

Are there Genuine Physical Explanations of Mathematical Phenomena?*

Bradford Skow

Abstract

There are lots of arguments for, or justifications of, mathematical theorems that make use of principles from physics. Do any of these constitute explanations? On the one hand, physical principles do not seem like they should be explanatorily relevant; on the other, some particular examples of physical justifications do look explanatory. In this paper I defend the idea that (some) physical justifications can and do explain mathematical facts.

1. Physical Arguments for Mathematical Truths
2. Preview
3. Mathematical Facts
4. Purity
5. Doubts about Purity, I
6. Doubts about Purity, II
7. How Physical Arguments Might Explain, I
8. How Physical Arguments Might Explain, II
9. Conclusion
10. Appendix A: Extracting an ‘Underlying’ Explanation
11. Appendix B: Another Example of an Explanatory Physical Argument

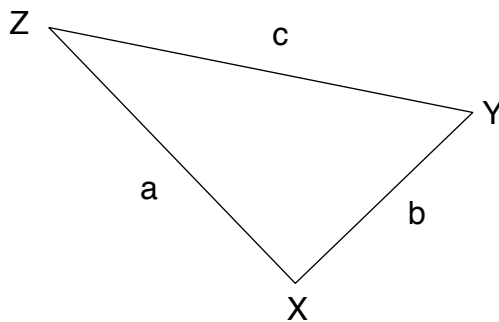
*Forthcoming in *British Journal for the Philosophy of Science* with the appendices removed.

1 Physical Arguments for Mathematical Truths

It is well-known that when it comes to solving physics problems mathematics can be a big help. Maybe less well-known is that this goes both ways. Sometimes reflecting on physical principles, or imagined physical set-ups, can lead to the discovery of mathematical truths. Sometimes physical principles can even help justify (provide a reason to believe, or evidence that favours) a mathematical truth. But can they do more? Can physical principles also help *explain* mathematical truths?

To get the question in focus let's look at an example: an argument for the Pythagorean theorem that uses physics. Imagine a right-triangular prism filled with gas and attached to a pole about which it can rotate. (Figure 1 depicts the prism from above; the right angle is at X; the pole passes through Z.) Its sides have lengths a , b and c . Imagine further that the prism is a closed system in equilibrium: the gas inside is not all bunched up against one of the sides, there are no giant masses nearby exerting gravitational forces on it, and so on. Since the prism is initially at rest the law of conservation of energy says that it will not begin to rotate. So the net torque on it must be zero.

Figure 1: The prism, from above



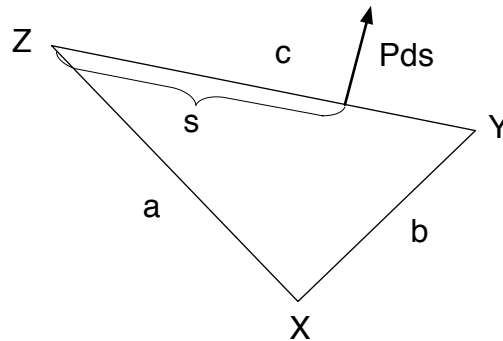
Let us find the individual torques that act on the prism. They come from the outward pressure of the gas on the prism's walls. (It may help to imagine that

the prism exists in a vacuum, so we can ignore the inward-directed force of the atmosphere.) Because the prism can only move by rotating about Z in the plane of the diagram we just need to find the torques about Z for walls ZY , ZX , and XY .

Now the torque about Z for a force of magnitude F acting on a portion of a wall is equal to $\pm Fs$ where s is the distance from Z to the line along which the force acts. The torque is positive if the force is trying to rotate the prism counter-clockwise; otherwise negative.

The force acting on a small portion of wall ZY (the hypotenuse), a portion of width ds , has magnitude Pds where P is the pressure of the gas, and acts at a right-angle to the wall (see figure 2). So the torque about Z for this force is equal to Pds . The total torque for ZY then is $\int_0^c Pds = Pc^2/2$. A similar calculation shows that the torque for ZX is $-Pa^2/2$.

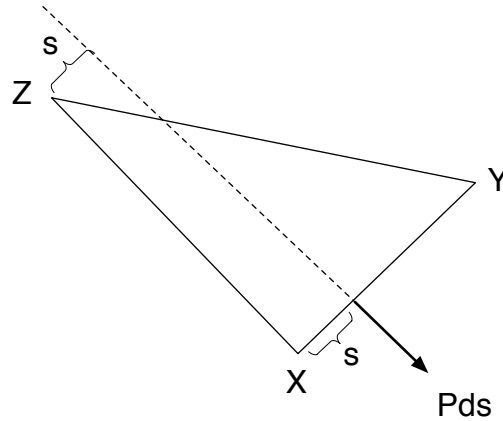
Figure 2: Calculating the torque for the hypotenuse



It is in the calculation for the torque due to the force on wall XY that the fact that the prism is right-triangular matters. The torque about Z for the force acting on a portion of that wall of width ds is $-Pds$ times the distance from Z to the line along which the force acts. It is because the angle at X is a right angle and the force is perpendicular to the wall that this distance is equal to the distance s from X to that portion of the wall. So the torque for that portion is (again) $-Pds$ and the total

torque for XY is $-Pb^2/2$. See figure 3.

Figure 3: Calculating the torque for side XY



Since the torques add up to zero,

$$Pa^2/2 + Pb^2/2 = Pc^2/2,$$

Or

$$a^2 + b^2 = c^2.$$

So there it is: an argument for the Pythagorean theorem that uses the law of conservation of energy. (It does not *just* use conservation of energy, of course. It also uses Newtonian mechanics and presupposes that space is Euclidean.)

Now: does the law of conservation of energy here help explain the Pythagorean theorem? The suggestion may sound absurd. But it is not universally denied. There are many books full of arguments for mathematical truths that are based on physical principles, and in at least one of those books — Mark Levi's *The Mathematical Mechanic* — the author repeatedly says that such arguments are explanatory, either by writing things like 'The Pythagorean theorem can be explained by the law of conservation of energy' (Levi [2009], p. 4) or by presenting a physical argument as

an answer to a why-question.¹ And I think Levi is on to something. When I myself read arguments like this they do not feel like mere tricks. So I want to give the idea that physical arguments can explain mathematical truths a run for its money.

2 Preview

For me this investigation starts from the fact that some physical arguments, like the Pythagorean theorem example, *seem* explanatory. It seems to me that I better understand why the theorems proved are true after reading those arguments. It would be great if I could back this up by presenting a general theory of explanation, complete with necessary and sufficient conditions for a body of fact to be an explanation, and then show that these physical arguments meet those conditions. But I do not have a general theory of explanation to defend.

Or, at least, I do not have one that can help with my project. It might well be true that something is an explanation if and only if it produces understanding, or that something is an explanation if and only if answers a why-question, though these have been disputed. But even if these are true, by themselves they are neither very illuminating nor very useful in this context. Someone who doubts that physical arguments can explain mathematical truths will also doubt that they answer why-questions or produce understanding.²

¹The first person I know of to reason to mathematical conclusions using physical principles was Archimedes. His treatise *Geometrical solutions derived from mechanics* was written in the third century BC (Archimedes [1909]). Later works of this sort include (Uspenskii [1961]) and (Kogan [1974]). Other places where Levi says things that entail that physical arguments can explain may be found on pages 4, 29, and 60 of his book. The justification of the Pythagorean theorem I am discussing appears in (Levi [2009], pp. 9-11) and in (Kogan [1974], p. 17).

²A referee said that if *all I mean* by ‘explanation’ is ‘answer to a why-question’ or ‘something that produces understanding’ then my thesis is not controversial, for almost anything can be an answer to a why-question, and one can understand why P is the case without knowing any reasons why P is the case. So, the referee continued, if I mean to be defending a controversial thesis (and I do) then I must mean something else — something stronger — by ‘explanation’, and I should say what that is. I have several things to say about this. First, by ‘why-question’ I mean ‘explanation-seeking why-question’, and I do not think that almost anything can

Since I lack a general theory of explanation I will take a less direct approach. I am going to begin by rebutting some arguments that physical principles cannot help to explain mathematical truths. Then I will make a positive case that they can explain. I will claim that at least some physical arguments that *seem* explanatory resemble in their structure things that are *certainly* explanatory. Of course the fact that X resembles an explanation in some respect or other is not much evidence that X is itself an explanation. But the kind of resemblance I will focus on makes the evidence stronger: the features that the physical arguments share with certain explanations are the features in virtue of which those explanations are explanatory.

