathematics; so we can imagine an isolated mathematician with little institutional training writing in his or her own formal language and developing a consistent or coherent 'mathematical' *faute de mieux* (in the absence of any better way of describing the activity)

instead of focusing on the explanatory indispensability of some mathematical objects or theorems. This option has not yet been explored in the literature but is it not *prima facie* uninteresting.

(II): Secondly, the test case that we have presented gives further support for the claim that mathematics explains physical facts and sometimes reaches further in terms of explanatory power (when compared to the traditional physical explanation). As we have seen, the Andréka-Németi group are able to explain mathematically phenomena that are poorly explained in the standard treatment of special relativity. This fact has a twofold effect on the dispute about the enhanced indispensability argument: (i) it reinforces the idea that there exist mathematical explanations of physical phenomena, thus giving support for the premise that 'mathematical objects play an indispensable explanatory role in science'; (ii) it makes a problem for the nominalists who want to rephrase some mathematical explanations that, as in the case of our example, are not obtainable or, at least have not been obtained, in purely physical terms. As we have seen, through the usual treatment of special relativity we are not able to explain some phenomena that the Andréka-Németi group explain mathematically. This is interesting for the philosophy of mathematics because *the types of questions asked* are different depending on whether we explain a science using causal laws and physical constants or pure mathematics. The tool we use, the type of explanation we give, influences the direction of further mathematical and scientific exploration. We hope that our efforts might give a fresh and stimulating impetus to the debate about mathematical explanation in science.

REFERENCES

Andréka, H., J.X. Madarász, and I. Németi [2002]: *On the Logical Structure of Relativity Theories*. E-book. Budapest: Alfréd Rényi Institute of Mathematics. With contributions from A. Andai, G. Sági, I. Sain, Cs. Tőke. http://www.math-inst.hu/pub/algebraic-logic/olsort.html, accessed August 2015.

——— [2007]: 'Logic of spacetime and relativity', in M. Aiello, I. Pratt-Hartmann, and J. van Benthem, eds, *Handbook of Spatial Logics*, pp. 607–711. Springer Verlag.

Andréka, H., J. Madarász, I. Németi, and G. Székely [2012]: 'A logic road from special relativity to general relativity', *Synthese* **186**, 633–649.

Andréka, H., J.X. Madarász, I. Németi, M. Stannett, and G. Székely [2014]: 'Faster than light motion does not imply time travel', *Classical and Quantum Gravity* **31**, 095005. DOI:10.1088/0264-9381/31/9/095005.

Ax, J. [1978]: 'The elementary foundations of space-time', *Foundations of Physics* **8**, 507–546.

Baker, A. [2005]: 'Are there genuine mathematical explanations of physical phenomena?', *Mind* **114**, 223–238.

——— [2009]: 'Mathematical explanation in science', *British Journal for the Philosophy of Science* **60**, 611–633.

'theory' that is later discovered by someone with more institutional training, and shows the relationship of that theory with other theories of mathematics.

Baker, A., and M. Colyvan [2011]: 'Indexing and mathematical explanation', *Philosophia Mathematica* (3) **19**, 323–334.
Batterman, R. [2010]: 'On the explanatory role of mathematics in empirical science', *British Journal for the Philosophy of Science* **61**, 1–25.
Benda, T. [2008]: 'A formal construction of the spacetime manifold', *Journal of Philosophical Logic* **37**, 441–478.
Bueno, O., and M. Colyvan [2011]: 'An inferential conception of the application of mathematics', *Noûs* **45**, 345–374.
Bueno, O., and S. French [2012]: 'Can mathematics explain physical phenomena?', *British Journal for the Philosophy of Science* **63**, 85–113.
Daly, C., and S. Langford [2009]: 'Mathematical explanation and indispensability arguments', *Philosophical Quarterly* **59**, 641–658.
Friend, M. [2014]: 'On the epistemological significance of the Hungarian project', *Synthese*. DOI 10.1007/s11229-014-0608-x:1–17.
Goldblatt, R. [1987]: *Orthogonality and Space-Time Geometry*. Springer Verlag.
Lyon, A., and M. Colyvan [2008]: 'The explanatory power of phase spaces', *Philosophia Mathematica* (3) **16**, 227–243.
Maddy, P. [1997]: *Naturalism in Mathematics*. Oxford University Press.
Mancosu, P. [2011]: 'Explanation in mathematics', in E.N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Summer 2011 edition.
Molinini, D. [2011]: *Towards a Pluralist Approach to Mathematical Explanation of Physical Phenomena*. Lille: Atelier national de reproduction des thèses.
Pincock, C. [2007]: 'A role for mathematics in the physical sciences', *Noûs* **41**, 253–275.
——— [2012]: *Mathematics and Scientific Representation*. Oxford University Press.
Rizza, D. [2011]: 'Magicicada, mathematical explanation and mathematical realism', *Erkenntnis* **74**, 101–114.
Robb, A.A. [1911]: *Optical Geometry of Motion : A New View of the Theory of Relativity*. Cambridge: W. Heffer.
Saatsi, J. [2011]: 'The enhanced indispensability argument: Representational versus explanatory role of mathematics in science', *British Journal for the Philosophy of Science* **62**, 143–154.
Steiner, M. [1978]: 'Mathematics, explanation and scientific knowledge', *Noûs* **12**, 17–28.
——— [1992]: 'Mathematical rigor in physics', in M. Detlefsen, ed., *Proof and Knowledge in Mathematics*, pp. 158–170. New York: Routledge.
Suppes, P. [1959]: 'Axioms for relativistic kinematics with or without parity', in L. Henkin, A. Tarski, and P. Suppes, eds, *Symposium on the Axiomatic Method with Special Reference to Geometry and Physics*, pp. 291–307. Amsterdam: North-Holland.
Székely, G. [2009]: *First-Order Logic Investigation of Relativity Theory with an Emphasis on Accelerated Observers*. PhD Thesis, Eötvös Loránd University.
——— [2011]: 'On why-questions in physics', in A. Máté, M. Rédei, and F. Stadler, eds, *Der Wiener Kreis in Ungarn / The Vienna Circle in Hungary*, pp. 181–189. Veröffentlichungen des Instituts Wiener Kreis. Vol. 16. Springer.
——— [2012]: 'The existence of superluminal particles is consistent with the kinematics of Einstein's special theory of relativity', *ArXiv e-prints*.
Taylor, E.F., and J.A. Wheeler [1966]: *Spacetime Physics*. San Francisco: W.H. Freeman.
Truesdell, C. [1991]: *A First Course in Rational Continuum Mechanics*. Vol. 1. 2nd ed. Pure and Applied Mathematics; 17. Boston: Academic Press.