231–248
10.1093/philmat/nku003
Philosophia Mathematica
Advance Access publication March 9, 2014

Mechanistic Explanation and Explanatory Proofs in Mathematics†

Joachim Frans* and Erik Weber**

*Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, B-1050 Brussels, Belgium.
E-mail: Joachim.Frans@vub.ac.be
**Centre for Logic and Philosophy of Science, Ghent University, B-9000 Gent, Belgium.
E-mail: Erik.Weber@ugent.be

ABSTRACT

Although there is a consensus among philosophers of mathematics and mathematicians that mathematical explanations exist, only a few authors have proposed accounts of explanation in mathematics. These accounts fit into the unificationist or top-down approach to explanation. We argue that these models can be complemented by a bottom-up approach to explanation in mathematics. We introduce the mechanistic model of explanation in science and discuss the possibility of using this model in mathematics, arguing that using it does not presuppose a Platonist view of mathematics and allows one to gain insight into why a theorem is true by answering what-if-things-had-been-different questions.

1. INTRODUCTION

There is an increasing consensus among philosophers of mathematics and mathematicians that mathematical explanations exist. Mancosu [2008] draws attention to two different senses of mathematical explanation: mathematical explanation in the natural or social sciences on the one hand, and explanation within mathematics itself on the other hand. In this paper we will only discuss the latter case. Furthermore, the scope of this paper is limited to explanatory proofs in mathematics. We do not address the explanatory power of mathematical techniques that go beyond traditional proof, for example the use of diagrams.

Philosophers such as Mancosu [2001] have pointed to the topic of explanation in philosophy of mathematics, tracing it back to Aristotle and the community of mathematicians. The central idea is that mathematical activity is not merely driven

†The research for this paper was supported by the Fund for Scientific Research – Flanders (FWO) through research project no. G.0031.09. The authors thank Liesbeth De Mol, Laszlo Kosolosky, Bert Leuridan, Liesbeth Martens, Jean Paul Van Bendegem, Bart Van Kerkhove, and Pieter Van Nuffel for comments on draft versions of this paper.

by justificatory aims such as the collection of mathematical truths. In many cases mathematicians will search for alternative proofs of known results in order to find a better explanation of the mathematical facts. Mancosu [2008, p. 142] refers to the mathematician Mordell, who describes this phenomenon:

> Even when a proof has been mastered, there may be a feeling of dissatisfaction with it, though it may be strictly logical and convincing; such as, for example, the proof of a proposition in Euclid. The reader may feel that something is missing. The argument may have been presented in such a way as to throw no light on the why and wherefore of the procedure or on the origin of the proof or why it succeeds. [Mordell, 1959, p. 11]

The underlying idea is that mathematical proofs can do more than establish the truth of a mathematical claim. While all proofs of theorem p show that p is true, some proofs also reveal why p is true. In contrast with the extensive discussion of scientific explanation, only a few authors have proposed accounts of mathematical explanation.

In philosophy of science, the goal of explicating what explanations are has led to two traditions: the causal tradition and the unificationist tradition. Salmon [1989] describes this distinction as bottom-up and top-down explanation. The top-down approach subsumes the explanandum under general principles, aiming at unifying several phenomena under these same general principles. With respect to explanation in mathematics, all theories developed up till now are top-down approaches, leaving the bottom-up approach to mathematical explanation — if any — untouched. Starting from a geometrical proof, we will argue that the bottom-up approach is also fruitful in mathematics.

In Section 2 we show that the seminal theory of Steiner [1978] and the improved version of it in Weber and Verhoeven [2002] fit into the top-down approach. This section serves as background against which we develop a bottom-up account of explanation in mathematics. This account is meant to be complementary to the top-down approach; they are not rivals. We develop this account in three steps: first we introduce the mechanistic model of explanation (Section 3); then we give an example (section 4) and discuss how this model could be used in mathematics (Section 5). Then we discuss the pros and cons of applying this model to mathematical explanation (potential objections in Section 6, advantages in Section 7).

2. THE TOP-DOWN APPROACH IN THE PHILOSOPHICAL ANALYSIS OF MATHEMATICAL EXPLANATION

2.1. Steiner's Theory and Its Refinement by Weber and Verhoeven

Steiner [1978, p. 143] uses the concept of *characterizing property* to draw a distinction between explanatory and non-explanatory proofs. A characterizing property is a property unique to a given entity or structure within a *family* or domain of such entities or structures. The concept of a family is left undefined. According to Steiner, an explanatory proof always makes reference to a characterizing property of an entity or structure mentioned in the theorem. Furthermore, it must be evident that the result depends on the property (if we substitute for the entity another entity in the family which does not have the property, the proof fails to go through) and that by suitably 'deforming' the proof while holding the 'proof-idea' constant, we can get a proof of a related theorem. Though many of Steiner's concepts (family, deformation, proof-idea) are vague, we can construct examples which beyond any doubt would classify as explanatory proofs by his criterion. Take the following proof of the Pythagorean Theorem:

Proof T1

(1)	For every triangle ABC: $c^2 = a^2 + b^2 - 2ab\cos(a, b)$	Prem
(2)	For every angle θ: $\cos\theta = 0$ if $\theta = 90^\circ$	Prem
(3)	For every right-angled triangle ABC with hypotenuse c: $(a, b) = 90^\circ$	Prem
(4)	For every right-angled triangle ABC with hypotenuse c: $c^2 = a^2 + b^2$	1, 2, 3

This proof makes reference to characterizing properties of right-angled triangles with hypotenuse c, namely that $(a, b) = 90^\circ$. It is evident that the proof fails to go through if another kind of triangle is considered, since (3) is false for all other types of triangles. Furthermore, we can easily imagine similar proofs of related theorems. *E.g.*, obtuse triangles contain exactly one angle $\theta > 90^\circ$. Because $\cos\theta < 0$ if $90^\circ < \theta < 180^\circ$, we can derive that for all obtuse triangles, $c^2 > a^2 + b^2$.