3 Mathematical Facts

My thesis, that physical principles can help explain mathematical truths, can use some clarification. What distinguishes mathematical claims from physical ones, and more generally from non-mathematical ones?

I do not have a general answer to this question. But I do not think I need one, since I plan to stick to examples in which it is uncontroversial which claims are mathematical and which are not. But the example I started with may seem controversial, so let me say something about it.

The (alleged) mathematical truth derived in the example belongs to geometry. But understood one way geometrical claims are claims about *space*, a concrete, physical thing. Someone might say that understood this way the Pythagorean theorem is not a mathematical truth.

So far there is no problem; even if this is right there is surely also some way to understand geometrical claims so they do belong to mathematics. *That* is how

be an answer to one of those (though some philosophers, notably Van Fraassen [1980], think otherwise). Similarly, I do not agree that one can understand why P while being ignorant of any reasons why P (though some philosophers, notably Lipton [2009], have argued otherwise). These claims are worthy of defence, but this is not the place to defend them. What I will do is avoid resting any argumentative weight on the claims that something is an explanation if and only if it answers a why-question or produces understanding. I cannot say more than this about what I mean by 'explanation' since, as I said, I do not have a more substantive general theory of explanation to defend.

I want them to be understood in this paper. *One* way to do this is to interpret the Pythagorean theorem as a claim about ‘ideal Euclidean right triangles’, which exist in ‘the ideal three-dimensional Euclidean space’, an abstract object of some kind.³ (I do not say that this is the only way to understand the Pythagorean theorem so that it is a mathematical claim, but it is a useful one for us to have in mind.)

I take it that the distinction between ‘concrete’ and ‘abstract’ interpretations of the Pythagorean theorem is just common sense, but drawing it may just seem to make it easier to raise doubts about my example. Maybe the example seems explanatory only when we interpret the Pythagorean theorem concretely. Then if the example is an explanation, it does not explain a mathematical truth.

I do not think this is right. But what I think is wrong with it is a bit of a long story. I will tell it when I describe the structure I think the explanation has, in sections 7 and 8.

4 Purity

A simple and appealing argument against the idea that physical principles can help explain mathematical truths goes like this: a mathematical truth has only mathematical explanations (if it has any). The *reasons why* a mathematical truth is true are all mathematical. If this is so then no physical principle or law can help explain any mathematical truth, including the Pythagorean theorem.

Call the general principle this argument uses Explanatory Purity:

(EP) Mathematical truths have only mathematical explanations.

Is EP true? What arguments might be given in its favour? While I do not think EP is true I am going to postpone discussing these questions until the next section. Right

³Probably few platonists (believers in the existence of mathematical objects, and abstract objects more generally) think that there is some single abstract object that qualifies as *the* ‘ideal three-dimensional Euclidean space’, even though this seems like the sort of thing Plato himself would have believed. Even a categorical set of axioms for Euclidean geometry has infinitely many models. What could make one of the models the ‘right’ model (and what could ‘right’ even mean)? I suspect these questions can be answered, but if you like you can substitute ‘an’ for ‘the’ in the text.

now I want to explore what status physical arguments can have if EP is true. Even if EP is true such arguments need not be merely tricks. EP is compatible with the idea that physical arguments can be ‘illuminating’, indeed even with the idea that they can be explanatory. EP just entails that the physical principles in those arguments are not doing any of the explaining.

How can a physical argument be explanatory even if the physical principles in the argument are not doing any explaining? Like this: the physical argument ‘conveys explanatory information’ about the Pythagorean theorem, but the explanatory information it conveys is purely geometrical. The physical argument contains a purely geometrical explanation of the Pythagorean theorem ‘dressed up’ in non-geometrical garb.

There is precedent for this idea. Mathematics appears in many explanations of physical phenomena. But this fact alone does not establish that the mathematics is itself explanatory. In the course of giving an explanation there is often non-explanatory work that needs to be done and sometimes it is more efficient to subcontract that work out to mathematics. The mathematics might be there so that the explanatory facts can be expressed in fewer sentences of English. More generally, it might be there to make it easier for the listener to grasp the explanatory facts. Either role is consistent with the claim that the explanatory facts themselves are not mathematical.⁴ Maybe physical principles in physical arguments are similarly there just to facilitate the union between the mind that seeks understanding and the (purely mathematical) facts that confer that understanding.

How does one tell whether or not a physical principle is doing ‘real explanatory work’ in an explanation of a mathematical truth? The claim that the physical principles in an apparently explanatory physical argument are *not* doing any of the explaining is best justified by exhibiting the purely mathematical explanation that is hiding in the physical argument. Of course this may be very difficult to do. Maybe there are even cases in which it is impossible. Our inability to find such an explana-

⁴Field ([1980]) and Yablo ([2002]) emphasize these roles for mathematics. In a much-discussed article Baker ([2005]) argues that in some explanations of physical phenomena mathematics is not just doing some non-explanatory work but is also helping explain.

tion is not itself proof that the physical principles *are* explanatory. But it is at least good evidence.

So one can accept EP without thinking that it is just an illusion that going through a physical derivation of a mathematical truth sometimes improves one's understanding of that truth. EP need not be completely at odds with one's initial judgments about examples. But for me some conflict between EP and my judgments about examples remains. The physical argument for the Pythagorean theorem is explanatory. That's how it seems to me anyway. (And I think this about other physical arguments for other theorems that I have not discussed.) But I have not found a translation of the physical argument for the Pythagorean theorem into a purely geometrical proof.

You can strip away a lot of the physical principles from the argument. Once they are stripped away the argument starts like this: associate with each point on the triangle (except the vertices) an outward-pointing vector orthogonal to the side containing that point. Make all these vectors the same length Pds . For each side 'add up' the quantities Pds (the length of the vector times the distance s from Z to the line containing the vector); these are the integrals $\int Pds$. For each side make the resulting number associated with that side positive (/negative) if the vectors attached to that side point counterclockwise (/clockwise) relative to Z . You can get that far. To complete the proof we need the claim that the three numbers associated with the sides of the triangles sum to 0. The physical reasoning that give us that claim is (i) a triangle at rest (in these circumstances) will not begin to move, so (ii) the net torque on the triangle is zero, and (iii) the numbers we have associated with the sides of the triangle represent the torques on those sides. What mathematical principles can replace (i) through (iii)? I do not know.

5 Doubts about Purity, I

Of course it may be that the mathematical explanation is *there* but I have been unable to find it. Or it may be that I am wrong to think that the physical argument in my example is explanatory in the first place. None of what I just said is meant to convince someone who is drawn to EP that EP is false and my thesis is true. I

have only been exploring the extent to which physical arguments can be explanatory even if EP is true. But now I want to raise some doubts about EP itself and thereby clear the ground for the idea that the physical principles in a physical argument can themselves help explain.

First I want to discuss an argument directly for the conclusion that no physical fact can help explain a mathematical truth. This argument, even if sound, does not establish EP. To do that we need to show that *no* non-mathematical fact can explain a mathematical fact; and there are non-mathematical facts other than physical facts. But this argument is clearly still relevant. (It is also connected to a more general motivation for EP, as I will explain in the next section.)

The argument is simply this: (i) mathematical truths are *necessary* and physical facts are *contingent*. But obviously (ii) no contingent fact can help explain a necessary one.

Why might one believe (ii)? One might defend it by asserting, in general, that

(A) If F helps explain G then necessarily if G obtains so does F.

But (A) is false for well-known reasons. In a standard case of causal preemption Billy's throw explains the breaking of the window, even though in nearby possible worlds Billy does not throw but the window is still broken (because Suzy threw instead). True, (A) is stronger than needed to get (ii); all we need is the weaker claim

(B) *If G is necessary* and F helps explain G, then necessarily if G obtains so does F.

The retreat from (A) to (B) strikes me as ad hoc, but never mind. There are counterexamples to (B) too. A 'grounding' explanation explains why some fact F obtains by citing the 'more basic' facts in virtue of which F obtains.⁵ Now a disjunction is grounded in its true disjunct(s). So the fact (G) that either it is raining or it is not raining is grounded in (and so explained by) the fact (F) that it is not raining (supposing that this is a fact). But of course it is possible that G obtain while F does not.

⁵A thorough discussion of grounding may be found in (Rosen [2010]).

So I do not think (ii) has much going for it. But technically (ii) is also a stronger premise than the argument needs. It can get by with the weaker

(iii) No contingent fact can help explain a mathematical fact.

