The Mathematical Route to Causal Understanding

Michael Strevens

To appear in *Explanation Beyond Causation*,
edited by Alexander Reutlinger and Juha Saatsi

In some scientific explanations, mathematical derivations or proofs seem to be the primary bearers of enlightenment. Is this a case, in science, of "explanation beyond causation"? Might these explanations be causal only in part, or only in an auxiliary way, or not at all? To answer this question, I will examine some well-known examples of explanations that seem to operate largely or wholly through mathematical derivation or proof. I conclude that the mathematical and the causal components of the explanations are complementary rather than rivalrous: the function of the mathematics is to help the explanations' consumers better grasp relevant aspects of the causal structure that does the explaining, and above all, to better grasp how the structure causally makes a difference to the phenomena to be explained. The explanations are revealed, then, to be causal through and through.

It does not follow that all scientific explanation is causal, but it does follow that one large and interesting collection of scientific explanations that has looked non-causal to many onlookers in fact fits closely with the right kind of causal account of explanation. In that observation lies my contribution to the present volume's dialectic.

1. Mathematics Gives Us the Gift of Scientific Understanding

Heat a broad, thin layer of oil from below, and in the right circumstances, Rayleigh-Bénard convection begins. At its most picturesque, the convecting fluid breaks up into many hexagonal convection cells, taking on the appearance of a honeycomb. Why that particular shape? An important part of the explanation, it seems, is that the densest possible lattice arrangement of circles in two dimensions is the hexagonal packing: for unrelated reasons, the fluid forms small, circular convection cells; these cells then distribute themselves as densely as possible and fill the interstitial spaces to take on the hexagonal aspect.

The explanation of the honeycomb structure has many parts: the explanation of circular convection cells; the explanation of their tendency to arrange themselves as densely as possible; the explanation of their expanding to fill the interstitial spaces. One essential element among these others is, remarkably, a mathematical theorem, the packing result proved by Lagrange in 1773. To understand the honeycomb structure, then, a grasp of the relevant causal facts is not enough; something mathematical must be apprehended.

* * *

Northern elephant seals have extraordinarily little genetic diversity: for almost every genetic locus that has been examined, there is only one extant allele (that is, only one gene variant that can fit into that genetic "slot"). The reason, as is typical in such cases, is that the seals have recently been forced through a "population bottleneck". In the late nineteenth century, they were hunted almost to extinction; as the population recovered, it was extremely small for several decades, and in populations of that size, there is a high probability that any perfectly good allele will suffer extinction through simple bad luck—or as evolutionary biologists say, due to random genetic drift.

To explain the genetic homogeneity of contemporary Northern elephant seals, you might in principle construct a real-life seal soap opera, first relating the devastation caused by hunting death after death, and then the rebuilding of the population birth after birth, tracking the fate of individual alleles as the seals clawed their way back to the numbers they enjoy today. But even if such a story should be available—and of course it is not—it would be no more explanatory, and some would say less explanatory, than a suitably rigorous version of the statistical story told above, in which what is cited to explain homogeneity is not births and deaths or even the extinction of individual alleles, but rather the impact of population size on the probability of extinction (and then, not the precise change for any particular allele but just the general trend, with the probability of extinction increasing enormously for sufficiently small populations). The derivation of the fact of this impact takes place entirely within the mathematics of probability theory. Though the explanation also has causal components, it seems to revolve around the mathematical derivation.

* * *

Consider an unusually shaped container—say, a watering can with all openings closed off. Inside the container is a gas, perhaps ordinary air. How does the gas pressure vary throughout the container after the gas is left to "settle down", that is, after the gas reaches its equilibrium state? The answer is not obvious. Gas pressure is caused by a gas's molecules pounding on a container's surfaces. Perhaps the pressure is lower in the neck of the watering can, where there is much less gas to contribute to pressure over the available surface area? Or perhaps it is higher, because at any given moment more of the gas in the can's neck than in its main body is close to a surface where it can contribute to the pressure?