Proof T2

(1)	For every triangle ABC: $c^2 = a^2 + b^2 - 2ab\cos(a, b)$	Prem
(2)	For every angle θ: $-1 < \cos\theta < 0$ if $90^\circ < \theta < 180^\circ$	Prem
(3)	For every obtuse-angled triangle ABC with obtuse angle in c: $90^\circ < (a, b) < 180^\circ$	Prem
(4)	For every obtuse-angled triangle ABC with hypotenuse c: $c^2 > a^2 + b^2$	1, 2, 3

This proof makes reference to the characterizing property of obtuse-angled triangles with obtuse angle at C, namely that $90^\circ < (a, b) < 180^\circ$. It can also be 'deformed' back to proof T1. The couple (T1,T2) answers the question "Why does it hold for right-angled triangles ABC with hypotenuse c that $c^2 = a^2 + b^2$, while for obtuse-angled triangles ABC with obtuse angle at C it holds that $c^2 > a^2 + b^2$?".

Both proofs are individually explanatory according to Steiner's criterion, though their explanatory power depends on the possibility of 'deforming' the proofs, or, in practice, on the existence of such 'deformed proofs and the related theorems they prove. For this reason Weber and Verhoeven [2002] argue that Steiner's account in fact sees mathematical explanations as answers to the question: "Why do mathematical objects of class X have property Q, while those of class Y have property Q'?". Questions of this form have to be answered by means of a couple of proofs; one individual proof is not enough. The proofs in the couple must use the same general theorem (*cf.* premise (1) in our proofs) and must use the characteristic property of respectively class X and class Y). Once we take this step, it is easy to see that Steiner fits within the unificationist tradition. We discuss the issue below.

2.2. Explanation as Unification

One of the generally accepted aims of explanation in the natural and social sciences is unification. Unifying events consist in showing that two or more different events are instances of the same (set of) law(s) of nature. As an example, assume that we have observed the following:

> Pendulum a had a period in the interval $[1.99s, 2.02s]$.
> Pendulum b had a period in the interval $[2.44s, 2.47s]$.

These events can be unified by deriving them from the pendulum law. The first derivation is:

(1) For all pendulums[1] $P = 2\pi\sqrt{L}/g$,
(2) a has a length in the interval $[0.99m, 1.01m]$.
(3) All pendulums that have a length in the interval $[0.99m, 1.01m]$, have a period in the interval $[1.99s, 2.02s]$.
(4) Pendulum a has a period in the interval $[1.99s, 2.02s]$.

(1) and (2) are premises. (3) is derived from (1), the explanandum (4) is derived from (2) and (3). The second derivation is

(1) For all pendulums $P = 2\pi\sqrt{L}/g$,
(2′) b has a length in the interval $[1.49m, 1.51m]$.
(3′) All pendulums that have a length in the interval $[1.49m, 1.51m]$, have a period in the interval $[2.44s, 2.47s]$.
(4′) Pendulum b has a period in the interval $[2.44s, 2.47s]$.

Note that the two derivations in the case of the pendulum have the same structure, use the same law, and differ only in the 'characterizing property' of each pendulum. Philip Kitcher, one of the developers of the explanatory unification approach, defends the idea that unification can cover mathematical explanations as well:

> For even in areas of investigation where causal concepts do not apply — such as mathematics — we can make sense of the view that there are patterns of derivation that can be applied again and again to generate a variety of conclusions [Kitcher, 1989, p. 437]

It is clear that proofs that give answers to the question "Why do mathematical objects of class X have property Q, while those of class Y have property Q'?" and satisfy the criteria of explanation discussed above, unify mathematical facts in a way similar to the way two physical explanations unify the two physical facts.

2.3. Link with Hempel's Covering-Law Model

One of the central tenets of the covering-law model of explanations as it was developed by Carl Hempel [1965] is that empirical regularities have to be explained by subsuming

[1] This equation is a simplification which only holds if the angular displacement is small enough that the small-angle approximation is true. [Serway, 2008, p. 432]

them under other laws (*i.e.*, by showing that the explanandum regularity could have been expected given the law(s) in the explanans). Here is an example of what Hempel had in mind:

All waves reflect	(Covering law)
All sounds are waves	(Auxiliary hypothesis)
All sounds reflect	(Explanandum)

It is clear that Steiner's account fits into this covering law idea: explanatory proofs use an overarching, more general theorem than the one that is to be explained. As we will see immediately, Hempel's model has been criticized and an alternative has been developed: the mechanistic account of explanation of capacities. The rest of this paper is an investigation of the potential of this account (which is certainly very valuable for the analysis of explanations in many empirical sciences) for the analysis of mathematical explanation.