Now is there any reason to accept the weaker (iii) and not the stronger (ii)? I actually think that there is. It starts with the principle:

(C) If F helps explain G then in any world in which G obtains while F does not G *does not obtain for all the same reasons*.

This conditional is not just true but analytic (factors that explain G just are reasons why G obtains). From (C) it is easy to get to (iii): just use the premise that the reasons why mathematical truths are true do not vary from world to world.⁶

The argument we are considering now is that (iii) contingencies cannot explain mathematical truths, but (i) physical principles are contingent. It is better than the argument we started with but it still does not work. In sections 7 and 8 I will suggest that it is not really the fact that the law of conservation of energy is *true* that helps (or appears to help) explain why the Pythagorean theorem is true; it is the fact that it is *a law in the prism scenario*. But this is just the fact that the law of conservation of energy is *consistent with* Euclidean geometry, Newtonian mechanics, and the behaviour of this prism. And *this* fact obtains of necessity.

If I am right about all this then maybe the problem with the argument from (i) and (iii) is that (i) is false. Whether this is the right diagnosis depends on what ‘physical fact’ means as it occurs in (i). There are (at least) two things one might mean by ‘physical fact’. According to the first,

(M1) A physical fact is a fact about the goings-on in the physical (concrete) world.

Facts about where hydrogen atoms are, facts about how massive the planets are: these are paradigm physical facts in this sense. But in another sense physical facts are not identified by their subject matter but by the vocabulary used to express them:

⁶Why accept *this* premise? One might defend it by appealing to EP, but that would be question-begging in this context. Still, I accept the premise (even though I reject EP), so will not challenge it.

(M2) A physical fact is a fact statable in ‘physical vocabulary’.

This is not much help without a definition of ‘physical vocabulary’, but there is surely some definition on which words like ‘energy’, ‘force’, ‘mass’, and so on count as physical vocabulary.

Getting back to the argument from (i) and (iii), if we go with (M2) then the fact that the law of conservation of energy is consistent with Euclidean geometry (and so on) counts as a physical fact and (i) is false. If we go with (M1) then it is unlikely that this fact counts as physical.⁷ If it does not then even if (i) and (iii) are true the argument is irrelevant. It does not show that the example I started with is not explanatory.

I do not think that one of (M1) and (M2) gives the correct sense of ‘physical fact’ in some absolute sense. But I do think that (M2) is more relevant for our purposes. If it were to turn out that energy is not conserved in the concrete, physical world, and moreover (for strange reasons we cannot now grasp) that it is *metaphysically impossible* that energy be conserved in the concrete, physical world, that would not show that the my example fails to be explanatory. All the explanation really requires is the ‘formal consistency’ (or truth in some ‘abstract model’) of energy conservation with Euclidean geometry, Newtonian mechanics, and the behaviour of the prism. Still, even though it is formal consistency that matters terms like ‘energy’ and ‘force’ count as physical because the main interest we have in them and the theories in which they appear is their usefulness in predicting and explaining the behaviour of physical things.⁸

I should admit that my response to the argument from (i) and (iii) reveals my thesis to be somewhat less radical than it may first appear. It is not the thesis that the truth of the Pythagorean theorem (abstractly interpreted) depends in some way on the actual behaviour of sofas, stars, and galaxies. Still, it remains the case that the explanation of the Pythagorean theorem works by investigating how a possible (or even a ‘merely formally possible’) right-triangular prism behaves, and also how

⁷It certainly does not seem to be about the goings-on in the physical world on any ordinary sense of ‘about’. But some philosophical theories of aboutness, like Lewis’s ([1998]), have it that a necessary truth is about every subject matter.

⁸I thank a referee for pressing me to clarify this.

such a prism would behave under different conditions. (More on this in sections 7 and 8.) This certainly warrants calling it a physical explanation.

6 Doubts about Purity, II

The argument from (i) and (iii) focuses on the (apparent) ‘modal independence’ of mathematical fact from physical fact. It relies on the idea that changing the physical facts would not change the mathematical facts. The argument does not work, but someone who opposes my thesis might try to employ the strategy behind the argument in a slightly different way. They might keep the idea of using *independence* to argue for *explanatory irrelevance* but focus on a different kind of independence. Even if some physical facts are necessary, so that the mathematical facts are not modally independent of them, the mathematical facts are still *conceptually* independent of them. No concepts from mechanics (like energy, force, torque, and so on) are needed to state any mathematical facts. Mathematical facts can be expressed using only mathematical concepts. That is conceptual independence.

One way to turn this into an argument for EP is to use the premise

(P1) If fact F is conceptually independent of fact G then G cannot help to explain F.⁹

But this is false. The concept of gravity appears in an explanation of the fact that the moon orbits the earth. But the fact that the moon orbits the earth can be expressed without using the concept of gravity.

Maybe a better way to run the argument is to use a restricted version of (P1), one that applies only to potential explainers of *mathematical facts*. This new premise says

(P2) If a mathematical fact F is conceptually independent of fact G then G cannot help to explain F.

⁹Conceptual independence is a symmetric relation, so if (P1) is true then F cannot help to explain G either. (But not all kinds of independence are symmetric: event E may counterfactually depend on event C while C is counterfactually independent of E.)

What might motivate accepting (P2) and not (P1)? One idea is that conceptual independence is only a guide to explanatory relevance when the domain of fact being explained consists entirely of necessary truths (as mathematics does).

I think that (P2) is also false. For it entails a thesis stronger than EP that I will call ‘Strong Explanatory Purity’ (SEP). Let us say that an argument for a mathematical truth that uses only concepts that appear in the statement of that truth is *pure*. Then SEP is the thesis that explanations of mathematical truths must be pure. SEP is stronger than EP because not every mathematical proof of a mathematical theorem is pure.¹⁰ It is compatible with EP that impure mathematical proofs are explanatory; but SEP says they are not. I think that SEP is certainly not right.

There is a way of thinking about mathematical explanation that may validate SEP. On this way of thinking, for every branch of mathematics there is an elite set of propositions from that branch — the ‘fundamental’ propositions. And only a proof that shows how a proposition is made true by the fundamental propositions of the part of mathematics to which it belongs (or at least shows how it is made true

¹⁰Technically, SEP only entails EP in conjunction with the premise that mathematical facts are conceptually independent of non-mathematical ones.

by propositions more fundamental than it) is an explanation of that proposition.¹¹ Perhaps these proofs are all and only the pure proofs of that proposition.

But is this really how mathematical explanations work? This question has two parts. First: is this really *a* way for a mathematical explanation to work? And second: is this the only way? Let's take the two parts in order.

Regarding the first part of the question. It takes only a passing familiarity with the variety of ways a given part of mathematics can be axiomatized to come to doubt that exactly one of those axiomatizations is getting the fundamental truths right. (Are the fundamental topological propositions propositions about the behaviour of open sets or of closed sets?) The suspicion here is not that the best axiomatization has yet to be written down; it is the suspicion that the idea that some mathematical truths are fundamental is mistaken. And if there are no fundamental truths of topology then an explanation cannot show how any proposition about topology is made true by them.

Still, I am partial to the idea that there are fundamental mathematical truths and that proving P from the fundamental truths is a way to explain P. Is it the only way? The answer to the analogous question about explanations of physical events

¹¹The most well-known theory of mathematical explanation is perhaps Steiner's. He states his theory like this:

My proposal is that an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property. (Steiner [1978], p. 143)

Does this theory entail SEP? That depends on what a characterizing property is. Two of Steiner's examples are: (i) the property of being a product of such-and-such prime numbers is a characterizing property of a natural number; (ii) the property of being decomposable into two triangles each similar to the whole is a characterizing property of a right triangle. So the characterizing property of a natural number is a number-theoretic property and the characterizing property of a right triangle is a geometrical property. If all characterizing properties are like this then Steiner's view may entail SEP. Still, in his paper Steiner seems hostile to talk of essential properties of mathematical objects. That hostility suggests to me that he would not require that characterizing properties be expressible using only concepts used in the statement of the theorem being proved.

is clearly not true. Here are two answers to ‘Why is the stoplight red?’:

- Because it is emitting photons with frequencies near 450 THz.
- Because the car heading toward the light is speeding and the light is rigged to turn red when people speed toward it.

The first answer is a ‘grounding’ explanation. It explains why the light is red by citing the more fundamental facts in which its redness consists. The second answer is not a grounding explanation. Instead it is a causal explanation. So here a physical event has two explanations, one of which is not a grounding explanation, and the explanations coexist without conflict. Why can’t mathematical propositions also have non-grounding explanations? If they can, then even if all grounding explanations of mathematical truths are pure, mathematical truths may also have impure explanations.