3. MECHANISTIC EXPLANATION OF CAPACITIES

3.1. Introduction

Though probably most philosophers of explanation realized that covering-law explanations of regularities are difficult to find outside physics, an alternative was not developed until the end of the twentieth century. This is how William Bechtel and Adele Abrahamsen describe the situation:

> The received view of scientific explanation in philosophy (the deductive-nomological or D-N model) holds that to explain a phenomenon is to subsume it under a law. However, most actual explanations in the life sciences do not appeal to laws specified in the D-N model. [2005, pp. 421–422]

The life sciences include biology (cell biology, genetics, ...) but also, *e.g.*, neuroscience. Bechtel and Abrahamsen claim that the discrepancy is due to a focus on physics:

> Given the ubiquity of references to mechanism in biology, and sparseness of reference to laws, it is a curious fact that mechanistic explanation was mostly neglected in the literature of 20th century philosophy of science. This was due both to the emphasis placed on physics and to the way in which explanation in physics was construed. [2005, p. 423]

In this section we discuss the alternative that was developed from 2000 onwards: the mechanistic model of explanation. We focus on explanations of capacities. A capacity can be ascribed to a system or to a class of systems. By a capacity we mean the ability of a system to react in a specific way (*i.e.*, to produce a specific output) given certain inputs. So if we claim that a system has a capacity, we claim that there is a regularity (a systematic connection between inputs and outputs) in its behavior.

3.2. The Mechanistic Model of Explanation

Many generalizations in biology, cognitive science, psychology, and engineering sciences ascribe capacities to classes of systems. And one of the main tasks which scientists in these disciplines set themselves (besides establishing the regularities) is explaining them. So these scientists ask why-questions about capacities ascribed to classes of systems. Here are some examples:

> Why do plants and bacteria have the capacity to convert carbon dioxide into organic compounds?
> Why do humans have the capacity to see depth?

Starting with the seminal paper [2000] of Machamer, Darden, and Craver, many philosophers have claimed that these capacities should be explained by means of a mechanistic explanation, which does not look like a covering-law explanation at all. So mechanistic explanations are put forward as an alternative to covering-law explanations. They are defined as follows:

> A *mechanistic explanation* of a capacity is a description of the underlying mechanism.

We define mechanisms as follows:

> A *mechanism* is a collection of entities and activities that are organized such that they realize the capacity.

This definition includes the three key terms which mechanists use: entities, activities, and organization. A description of a mechanism is usually called a model of the mechanism. The core idea of the mechanists is that, in order to have explanatory value, the model has to describe the mechanism in terms of its entities, its activities, and the way these entities and activities are organized. Before we present an example, some characteristic quotes. Bechtel and Abrahamsen write:

> A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena. [2005, p. 423]

Carl Craver writes:

> [M]echanisms are entities and activities organized such that they exhibit the explanandum phenomenon. [2007, p. 6, italics removed]

These quotes show that mechanists have no unique way of defining what a mechanism is. However, there is a common core idea and our definition captures this core idea.

3.3. An Example

Our example is about an electrical circuit, which we label E. Assume that everything inside the large rectangle is contained in an opaque box, so that only the three input wires and two output wires are visible. Assume also that we can somehow measure

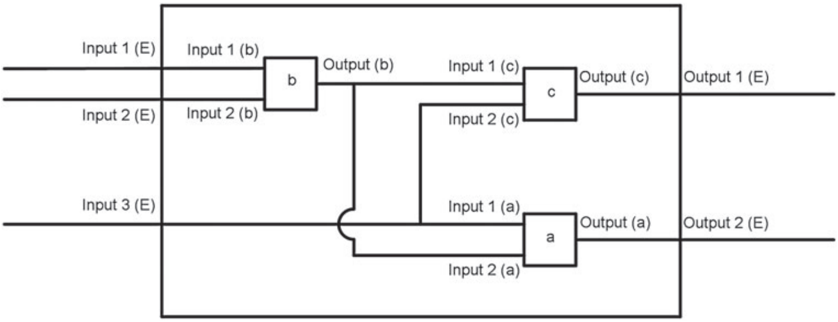

Fig. 1. Electrical circuit E.

whether these wires are charged or not. Then an experiment can be performed to see whether there is a regularity connecting the states of the input wires with the states of the output wires. Suppose that such an experiment yields the following claim:

If $\text{input}_1(\text{E}) = 1$, $\text{input}_2(\text{E}) = 0$, and $\text{input}_3(\text{E}) = 1$,
then $\text{output}_1(\text{E}) = 0$ and $\text{output}_2(\text{E}) = 1$.

"$\text{Input}_1(\text{E}) = 1$" is shorthand for "The first input wire of E is charged", $\text{input}_2(\text{E}) = 0$ for "The second input wire of E is not charged", *etc*. This claim ascribes a capacity (the capacity to produce a specific output given a specific input) to the system E. In order to give a mechanistic explanation of this capacity, we have to open the box. If we do this, we discover that E contains three binary gates (a, b, and c) and several wires. This is what we find out about the entities of the mechanism. Each of the gates can be taken out of the circuit; so we can investigate their individual behaviour. Assume that such tests give the following results:

a is an AND gate.
b is an XOR gate.
c is an XOR gate.

An AND gate has output 1 if and only if both inputs are 1. And XOR gate (exclusive OR) has output 1 if and only if the values of the inputs are different. These are claims about the activities of the entities. Finally, we can describe the organisation of the circuit:

The circuit is wired such that:

$\text{output}(b) = \text{input}_2(a)$.
$\text{output}(b) = \text{input}_1(c)$.
$\text{input}_1(\text{E}) = \text{input}_1(b)$.
$\text{input}_2(\text{E}) = \text{input}_2(b)$.
$\text{input}_3(\text{E}) = \text{input}_1(\text{a}) = \text{input}_2(c)$.
$\text{output}_1(\text{E}) = \text{output}(c)$.
$\text{output}_2(\text{E}) = \text{output}(a)$.

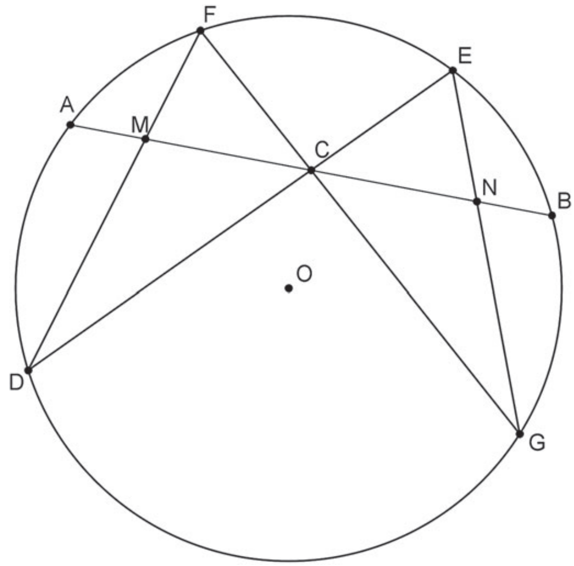

Fig. 2. Butterfly theorem.

4. A MATHEMATICAL EXAMPLE

4.1. The Butterfly Theorem

In this section we will discuss the possibility of using the mechanistic model of explanation in mathematics. The butterfly theorem is a result in Euclidean geometry. Let C be the midpoint of a chord AB of a circle, through which two other chords FG and ED are drawn; FD cuts AB at M and EG cuts AB at N. Then C is the midpoint of MN. The name is evidently linked with the apparent image of a butterfly in the configuration of the problem. Numerous proofs have been developed, varying in difficulty and length. Bankoff [1987] traces the presumably earliest proof of this problem back to 1815. We will use a proof that involves the identification of entities that are part of the original figure. We will argue that such a proof is explanatory as well.

4.2. Proof

(1)	$\angle FDE$ and $\angle EGF$ are equal	Inscribed angles
	$\angle DFG$ and $\angle DEG$ are equal	Inscribed angles
(2)	$\triangle FCD$ and $\triangle ECG$ are similar	(1: Property similar triangles)
(3)	$\frac{FD}{FC} = \frac{EG}{EC}$	(2: Property similar triangles)
(4)	Construct point H on DF such that OH is perpendicular to DF, and construct point J on EG such that OJ is perpendicular to EG	
(5)	$FD = 2.FH$	(4: Property of a chord)
	$EG = 2.EJ$	(4: Property of a chord)

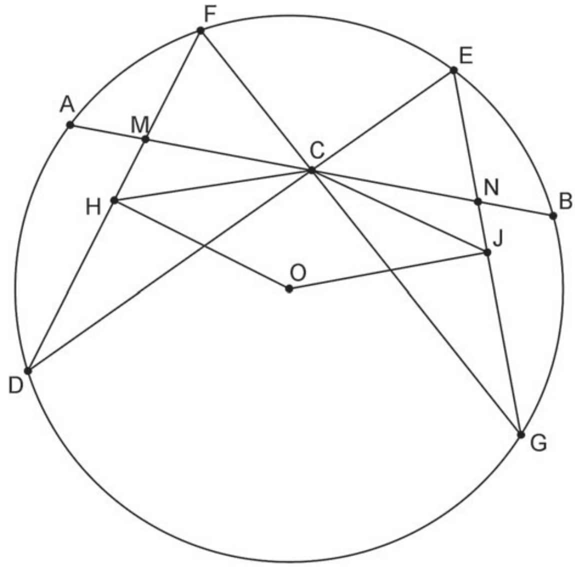

Fig. 3. Butterfly theorem: Intermediate steps.

(6)	$\frac{FH}{FC} = \frac{EJ}{EC}$	(3,5: Substitution)
(7)	$\triangle FCH$ and $\triangle ECJ$ are similar	(1, 6: Property of similar triangles)
(8)	$\angle EJC$ and $\angle FHC$ are equal	(Property corresponding angles)
(9)	$\square OCMH$ is cyclic	(4: $\angle FHO + \angle ACO = 180^\circ$)
	$\square OCNJ$ is cyclic	(4: $\angle EJO + \angle BCO = 180^\circ$)[2]
(10)	$\angle MHC$ and $\angle MOC$ are equal	(9: Inscribed angles)
	$\angle CON$ and $\angle CJN$ are equal	(9: Inscribed angles)
(11)	$\angle MOC$ and $\angle CON$ are equal	(8, 10: Substitution)
(12)	$\triangle OCM$ and $\triangle OCN$ are equal	(10, 11: Property equal triangles)
(13)	C is the midpoint of MN	(12: Property equal triangles)

5. ANALYSIS OF THE PROOF

5.1. Explanatory Proof

In Section 3, mechanistic explanation was defined in terms of entities, activities, and organization. We will demonstrate how the proof of the butterfly theorem is

[2]We thank the anonymous referee for pointing out that if point H and point M coincide, this step is not possible. It is nevertheless still possible to come to the conclusion that $\triangle OCM$ and $\triangle OCN$ are equal. In the case that H and M coincide, we know that $\angle FMC$ and $\angle CMO$ are complementary angles, since $\angle FMO$ is a right angle. Similarly, $\angle CJO$ and $\angle EJC$ are complementary angles. $\angle CNO$ and $\angle CJO$ are equal, since they are inscribed angles in $\bigcirc OCNJ$. Consequently, $\angle CNO$ and $\angle CHO$ are equal. Since we also know that $\angle HCO$ and $\angle NCO$ are equal(right angles), and they have a common side CO, we can derive that $\triangle OCM$ and $\triangle OCN$ are equal.