Of course, an explanation of a mathematical proposition that is not a grounding explanation cannot be a causal explanation. But it would be rash to assume that since mathematical truths cannot have causal explanations they can have only grounding explanations. Maybe there are yet other modes of explanation.

So far I have just been trying to undermine the appeal of SEP. Now for a more direct line of attack. Here is a counterexample to the principle.¹² If f is an infinitely-differentiable real-valued function on the real line then it has an associated power series $f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \dots$. This is the Taylor series for f . Obviously f is equal to its Taylor series at $x = 0$. But how far out does this equality extend?

¹²Arana and Mancosu ([2012]) discuss another potential counterexample. Desargues’ theorem says that if the lines through corresponding vertices of two triangles (lying in the plane) meet in a point then the intersection points of lines along corresponding sides lie on a single line (that is, if two triangles are in perspective from a point then they are in perspective from a line). To prove this theorem one must either use considerations from three-dimensional solid geometry or use metrical considerations — yet the statement of the theorem makes no use of metrical notions and is about plane figures. If any of these standard proofs are explanatory and impure then they are also counterexamples. I prefer the example I give in the body of this paper because I think it is easier to argue that that proof is impure.

For every power series $\sum c_n x^n$ there is some positive number R (possibly $R = \infty$) such that the series converges inside the interval $(-R, R)$ and diverges when $|x| > R$. So a function may fail to be equal to its Taylor series for large x because its Taylor series is no longer defined (no longer converges). The explanation I want to look at focuses on a particular example of this failure.

Let f be the infinitely-differentiable function on the real line defined by $f(x) = \frac{1}{1+x^2}$. The Taylor series for f is

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

This series converges when $|x| < 1$ and does not converge when $|x| \geq 1$. The function f , on the other hand, is perfectly well-defined for $|x| \geq 1$. What we want explained is: why does f fail to be equal to its Taylor series for $|x| \geq 1$?

A contrast with another function makes this explanatory request more urgent. The function $h = \frac{1}{1-x^2}$ has for its Taylor series the sum $1 + x^2 + x^4 + \dots$. This series also fails to converge for $|x| \geq 1$, but here it is not surprising that h fails to equal its Taylor series for $|x| \geq 1$: the function h itself blows up to infinity at $x = \pm 1$. Since h is not well-behaved when $|x| = 1$, it is not surprising that its Taylor series cannot be extended past those points. But nothing like this happens with f . The function f is defined and well-behaved everywhere, including at $|x| = 1$. So what ‘prevents’ f from having a Taylor series that extends to larger values of $|x|$? That is what we want to explain.

It is a commonplace in mathematics that we can only explain this if we use facts from complex analysis.¹³

The explanation goes like this. Consider the function $F(z) = \frac{1}{1+z^2}$, defined on the complex plane. Just as in real analysis, in complex analysis for every power series (in the variable z) there is an $R > 0$ such that the series converges for $|z| < R$ and diverges for $|z| > R$. That is, a complex power series converges everywhere inside some circle and diverges outside the circle; R is the radius of this circle. The

¹³See for example (Needham [1997], pp. 64-5) and (Spivak [2008], pp. 516, 563). Long ago the mathematician Hadamard said ‘The shortest path between two truths in the real domain often runs through the complex numbers’ (quoted in Wilson [2006], p. 314).

theorem in complex analysis that has no counterpart in real analysis, the theorem we are interested in, says that the radius R of this circle for the Taylor series of a function h is equal to the distance from 0 to the nearest point at which h is undefined.

So to explain the relationship between f and its Taylor series we first apply this theorem to F : the nearest point at which the function F is undefined is $z = i$, so the radius of convergence for the Taylor series for F is 1; outside that radius the Taylor series for F does not converge. It follows that this is true also for the restrictions of F and G to the real line; $f \neq g$ for $|x| > 1$.

What needed explanation was the fact that f 's Taylor series stopped converging at ± 1 , even though f is well-behaved everywhere on the real line. We see now that the reason the series stops converging there is that f , when regarded as a function on the complex plane, is *not* well-behaved everywhere; it blows up at i , a point distance 1 from the origin.

This explanation is not pure.¹⁴ The original question was a question about the behaviour of a real-valued function on the real line, and the answer looked at the behaviour of another function, a complex-valued function on the complex plane, that happens to agree with the original function on the real line. And complex analysis uses concepts (like the concept of an imaginary number) that are foreign to real analysis. Of course real analysis is part of complex analysis. But it is only a part, and the main theorem used in the explanation — that the radius of convergence of a function's Taylor series is the distance from the origin to the nearest singularity — comes from the part of complex analysis that goes beyond real analysis. This theorem cannot be 'translated' into statements of real analysis. Since the complex analysis cannot be removed from the explanation it is playing a real explanatory role. Strong Explanatory Purity is false.

One might respond to all this by looking for a principle weaker than SEP that escapes this counterexample but still precludes physical facts from explaining mathematical ones. Before looking at any weaker principles it is worth remembering that SEP did not arrive out of thin air. It was motivated by the idea that

¹⁴Mathematicians (Spivak and Needham, cited above) and philosophers (Lange [2010], p. 329 and Steiner [1978], pp. 18-9) agree that this example is explanatory. Dawson ([2006], p. 280) agrees that it is impure.

conceptual independence makes for explanatory irrelevance. This idea is not going to motivate any weaker principle. So there is a worry that a weaker principle will be ad hoc.

Some philosophers say that the complex numbers are the ‘natural setting’ for investigating — and understanding — the behaviour of functions from \mathbb{R} to \mathbb{R} . One might propose a cousin to SEP that uses this notion of a natural setting rather than the notion of purity: an argument is explanatory only if it draws on ideas or premises from the natural setting for the mathematical statement in question.¹⁵ This principle does not look too ad hoc. And my counterexample to SEP is not a counterexample to this principle. But physical facts do not seem to be part of the natural setting for any mathematical truth.

I am not convinced. For one thing, I worry about how this notion of a natural setting is to be understood. It may seem that what makes the complex numbers a natural setting for investigating functions from \mathbb{R} to \mathbb{R} is that the complex numbers contain the real numbers as a part. But I doubt that this definition will cover all uses of the notion. Nor do I think it works for this case: the quaternions also contain the real numbers as a part, and no one suggests that the quaternions are a natural setting for investigating these functions.¹⁶ It seems to me that the most straightforward reason for thinking that the complex numbers are the natural setting is this: it is only from the perspective of complex analysis that the behaviour of certain functions from \mathbb{R} to \mathbb{R} can be fully explained. This reason goes best with a definition of ‘natural setting’ in explanatory terms: the natural setting for investigating X *just is* the body of fact which provides for explanations of X. But if this is the definition then the cousin to SEP is of no dialectical use. It is begging the question to argue that the law of conservation of energy does not belong to the natural setting for the Pythagorean theorem.

¹⁵A referee pressed this response. Jamie Tappenden has discussed this notion of a natural setting in, for example, (Tappenden [1995]). He explicitly connects this notion to explanation, but does not there endorse any precise thesis about their connection. See also (Wilson [2006], pp. 313-5).

¹⁶The complex numbers, like the real numbers, are a field, while the quaternions are not. But then again the real numbers are an ordered field while the complex numbers are not. There are similarities and differences in both cases.

The cousin to SEP was offered as a thesis that avoided my counterexample to SEP and could still be used to show that physical arguments do not explain. But my case against SEP did not rest on that counterexample. Even if the cousin principle can be stated so that it is not question-begging to use it, the first problem I discussed for SEP is still a problem for it. It may be right that a *grounding* explanation of a mathematical truth may draw only on facts from that truth's natural setting. But I see no reason to extend this restriction to non-grounding explanations.

Let me say a final thing about SEP. (Remarks similar to the following apply to its cousin.) One thing that makes SEP appealing is that if true it would explain why mathematicians continue to seek pure proofs of theorems that have already been proved: they want an explanation, not just a proof.¹⁷ But there are a couple of alternative theses about purity in mathematical explanation that are close to SEP which, I think, might well be true and, if true, could also explain this fact about mathematical practice.¹⁸ Perhaps those who find SEP appealing will be happy with one of them instead. Here are the alternatives (one of which connects with what I said earlier about grounding explanation):

- A pure proof of a theorem is *more explanatory* than any impure proof.
- The only *grounding explanations* of a theorem are pure proofs.