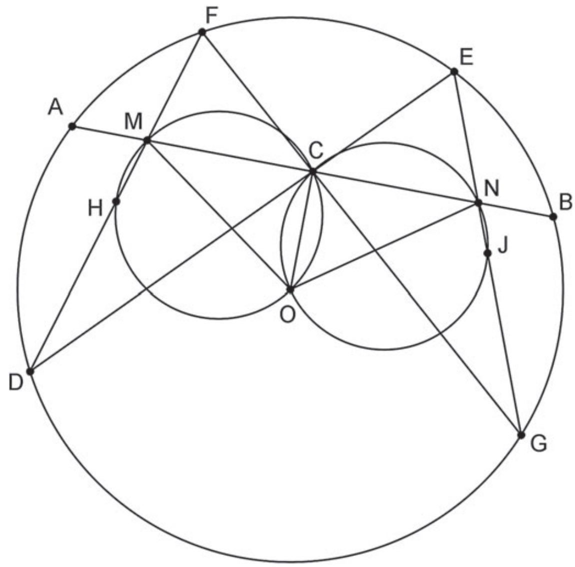

Fig. 4. Butterfly theorem: Final steps.

explanatory by (a) identifying a dependency to be explained, (b) identifying entities, (c) substituting the notion of activities with the notion of difference-makers, and (d) show that these difference-makers are organized such that the truth of the theorem is established. Finally, we will discuss a second proof that lacks this kind of information to clarify what is gained by our explanatory proof.

5.2. Capacity

All explanations start with the identification of the explanandum. In the case of mechanistic explanation the explanandum is a capacity of a system or a class of systems. A capacity is defined, parallel to the discussion above, as a systematic connection between a specific input and output. The class of systems that we ascribe a capacity to is here the set of quadrilaterals inscribed into a circle. The instructions used to construct the figure are the input:

C is the midpoint of a chord *AB* of a circle.
FG and *ED* are chords that go through point *C*.
FD cuts *AB* at *M*.
EG cuts *AB* at *N*.

If a quadrilateral *DFEG* inscribed in a circle satisfies these criteria, it will produce a specific output:

C is the midpoint of *MN*.

Within the set of quadrilaterals inscribed in a circle, one can find many figures that do not satisfy the input conditions. However, these specific conditions result in a specific output. Throughout the further discussion of using the mechanistic account of explanation in mathematics, we must be careful not to present an account in mathematics that sounds too causal. Rather than speaking of a capacity that is realized, we will speak of a dependency between the given input and output. The question why the butterfly theorem is true is thus the question why the system has the dependency between this result and the specific input.

5.3. Entities and Activities

Two important constituents of mechanistic explanation are entities and activities. The proof as well identifies certain entities of the original structure such as chords, triangles, midpoints, and angles. We can furthermore make claims about these entities or component parts:

> If a triangle has two angles that are equal to two angles of another triangle, these triangles are similar. (Step 2)
> A line from the centre of a circle, perpendicular to a chord, bisects the chord. (Step 5)

It would be peculiar to argue that these are claims about the activities of the entities. An activity would suggest that the magnitude of an angle produces similar triangles. We do not defend such a productive relation between mathematical entities. However, rather than speaking of some sort of active exchange between properties of mathematical entities, one can speak of a dependency between these properties. The property that the line is from the centre of a circle is dependently related to the property that the intersection is the midpoint of the chord.

5.4. Difference-makers

One can investigate this relationship between entities by varying the properties of the entities. If we vary a property of a mathematical entity,[3] we see how other properties change in response. It is for example possible to vary the property of the line from the centre of the circle to a chord:

> If a line from the centre of a circle is perpendicular to a chord, the chord is divided equally. If a line from the centre of a circle is not perpendicular, the chord is not divided equally.

[3]We can intervene on the value of the properties, but not on the properties themselves. We cannot, for example, change the fact that if a line from the midpoint of a circle is perpendicular, the chord is divided equally. But by intervening on the value of, for example, the angle, we can make its relation with the division of the chord clear.

Such an investigation corresponds to the interventionist account of Woodward [2003]. Woodward defines causality and explanations in terms of manipulations or interventions. The meaning of a causal claim is that intervening on some variable would change the value of another variable. This allows one to answer what-if-things-had-been-different questions, leading to an account of explanation by Woodward and Hitchcock:

> On our account the aim of explanation is to provide the resources for answering what-if-things-had-been-different questions by making explicit what the value of the explanandum variable depends upon. [2003, p. 190]

Several mechanists, such as Craver [2007] and Glennan [2002], refer to the interventionist account of Woodward, accepting that mechanistic explanation is in fact answering certain kinds of what-if-things-had-been-different questions. We treat interventions in mathematics, such as described above, as imaginary. The proof makes explicit that the butterfly theorem depends upon properties of the identified entities. These can be subject of what-if-things-were-different questions:

> What if the angles were not equal to the angles of another triangle?
> What if the line from the centre of the circle was not perpendicular to the chord?

These questions not only demonstrate the motivation for further steps of the proof showing how properties of specific parts are related to each other. The proof also identifies which entities and properties are relevant in order to explain why the theorem holds, since imaginary manipulation of an entity shows how the system, of which a dependency between an input and output is described in the theorem, changes in response. An entity and a property are relevant if changing the property would change the outcome of the proof. In other words, such entities and properties are difference-makers. The proof identifies which entities and properties are difference-makers. Interventions also allow one to see which entities and properties are not difference-makers. Answering the question “what if the slope of a chord is different” shows that the slope of the chord is irrelevant since intervention on it does not have any effect on further steps or the conclusion of the proof. The proof is consequently only a part of the explanation. The proof gives a person the relevant difference-makers. We argue that a person can gain deeper understanding by asking the appropriate questions. If a person succeeds in this task, (s)he understands how each separate part of the original figure contributes to the truth of the theorem, similar to the case of the electrical circuit discussed earlier. Such information, we argue, is genuinely explanatory information.

This approach further implies that the notion of activities does not have to be assumed in order to defend mechanistic explanation in mathematics. Both scientists and mathematicians search for difference-makers. In science this entails identifying entities and activities. Mathematicians on the other hand identify entities and dependencies between properties of these entities. Both approaches allow the investigator to answer what-if-things-had-been-different questions by intervention, and increase their understanding in how difference-makers are relevant for the explanandum. It should

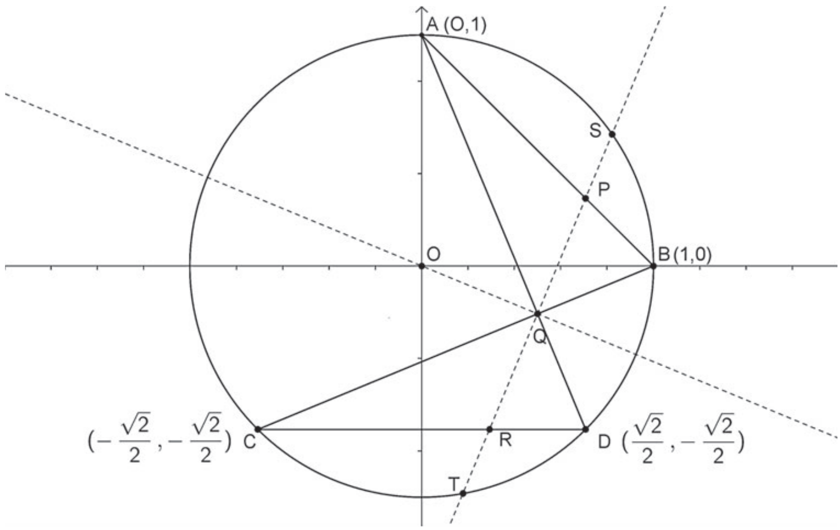

Fig. 5. Butterfly theorem.

nonetheless be clear that the concepts of difference-makers in science and in mathematics are not identical. Firstly, an intervention in science holds only *ceteris paribus*. This means that a scientist works in a framework which includes the rest of the mechanism and other scientific laws. If a part of the mechanism, for example, does not remain intact, a certain outcome of an intervention could no longer hold. The mathematical statement that "If a line from the centre of a circle is perpendicular to a chord, the chord is divided equally" holds in contrast for every circle in Euclidean space. Secondly, in contrast to science, there is no real observational movement in mathematics. We will discuss this point in Section 6.

5.5. Organization

The final element in the characterization of mechanistic explanations is organization. What does it mean to say that the entities and difference-makers in the case of the butterfly theorem are organized such that the theorem is true? Firstly, we need all difference-makers in order to establish the truth of the theorem. One difference-maker, or even all but one, do not yield an explanation of the butterfly theorem. In order to give such an explanation, we need to see how all difference-makers collectively explain the theorem. Secondly, there is a certain sequence in the steps of the proof that can not be changed. Likewise, there is a certain structure in the mechanistic explanation of the butterfly theorem. One could start with constructing points H and J (step 5), but one cannot come to step 6 without going through steps 1 to 4. Thirdly, the spatial organization is crucial. The same set of difference-makers could be applied on a similar structure that does not prove the butterfly theorem. Hence, we can see how not only the values of the difference-makers have effects on the proof, but that the entities and properties are organized such that the truth of the mathematical theorem is established.

5.6. A Second Proof

We can clarify what explanatory power we gain using the account by looking at a second proof of the butterfly theorem.[4] It should be noted that this is only a proof of a special case, and that we use specific co-ordinates. It is possible to discuss other cases and work with arbitrary co-ordinates, but that would significantly increase the length and complexity of the proof. We limit ourselves to the discussion of this case here, in order to clarify our account further. Let us take a look at the proof.

(1) Assume that:
Q is the midpoint of ST;
AB goes from $A(0, 1)$ to $B(0, 1)$; and
CD is parallel to the x-axis.

(2) The equation of AB is: $x + y = 1$.
The equation of CD is: $y = -\frac{\sqrt{2}}{2}$.

(3) Since we know the co-ordinates of both A and D, we know the equation of AD is:
$y - 1 = -\frac{2+\sqrt{2}}{\sqrt{2}}x$.
Similarly, we find the following equation of BC:
$y = \frac{\sqrt{2}}{2+\sqrt{2}}(x - 1)$.