(These claims may be connected; one might think that the best explanations of mathematical truths are grounding explanations.) If these claims are true then the search for pure proofs is a search for more explanatory proofs, or for proofs that explain in a distinctive kind of way. And those are certainly intelligible motives. But these claims leave it open whether physical arguments can explain. They are no threat to my thesis.

¹⁷Of course, there may be reasons for seeking pure proofs that have nothing to do with explanation. Detlefsen and Arana ([2011]) argue that pure proofs better justify their conclusions (in a sense of 'justify' that they describe). If they are right then this is another reason to prefer pure proofs.

¹⁸I thank two anonymous referees for getting me to think about these alternatives.

7 How Physical Arguments Might Explain, I

So far my arguments have been negative: I have tried to undermine some reasons to doubt that physical arguments explain. But doubt that physical arguments can explain will remain as long as we have no idea how they might go about doing it. So now I want to look for some kind of ‘model’ of how physical arguments might explain.

Let’s go back to the derivation of the Pythagorean theorem in the example. What the calculation I went through establishes is this:

- (1) Necessarily, if Euclidean geometry, Newtonian mechanics, and conservation of energy are all true, and there is a prism filled with gas (and so on), then the Pythagorean theorem is true.

But the conclusion we really want is this:

- (5) Necessarily, if Euclidean geometry is true then the Pythagorean theorem is true.

It is not, however, very hard to add premises that get us from (1) to (5). After doing so the whole argument looks like this:

- (1) Necessarily, if Euclidean geometry, Newtonian mechanics, and conservation of energy are all true, and there is a prism filled with gas (and so on), then the Pythagorean theorem is true.
- (2) It is possible that Euclidean geometry, Newtonian mechanics, and the law of conservation of energy all be true, and that there be a prism filled with gas (and so on).
- (3) Therefore, it is possible that Euclidean geometry and the Pythagorean theorem be true.
- (4) If it is possible that Euclidean geometry and the Pythagorean theorem be true then necessarily, if Euclidean geometry is true then the Pythagorean theorem is true.

(5) Therefore, necessarily, if Euclidean geometry is true then the Pythagorean theorem is true.

Earlier, in section 3, I raised the concern that all this argument really explains is why the Pythagorean theorem is true when interpreted concretely. But (1) is true whether we are interpreting (all) the relevant theories concretely or abstractly. And I have said that even if (2) is (for some reason) false when interpreted concretely it is certainly true when interpreted abstractly. Finally, the modalities in (4) are intended to cover both kinds of interpretations (so its form is: if R is possible on *one* kind of interpretation then S is necessary on *both* kinds of interpretation). The conclusion (5), therefore, *is* about the Pythagorean theorem interpreted abstractly.

I think, however, that the structure of this argument is a poor guide to the structure of the explanation the argument conveys. This is a familiar idea: although Hempel ([1965]) claimed that explanations are arguments, many philosophers since then have persuasively defended the claim that an explanation is, roughly speaking, a body of fact about *dependencies*. An argument *may* contain a body of fact about dependencies, but not all arguments do, and an argument is not the only way to convey such a body of fact. A physical event, for example, may be explained by information about its causes, and this information need not be packaged into an argument.¹⁹

So what structure does the explanation of the Pythagorean theorem have? The following is a bit speculative. My idea is that the explanation has two parts. I will discuss the first part now and get to the second in a little while. The first part is a counterfactual:

(E) If the Pythagorean theorem were not true in Euclidean geometry then, in the prism situation, the net torque on the prism would not be zero, so the prism would rotate.

(‘The prism situation’ includes not just the fact that there is a right-triangular prism initially at rest, but also the fact that Euclidean geometry is true and the laws are the

¹⁹This point is prominent in (Lewis [1986b]) and (Woodward [2003]). Of course all (deductive) arguments convey information about *logical* dependencies, but these are not the kind relevant to explanation.

laws of Newtonian mechanics and the law of conservation of energy.)

Why is this counterfactual true and a good thing to say? There is some pressure to say that (E) is at best misleading. If the Pythagorean theorem were false then the prism situation would be inconsistent. So why single out the torque on the prism as the thing that would be different? In fact, shouldn't we say that if the Pythagorean theorem had been false then *everything* would have been true in the prism situation? Newtonian mechanics would have been both true and false, the prism both would and would not have been right-triangular, and so on.

There are two questions here. Why single out the torque, rather than some other feature of the situation, as what would have been different? And why single out *any* particular thing as what would have been different, rather than say that everything would have been different (and also the same)? The second question is easier. I deny that all counterfactuals with impossible antecedents are vacuously true. It is, perhaps, impossible that I be a fried egg. But it is nevertheless false that if I were a fried egg I would have one hundred legs.

I do not have a detailed theory of the workings of counterpossibles to give in response to the first question.²⁰ But this seems to be a relevant consideration: in the derivation of the Pythagorean theorem (the defence of (1)) the quantity that the Pythagorean theorem is most closely connected to is the net torque. The Pythagorean theorem is deduced more or less immediately from the equation expressing the claim that the net torque is zero. I suspect that their proximity in the derivation is one factor that makes (E) true (in a context in which that derivation is salient, of course).

I have been arguing that (E) is true and interesting. But even if it is how can it be part of an explanation of the Pythagorean theorem? Although it is generally agreed that there is some connection between counterfactuals and explanation, (E) looks like the wrong kind of counterfactual. Usually we explain X by citing factors

²⁰Some philosophers hold that even if A is impossible we should still say that 'If it had been that A it would have been that B' is true if B is true at the nearest A-world; we must just allow for impossible worlds in addition to possible worlds. Then not all counterpossibles are vacuously true. This still leaves us with the task of saying something informative about what makes one impossible world closer than another. See for example (Nolan [1997]).

on which X depends. But that is not what (E) does. It does not say: if such-and-such had not obtained then the theorem would have been false (in Euclidean geometry). Instead it tells us something about *what depends on the theorem*, namely the behaviour of the prism.²¹

Can something be explained by citing factors that depend on it? There is precedent for this idea. Teleological explanations (some of them anyway) work like this. Why is this acorn sprouting this way? Because if it sprouted some other way it would not be able to reach its end — to become a mighty oak. It is easy to see why in some teleological explanations a fact or event is explained by saying something about what depends on it. In a teleological explanation a behaviour is explained by its end, its final cause. But the final cause is often an effect of that behaviour. And effects depend on their causes.

Now maybe no teleological explanation like this is true.²² Maybe in fact no teleological explanation like this *could* be true; maybe there is no way the world could have been in which it is true. It does not follow from this that teleological explanations exemplify a defective style of explanation.²³ Maybe facts about what depends on P *can* explain P (for some values of P); but the conditions under which these kinds of facts get to be explanatory cannot be realized for plants.

I want to suggest that they can be realized for mathematical truths.

²¹The worry here is that (E) is the wrong kind of counterfactual because it has the wrong kind of antecedent and consequent. There is another way a counterfactual can be of the wrong kind for explanation: it might be true only in the wrong kind of context. David Lewis ([1986a]) claimed that while counterfactuals like ‘If the bomb had not gone off at 12, no one would have lit the fuse at 11’ are true in some contexts, those contexts are not contexts in which there is a connection between counterfactuals and causation, so we do not have to say that the explosion of the bomb explains the lighting of the fuse. But I see no reason to suspect that (E) is true only in the wrong kind of context.

²²Some philosophers have attempted to show that some teleological explanations are really just (efficient) causal explanations. On this view ends or final causes do not play an *irreducible* explanatory role. By ‘true teleological explanation’ I mean ‘true, irreducible teleological explanation.’

²³Hawthorne and Nolan ([2006]) try to describe what the laws would have to be like for there to be teleological explanations of natural phenomena.

But before I go on I should emphasize that teleological explanations are here just an *example* of a style of explanation in which a fact is explained by citing other facts that depend on it. I do not mean to say that the physical explanation of the Pythagorean theorem is itself a teleological explanation. In a teleological explanation something's end plays an important role. The concept of an end plays no role in the physical explanation of the Pythagorean theorem — though, as we will see, there is something that does play a similar role.

Still, the analogy with teleological explanation is fruitful. Suppose that the teleological explanation of the acorn's sprouting is true (and irreducible). It seems to explain the sprouting indirectly. It shows that the acorn *has* to sprout, if it is to achieve its end. But it says nothing about the physical process that went on inside the acorn and caused it to sprout that way. So the existence of this teleological explanation is compatible with the existence of an ordinary (efficient) causal explanation of the acorn's sprouting. A causal explanation will explain the sprouting more directly by saying something about that process.