(4) We now have a system of the equations AD and BC, which allows us to find the co-ordinates of point Q:
$x_Q = \frac{1}{2}; \quad y_Q = -\frac{\sqrt{2}}{2(2+\sqrt{2})}$.

(5) Using the co-ordinates of O and Q, we determine the equation of OQ:
$y = -\frac{\sqrt{2}}{2(2+\sqrt{2})}x$.

(6) Since we know that Q is the midpoint of ST, OQ is perpendicular to ST. This allows us to find the slope, and thus the equation of PR:
$y = \frac{2+\sqrt{2}}{\sqrt{2}}x - \frac{2+\sqrt{2}}{1+\sqrt{2}}$.

(7) Solving the system of equations of AB and PR gives us the co-ordinates of P:
$x_P = \frac{4+3\sqrt{2}}{6+4\sqrt{2}}; \quad y_P = 1 - \frac{4+3\sqrt{2}}{6+4\sqrt{2}}$.

(8) Similarly, solving the system of the equations of CD and PR gives us the co-ordinates of R:
$x_R = \frac{\sqrt{2}}{2+\sqrt{2}}; \quad y_R = -\frac{\sqrt{2}}{2}$.

(9) If we have $x_P - x_Q = x_Q - x_R$ and we have $y_P - y_Q = y_Q - y_R$, we can conclude that $|QP| = |QR|$. For the values of x we find:
$x_P - x_Q = \frac{4+3\sqrt{2}}{6+4\sqrt{2}} - \frac{1}{2} = \frac{\sqrt{2}}{4+2\sqrt{2}} = \frac{1}{2} - \frac{\sqrt{2}}{2+\sqrt{2}} = x_Q - x_R$.

[4]This proof was written by Jean Paul Van Bendegem after a discussion of this paper. We thank him sincerely for this contribution.

(10) For the values of y we find:
$y_P - y_Q = 1 - \frac{4+3\sqrt{2}}{6+4\sqrt{2}} + \frac{\sqrt{2}}{2(2+\sqrt{2})} = \frac{1}{2} = -\frac{\sqrt{2}}{2(2+\sqrt{2})} + \frac{\sqrt{2}}{2} = y_Q - y_R$.

(11) Since we have $x_P - x_Q = x_Q - x_R$ and we have $y_P - y_Q = y_Q - y_R$, we can now conclude that $|QP| = |QR|$.

This analytical proof gives us a justification of the butterfly theorem (of this particular case). Furthermore, it makes reference to certain entities of the figure such as points and lines. We do not, however, receive any information on how the properties of these entities are specifically organized such that the theorem holds. The only operations that are performed are determining equations, solving equations, and solving systems of equations. This is mathematically legitimate, and helps us to establish the truth of the theorem. However, it is not clear how this helps us understand why the theorem holds. What our proof, using the model of mechanistic explanation, contributes is showing how certain difference-makers underlie the theorem. By making interventions on the properties of the identified entities, it is demonstrated how and which other properties and the outcome of the proof change in response. As a result, given the identification of these difference-makers, the outcome of the proof is to be expected. We argue that this is genuine explanatory information.

However, we should perhaps not be too quick to conclude that the second proof cannot give such information. We argue that the first proof gives us clear information that we need to complete a deeper understanding of the theorem. While this is unclear in the second proof, perhaps someone with an expertise in analytical proofs can recognize difference-makers here as well. This is a final characteristic of our account. We argue that the use of concepts such as difference-makers and what-if-things-had-been-different questions, allows one to expect and explain a mathematical theorem. Since it involves certain insights from the person in question, one will not always extract an explanation from a potentially explanatory proof. Furthermore, every mathematician has one's own expertise, and this can lead to the fact that certain mathematicians find one proof identifies more satisfactory difference-makers than another. This results from personal preferences and knowledge, and we take it to be natural that these influence the appreciation of certain explanations.

6. POSSIBLE OBJECTIONS

In the previous section we already referred to the issue of 'capacities' and 'activities' in mathematics. It is important to stress as well that we do not make claims about the ontological status of mathematical entities. Approaching the proof in terms of mechanisms by referring to entities, difference-makers, and organization could appear to suggest that we are talking about actual objects, while mathematical objects are abstract and are not spatially or temporally localized. Paul Benacerraf [1973] raised this epistemological problem for Platonist positions towards mathematics. The mechanistic model can, however, be interpreted by both realist and nominalist accounts regarding the existence of mathematical objects.

Two types of realism about mathematics should be distinguished: metaphysical and semantic. Metaphysical realism is the belief that mathematical entities exist, coinciding with Platonist views of mathematics. Semantic realism on the other side is the belief that mathematical statements have objective truth-values. In order to identify entities and perform imaginary manipulations as discussed above, one can adopt a metaphysical-realist position but it suffices to adopt the semantic-realist position. One can answer questions about a triangle without adopting a metaphysical-realist position about this entity. Take for example the claim that an equilateral triangle cannot have a right angle. We cannot imagine manipulating an equilateral triangle in such a way that one of its angles is a right angle. In order to do so, we have to have a fixed meaning for an equilateral triangle and a right angle. One does not have to presuppose that these entities actually exist. Several nominalist approaches to mathematics, such as the modal-structural account of Geoffrey Hellman [1989], are semantic-realist without being metaphysical-realist.

7. ADVANTAGES

The account of Steiner does not allow us to grant any explanatory value to the discussed proof of the butterfly theorem. First, it is hard to see what the characterizing property of any entity mentioned in the theorem would be. The proof depends on properties of several entities such as circle, chord, and triangle. The theorem fails to hold if we drop any of its conditions, but it is impossible to isolate a property of an entity mentioned in the theorem. Furthermore, the proof is not deformable and consequently has no unifying value.

We argued that the proof can be seen as explanatory. More precisely, approaching the proof in terms of entities and imaginary manipulations allows us to answer what-if-things-had-been-different questions. Victor Gijsbers [2011] has also proposed an interventionist theory of mathematical explanation. He argues that a proof that is categorized as explanatory by Steiner can be reconstructed to be a Woodwardian explanation. In the case of the proof of the Pythagorean theorem discussed above, the explanatory power of the proof arises from the possibility of intervening on the right angle. This intervention shows how, starting from the cosine law that holds in every triangle, the Pythagorean theorem holds only in right-angled triangles. Such an account does not discuss bottom-up explanations, since it only addresses intervention on the class of objects mentioned in the theorem.

Our account goes further, stating that identifying parts of the mathematical object and showing how they make a difference for the truth of a theorem has explanatory power. This can be clarified by the notions of decomposition and localization, introduced by Bechtel and Richardson [1993]. Decomposition is the strategy where scientists divide the system into separate sub-processes. The authors assume that an activity of the system is a product of a set of subordinate functions performed in parts of the system. Localization subsequently means that scientists can indicate in which component parts these sub-processes occur. We argue that, at least for geometry, mathematicians likewise decompose mathematical figures and localize a set of component parts that are relevant difference-makers to the mathematical theorem. This approach goes beyond intervening on the class of objects mentioned in the theorem.

8. CONCLUDING REMARKS

We have argued that, while developed accounts of mathematical explanation fit top-down approaches to explanation, the bottom-up approach to explanation is also fruitful in mathematics, more precisely in geometry. Similar to a mechanistic explanation of an electrical circuit, a geometrical proof depends upon the identification of entities, properties, and difference-makers. The proof shows how the theorem depends upon properties of component parts of the mathematical structure. By imaginative manipulation of one property, one can see how properties of other entities and the truth of the mathematical theorem change. Such an approach allows one to answer what-if-things-had-been-different questions and gain insight into why the theorem is true. We have argued that such an approach does not necessarily entail postulating the existence of mathematical objects. Our approach furthermore shows that certain proofs are explanatory while having no unifying value.

We do not argue that the unificationist approach to mathematical explanation is wrong. We endorse a pluralist view of mathematical explanation. Both top-down and bottom-up explanations have a place in geometry. This does not render all proofs explanatory, since certain types of proof, such as proofs using *reductio ad absurdum*, fail to give either bottom-up or top-down explanations.

The benefits, limits, differences, and similarities of both types of explanation should be further analyzed in future research. This paper discussed the possibility of bottom-up explanation by looking at geometry. Future research should address whether this can be extended to mathematical domains outside geometry. Furthermore, it would be interesting to see whether our account could clarify the explanatory power of visual proofs.

REFERENCES

Bankoff, L. [1987]: 'The metamorphosis of the butterfly theorem', *Math. Magazine* **60**, 195–210.

Bechtel, W., and A. Abrahamsen [2005]: 'Explanation: A mechanist alternative', *Studies in History and Philosophy of the Biological and Biomedical Sciences* **36**, 421–441.

Bechtel, W., and R. Richardson [1993]: *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton University Press.

Benacerraf, P. [1973]: 'Mathematical truth', in P. Benacerraf and H. Putnam, eds, *Philosophy of Mathematics: Selected Readings*, pp. 403–420. Cambridge: Cambridge University Press.

Craver, C. [2007]: *Explaining the Brain*. Oxford: Clarendon Press.

Gijsbers, V. [2011]: *Explanation and Determination*. PhD thesis. Leiden University.

Glennan, S. [2002]: 'Rethinking mechanistic explanation', *Philosophy of Science* **69**, 342–353.

Hellman, G. [1989]: *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Clarendon Press.

Hempel, C. [1965]: 'Aspects of scientific explanation', in C. Hempel, ed., *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, pp. 331–496. New York: Free Press.

Kitcher, P. [1989]: 'Explanatory unification and the causal structure of the world', in P. Kitcher and W. Salmon, eds, *Scientific Explanation*, pp. 410–505. Minneapolis: University of Minnesota Press.

Machamer, P., L. Darden, and C. Craver [2000]: 'Thinking about mechanisms', *Philosophy of Science* **67**, 1–25.

Mancosu, P. [2001]: 'Mathematical explanation: Problems and prospects', *Topoi* **20**, 97–117.

——— [2008]: 'Mathematical explanation: Why it matters', in P. Mancosu, ed., *The Philosophy of Mathematical Practice*, pp. 134–150. Oxford: Oxford University Press.

Mordell, L.J. [1959]: *Reflections of a Mathematician*. Montreal: Canadian Mathematical Congress.
Salmon, W. [1989]: *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
Serway, R. [2008]: *Physics for Scientists & Engineers, with Modern Physics*. Philadelphia: Saunders.
Steiner, M. [1978]: 'Mathematical explanation', *Philosophical Studies* **34**, 135–151.
Weber, E. and L. Verhoeven [2002]: 'Explanatory proofs in mathematics', *Logique et Analyse* **45**, 299–307.
Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
Woodward, J., and C. Hitchcock [2003]: 'Explanatory generalizations: Part ii: Plumbing explanatory depth', *Noûs* **37**, 181–199.