The explanation of the Pythagorean theorem seems to be similar.²⁴ It shows that the Pythagorean theorem *has* to be true, if energy conservation is to be consistent with Euclidean geometry, Newtonian mechanics, and the behaviour of the prism. But it says nothing about the geometrical reasons why the Pythagorean the-

²⁴Baker's well-known example of an (alleged) mathematical explanation of a physical phenomenon (Baker [2005]) also works by citing a factor that the event being explained is a difference-maker for. Why do cicadas' life-cycles have lengths that, when measured in years, are prime numbers? Because if they did not the cicadas would encounter predators more frequently. But I do not think this suggests a deep similarity between mathematical explanations of physical phenomena and physical explanations of mathematical phenomena. The quasi-teleological structure of the explanation Baker discusses is an instance of the form 'organism X has trait Y because that trait is the fittest.' Explanations of this form do not need to invoke mathematics. And, I think, the mathematics in Baker's explanation does not primarily explain why cicadas have prime life-cycles. What explains that is the fitness of those life-cycles. If the mathematics explains any physical fact (and whether it does is in dispute — I myself do not think it does) it is the fact that life cycles with prime lengths are the fittest. And the explanation of this fact does not have a quasi-teleological structure. (Thanks to a referee for asking about this apparent similarity between Baker's example and the ones I discuss.)

orem is true. A pure (geometrical) explanation will explain this more directly by citing those geometrical reasons.

I am getting a little ahead of myself. I have said that the counterfactual (E) is only part of the explanation of the Pythagorean theorem. All by itself it explains nothing. What else is needed?

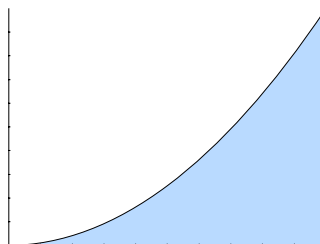
The analogy with teleological explanation suggests an answer. The counterfactual ‘If the acorn had not sprouted like this, it would not have become a mighty oak’ also explains nothing by itself. To get a teleological explanation we need to add to this counterfactual the claim that becoming an oak is the acorn’s end. If there is teleology in nature then a thing’s end acts as a constraint on how it may behave. The acorn does not sprout in some other way because doing so would violate that constraint. My hypothesis is that in the Pythagorean theorem example something is acting as a constraint on the prism’s behaviour. And the Pythagorean theorem is true because its being false would lead to behaviour (rotation) that would violate that constraint. That constraint, I suggest, is the law of conservation of energy.

To flesh out and defend this idea I want to compare the physical argument for the Pythagorean theorem with another physical argument that I think is not explanatory.

8 How Physical Arguments Might Explain, II

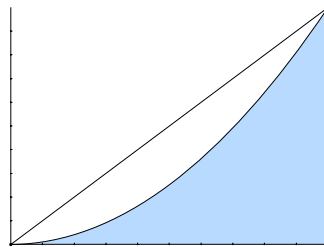
Archimedes gave a derivation of the area under a parabola between $x = 0$ and $x = 1$ (the area of the shaded region in figure 4 — the curve is the graph of $y = x^2$) that used principles from mechanics.

Figure 4: The shaded region



Imagine that the shaded region is a real physical thing, one with constant density. (If you find it difficult to imagine a real physical thing that is two-dimensional, imagine it to be, say, 1 meter thick.) Now consider the triangle got by drawing a diagonal across the unit square and imagine it too to be a real physical thing (again with constant density). Let us compare the mass of these two things (show together in figure 5).

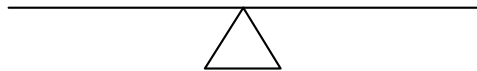
Figure 5: The shaded region and the triangle



We will compare their mass using a lever (figure 6). The law of the lever says

- Bodies placed on opposite sides of the fulcrum (in a uniform-downward gravitational field) will balance just in case $M_1a = M_2b$, where M_1 and M_2 are the bodies' masses and a and b are the distances from the fulcrum to their centres of mass.

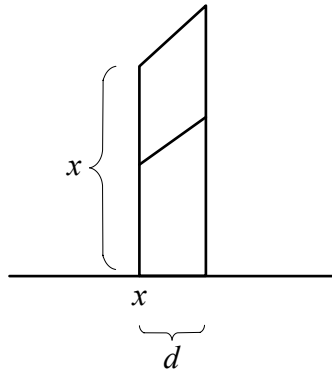
Figure 6: A balance



Now divide the triangle and the region under the parabola into thin vertical segments of width δ , where δ is very small. Consider the segment of the triangle and the

segment of the parabola that sit above some point x on the horizontal axis (shown blown-up in figure 7). The height of the short side of the segment of the triangle is

Figure 7: Segments of the triangle and shaded region over point x



x meters (in this derivation all distances are measured in meters) and so it's mass is approximately²⁵ $x\delta$ kg (I am assuming that the density is 1 kg/m^3). The height of the short side of the segment of the parabola is x^2 meters and so that segment's mass is approximately $x^2\delta$ kg. By the law of the lever, then, if the segment of the parabola is placed at a distance 1 from the fulcrum and the segment of the triangle is placed at a distance x on the opposite side of the fulcrum they will balance. But this is true for every segment. 'Add' these results and replace the segments of the parabola piled on top of each other with a point mass (see figure 8); it follows that

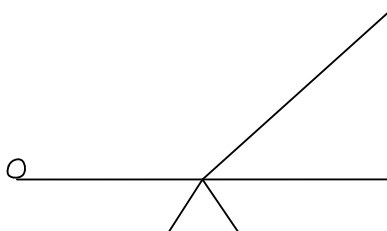
- A point-mass with the same mass as the shaded region, placed at distance 1 from the fulcrum, balances the triangle placed with its tip is over the fulcrum.

The mass of the triangle is $1/2$, and the effect of the triangle on its side of the lever is identical to the effect of an equally massive point particle sitting below the triangle's centre of mass. Again using the law of the lever we see that

- (Mass of the shaded region) = $\frac{1}{2}$ (horizontal distance from the fulcrum to the triangle's centre of mass).

²⁵Really the mass is $x\delta + \delta^2/2$. But δ is small so δ^2 is extremely small and we can neglect it. Similar remarks apply to the calculation of the mass of the segment of the parabola.

Figure 8: The point-mass balances the triangle



Now the triangle's centre of mass is located above $x = 2/3$.²⁶ So the mass of the shaded region is $1/2 \times 2/3 = 1/3$. Since its density is 1 kg/m^3 , the area of the shaded region is $1/3$ of a square meter.

This is a brilliant piece of reasoning that uses mechanical principles to derive a result that usually requires calculus²⁷ (which was unavailable to Archimedes). But I do not think it explains why the area under the parabola is $1/3$. Why not? How is it different from the physical argument for the Pythagorean theorem? Here is a stab at an answer.

Focus on the situation depicted in figure 8. We have a lever with a point-mass on one side and a triangle on the other. And we know that the mass balances the triangle. So far there are close parallels with the prism situation. We have a physical set-up (a balance, a prism) and know enough about the laws and the set-up to know how it will behave (it will balance, it will remain at rest).

The parallels continue: just as we have the counterfactual (E) about the prism situation, we seem to have this counterfactual about the balance situation:

²⁶The centre of mass of a triangle is the point at which its medians intersect (a median is a line from a vertex to the midpoint of the opposite side). It is a standard theorem of Euclidean geometry that this intersection point is $2/3$ of the way along a median from any vertex.

²⁷Some calculus is still needed: the derivation needs to be supplemented with an argument that the differences between the true areas of the segments and the approximate areas used in the calculation go to zero sufficiently fast as $\delta \rightarrow 0$.

(N) If the area of the shaded (parabolic) region were not $1/3$, then (since the point-mass has the same mass as that region) the point mass would not balance the triangle.

Like (E) this counterfactual tells us something about what depends on the fact we want explained: the fact that the area of the shaded region is $1/3$.

It is here that there begin to be differences between the two examples. I contend that there is nothing in the balance situation that ‘constrains’ the point-mass to balance the triangle. But (N) can only explain why the area is $1/3$ in the presence of some such constraint (if, that is, Archimedes’ argument is to explain in the same way that I think the argument for the Pythagorean theorem does).

Of course Archimedes *did* provide an argument that the point-mass balances the triangle. The argument was that the segment of the shaded region over a point x will balance the segment of the triangle over x if the segment of the shaded region is moved to a point one meter away from the fulcrum (and opposite the triangle segment). The point-mass then replaces the piled-up segments of shaded regions: it is located in the same place and its mass is the sum of their masses. Why is this not enough to constrain the point-mass to balance the triangle?

Because this argument does not ground the balancing independently of the fact being explained. I said that in the style of explanation I think the Pythagorean theorem example exhibits X is explained by the fact that the failure of X would lead to something Y that violates some ‘constraint’. What I intend by this is that Y is *independently forbidden*. For nothing can help explain itself;²⁸ but X would be doing just that if it played a part in forbidding Y . That is what is happening in the Archimedean example. The argument that the point-mass balances the triangle uses facts about the areas of segments of the shaded region. But those are just the facts that are called into question when we consider what would be the case if the area of the shaded region were not $1/3$.

The fact that the prism will not rotate, by contrast, is grounded independently of the fact being explained. It is grounded in the law of conservation of energy. And

²⁸Except in very strange situations, like time travel scenarios containing causal loops.

this law is not immediately called into question when we consider what would be the case if the Pythagorean theorem were false.

9 Conclusion

My claim has been that the physical argument for the Pythagorean theorem has within it a structure that resembles the structure of teleological explanations. This resemblance is the main positive reason I have to offer in favor of that argument's constituting an explanation. My description of that structure uses the notion of a constraint, and of a fact being 'independently grounded'. I recognize that these notions could use further clarification and analysis, especially when they are applied to necessary facts. But I have tried to give us some handle on them by using them to describe the two contrasting examples. My hope is that I have made these notions clear enough to tell whether my thesis is on the right track. If it is then there are lots of interesting questions, about these notions, about explanation in mathematics, and about explanation more generally, waiting up ahead.

Appendix A: Extracting an 'Underlying' Explanation

I said in section 4 that a physical argument might 'hide' a purely mathematical explanation. Here is an example to illustrate this kind of phenomenon. It is not an example in which a mathematical explanation is hiding in a physical argument. Instead it is an example in which a purely algebraic proof (of an algebraic fact) is hiding in a geometrical proof. That should be close enough to make the idea clear.

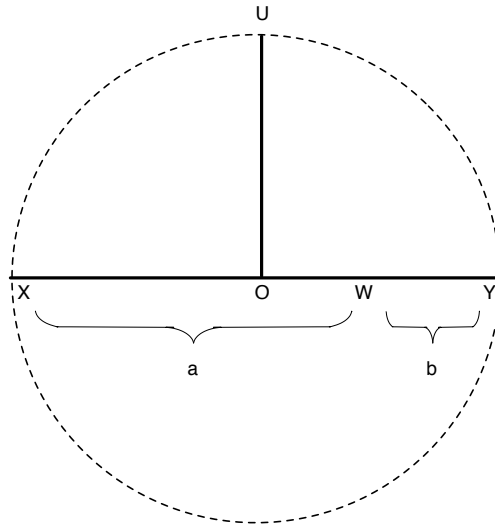
For any two positive real numbers a and b , $(a + b)/2$ is their arithmetic mean (what we ordinarily call their average) and \sqrt{ab} is their geometric mean. The arithmetic mean-geometrical mean (am-gm) inequality says that the geometric mean is never greater than the arithmetic mean:

$$\sqrt{ab} \leq \frac{a+b}{2}. \quad (1)$$

Here is a geometrical proof of the inequality, a proof that looks explanatory.

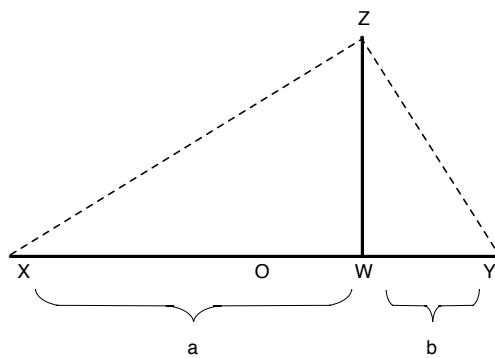
To start we need a geometrical representation of the arithmetic mean. Inter-

Figure 9:



pret a and b to be the lengths of two line segments. Then lay those line segments end to end along the diameter XY of a circle, as in figure 9. Since $a + b$ is the length of the diameter, the arithmetic mean $(a + b)/2$ is the length of the circle's radius, and so is the length of OU .

Figure 10:

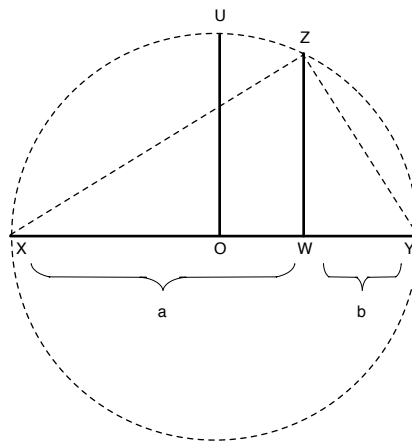


We also need a geometrical representation of the geometrical mean. It may be seen in figure 10. Erect a perpendicular at W and consider the right triangle that

has XY as its hypotenuse and its right-angled vertex Z on that perpendicular. Then triangles XWZ and ZWY are both similar to XZY and hence to each other. Thus the ratios of lengths of corresponding sides are the same: $XW/WZ = WZ/YW$. So $(WZ)^2 = (XW)(YW) = ab$. Conclusion: the length of segment WZ is \sqrt{ab} , the geometrical mean.

To complete the proof of the am-gm inequality just note that Z lies on the circle from figure 9 (see figure 11). (A standard theorem of Euclidean geometry says that this is true for any right triangle that has XY as its hypotenuse.) Since WZ is clearly no longer than OU (and is shorter iff $a \neq b$), $\sqrt{ab} \leq (a + b)/2$.

Figure 11:



This proof is explanatory. And it uses geometry to prove an algebraic statement. But one might say that the geometry is not itself doing any of the explaining, and try to back up this claim by translating the proof into a purely algebraic proof. If the proof loses no explanatory power in translation that is good evidence that the geometrical facts are not doing any explaining.

Here's one translation. Consider the function $f(x) = \sqrt{(\frac{a+b}{2})^2 - x^2}$, defined on the interval $(-\frac{a+b}{2}, \frac{a+b}{2})$. (The graph of f is a semi-circle centred at the origin with radius $(a + b)/2$.) It is easy to see (even without picturing the graph of f) that f has a single maximum, at $x = 0$; there its value is $f(0) = (a + b)/2$. But $f((a - b)/2) = \sqrt{ab}$, so $\sqrt{ab} \leq (a + b)/2$ with equality if $a = b$.

If one thought that this algebraic proof was also explanatory then one could maintain that the real explainers in the geometrical proof are certain algebraic facts. One could say that the geometry is there just to make it easier for us to grasp those algebraic facts and to make it easier for us to see which algebraic statements imply which others. The one proof uses a circle, the other a function whose graph is (half of) that circle. In some contexts a function is easier to grasp by looking at its graph than by looking at an algebraic definition. Maybe that is all that is going on here. (I myself am not sure whether the algebraic proof is, considered on its own, explanatory.)

Appendix B: Another Example of an Explanatory Physical Argument

Inside every ellipse are two points F and G , the foci, with the property that for any point P on the ellipse the sum of the lengths of the segments FP and GP is the same. Now ellipses have the following interesting property: for any point P on the ellipse, the angle that FP makes with the tangent at P is equal to the angle GP makes with the tangent at P (see figure 12). This is the ‘reflective’ property of ellipses. (Light reflects from mirrors so that the angle of incidence is equal to the angle of reflection. So If you and I stand at the two foci of an ellipse made of mirrors and I aim a flashlight in *any* direction (in the plane of the ellipse), the light will (after one reflection) shine on you.) Why do ellipses have this reflective property?

Here is an argument that relies on physical principles. Hammer two nails straight into a piece of wood and tie a string of length l between them. Orient the wood so that the nails lie in the horizontal plane. Put a frictionless pulley on the string and hang a mass from the pulley. If we keep the string taut while moving the pulley the pulley will trace out an ellipse. The lowest point on that ellipse is the point where the pulley will hang at rest (see figure 13). (The pulley can be made to rest on any point of the ellipse just by rotating the board so that that point is the lowest point of the ellipse.) The tangent to the point P where the pulley rests is the horizontal line through P .

Since the pulley is at rest the net force on it is zero. The two string segments

Figure 12: Reflective property of the ellipse.

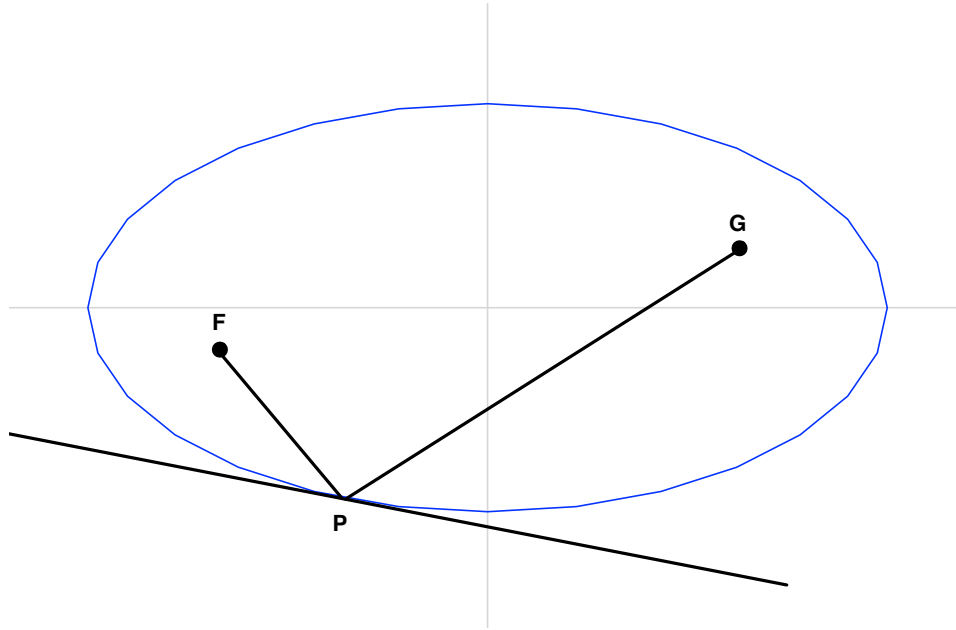
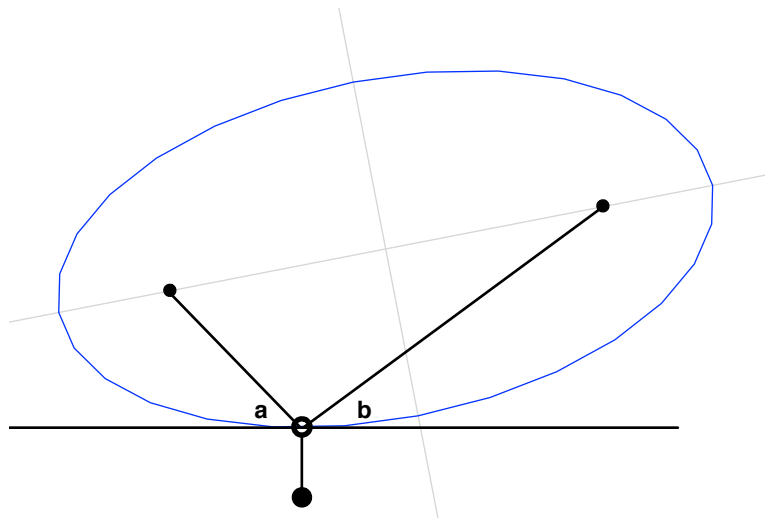


Figure 13: Diagram for the mechanical argument



pull on the pulley with forces of the same magnitude T . (The magnitudes must be the same: otherwise, since the pulley is frictionless, it would spin until they were equal.) The horizontal components of these forces are the only horizontal forces acting on the pulley. So they must be equal in magnitude. That is, $T \cos a = T \cos b$, where a and b are the angles the string segments make with the horizontal. Since both angles must be less than 180° , it follows that $a = b$, which is what was to be shown.

Levi ([2009], p. 30) contrasts this physical argument with a purely mathematical proof that requires ‘finger-breaking calculation.’ He says that ‘little understanding would be gained’ by performing the calculation but the physical argument shows ‘what’s going on.’ I agree: this argument, like the argument for the Pythagorean theorem, is explanatory.²⁹ And I think the explanations have the same kind of structure. We have a counterfactual: if ellipses did not have the reflective property then the pulley would not hang at rest. And the fact that the pulley does hang at rest is independently grounded, in the physical laws true of the situation.

Bradford Skow
Dept. of Linguistics and Philosophy
Massachusetts Institute of Technology
77 Massachusetts Ave., 32-D808
Cambridge, MA 02139 USA
bskow@mit.edu

Acknowledgements

Thanks to Steve Yablo, Agustin Rayo, Brendan Dill, everyone else at MIT, Kenny Easwaran, Jaime Tappenden, Chris Pincock, and the audience at the 2012 PSA

²⁹This argument also appears in (Uspenskii [1961], pp. 6-9). It is not true that *all* mathematical proofs of this theorem require finger-breaking calculation. See for example the short proof using vector analysis in (Flanders [1968]).

meeting. And thanks especially to my referees for their incredibly insightful comments. They forced me to re-think a lot of the things I say.

References

Arana, A. and Mancosu, P. [2012]: 'On the Relationship Between Plane and Solid Geometry', *The Review of Symbolic Logic*, **5**, pp. 294-353.

Archimedes [2009]: *Geometrical Solutions Derived From Mechanics*, J. L. Heiberg (trans), Chicago: Open Court.

Baker, A. [2005]: 'Are there Genuine Mathematical Explanations of Physical Phenomena?', *Mind*, **114**, pp.223-38.

Dawson, Jr., J. W. [2006]: 'Why do Mathematicians Re-prove Theorems?', *Philosophia Mathematica*, **14**, 269-86.

Detlefsen, M. and Arana, A. [2011]: 'Purity of Methods', *Philosophers' Imprint*, **11**.

Field, H. [1980]: *Science Without Numbers*, Princeton: Princeton University Press.

Flanders, H. [1968]: 'The Optical Property of the Conics', *The American Mathematical Monthly*, **75**, p. 399.

Hawthorne, J. and Nolan, D. [2006]: 'What Would Teleological Causation Be?', in J. Hawthorne, *Metaphysical Essays*, New York: Oxford University Press, pp. 265-84.

Hempel, C. [1965]: *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: The Free Press.

- Kogan, Y. [1974]: *The Application of Mechanics to Geometry*, Chicago: University of Chicago Press.
- Lange, M. [2010]: ‘What Are Mathematical Coincidences (and Why Does It Matter)?’, *Mind*, **119**, pp. 307-40.
- Levi, M. [2009]: *The Mathematical Mechanic*, Princeton: Princeton University Press.
- Lewis, D. [1986a]: ‘Counterfactual Dependence and Time’s Arrow’, in *Philosophical Papers, Volume II*, New York: Oxford University Press, pp. 32-66.
- Lewis, D. [1986b]: ‘Causal Explanation’, in *Philosophical Papers, Volume II*, New York: Oxford University Press, pp. 214-40.
- Lewis, D. [1998]: ‘Statements partly about observation’, in *Papers in Philosophical Logic*, New York: Cambridge University Press, pp. 125-55.
- Lipton, P. [2009]: ‘Understanding Without Explanation’, in H. W. De Regt, S. Leonelli, and K. Eigner (eds), *Scientific Understanding: Philosophical Perspectives*, Pittsburgh: University of Pittsburgh Press, pp. 43-63.
- Needham, T. [1997]: *Visual Complex Analysis*, New York: Oxford University Press.
- Nolan, D. [1997]: ‘Impossible Worlds: A Modest Approach’, *Notre Dame Journal of Formal Logic*, **38**, pp. 535-72.
- Rosen, G. [2010]: ‘Metaphysical Dependence: Grounding and Reduction’, in B. Hale and A. Hoffman (eds), *Modality: Metaphysics, Logic, and Epistemology*, New York: Oxford University Press, pp. 109-35.
- Spivak, M. [2008]: *Calculus*, 4th edition, Houston: Publish or Perish.

Steiner, M. [1978]: 'Mathematics, Explanation, and Scientific Knowledge', *Nous*, **12**, pp. 17-28.

Tappenden, J. [1995]: 'Extending Knowledge and 'Fruitful Concepts': Fregean Themes in the Foundations of Mathematics', *Nous*, **29**, pp. 427-67.

Uspenskii, V. A. [1961]: *Some Applications of Mechanics to Mathematics*, New York: Blaisdell Publishing Company.

Van Fraassen, B. C. [1980]: *The Scientific Image*, New York: Oxford University Press.

Wilson, M. [2006]: *Wandering Significance: An Essay on Conceptual Behavior*, New York: Oxford University Press.

Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.

Yablo, S. [2002]: 'Abstract Objects: A Case Study', *Philosophical Issues*, **12**, pp. 220-40.