Assume that at equilibrium, the gas is evenly distributed through the container, so that the density does not vary from place to place, and that the average velocity of gas molecules is the same in each part—a conclusion that it is by no means easy to derive, but the explanation of which I bracket for the sake of this example. Then a short mathematical derivation—essentially, the backbone of the explanation of Boyle's law—shows that the pressure in the container is the same everywhere. The key to the derivation is that the two factors described above exactly cancel out: there are many more gas molecules in the main section of the watering can, but proportionally more of the molecules in the neck are at any time within striking distance of a surface. The net effect is equal numbers of "strikes" on every part of the can's—or any container's—surface. This canceling out is, as in the case of the elephant seals, displayed by way of a mathematical derivation. Mathematics, then, again sits at the center of a scientific explanation.

* * *

An example used to great effect by Pincock (2007) begins with a question about the world of matter and causality: why, setting out on a spring day to traverse the bridges at the center of the city of Königsberg without crossing any bridge twice, would Immanuel Kant fail by sunset to accomplish this task? (The rules governing the attempt to trace what is called an Eulerian path are well known: the path must be continuous and rivers may be crossed only using the bridges in question. You may start and finish anywhere you like, provided that you cross each bridge once and once only.)

The explanation of Kant's failure is almost purely mathematical: given the configuration of the bridges, it is mathematically impossible to walk an Eulerian path. For any such problem, represent the bridges (or equivalent) as a graph; an Eulerian path exists, Leonhard Euler proved, only if the number of nodes in the graph with an odd number of edges is either two or zero. The graph for the Königsberg problem has four odd-edged nodes.

We could explain Kant's lack of success by enumerating his travels for the day, showing that no segment of his journey constitutes an Eulerian path. But that explanation seems quite inferior to an explanation that cites Euler's proof. Perhaps more clearly than in any of the cases described above, this explanation of a material event turns on a mathematical fact, the proof of which is essential to full understanding.

2. The Role of Mathematics in Scientific Explanation

How, then, does mathematics convey understanding of the hexagonal structure of Rayleigh-Bénard convection cells, of the genetic homogeneity of Northern elephant seals, of the uniform pressure of gases at equilibrium regardless of container shape, of the persistent failure of sundry flâneurs' attempts to traverse, Eulerianly, the bridges of Königsberg? Why, in particular, is it so tempting to say, in each of these cases, that the phenomenon in question holds *because* of such and such a mathematical fact—that a convection pattern is hexagonal because of Lagrange's theorem, or that an attempt on the bridges fails because of Euler's theorem—a locution that seems to place mathematical facts at the heart of certain scientific explanations?

Galileo famously suggested in *The Assayer* that "The Book of Nature is written in mathematical characters". The Book of Nature is the physical world; this metaphor suggests, then, that mathematics is embedded in nature itself. In that case, perhaps, mathematical properties could explain physical states of affairs by way of mathematical necessitation. I call this notion—it is too nebulous to be called a thesis—the Galilean view of the role of mathematics in explanation.

The Galilean view might be fleshed out in many ways. You might, for example, attribute to abstract mathematical objects—say, the number three—causal powers. Then mathematical necessitation could be understood as a kind of causation, and the examples of mathematical explanation given above as causal. That is not, however, a popular view.

Another possibility runs as follows. Consider a law of nature of the sort usually supposed to describe the effects of causal influence, such as Newton's second law (never mind that it has been superseded): $F = ma$. The law tells us how an object's position changes as a consequence of the total impressed force. On the Newtonian worldview, force is doing something in the world: it is making changes in objects' positions. Both force and position are physical rather than mathematical properties, so there is no mathematical causation here. But could mathematics shape the channels through which causal influence flows? What I have in mind is that there is mathematical structure in the physical world and that causation operates through this structure—in the Newtonian case, for example, perpetrating its effects along lines inscribed in reality by the mathematics of real-numbered second-order differential equations.

If something like this were the case (and of course I have offered only the barest sketch of what that might mean), then it would be no surprise to find mathematics at the core of scientific explanation, dictating the ways in which physical processes may or may not unfold. When we say that a phenomenon obtains *because* of some mathematical fact—say, that no traversal of the bridges can occur *because* of Euler's theorem—we would mean it literally. It is not that Euler's theorem is itself a cause (any more than Newton's second law is "a cause"), but rather that it exhibits a mathematical fact that plays a direct and essential role in the unfolding of the causal processes that constitute attempts at an Eulerian path, a fact that participates in the causal story in a raw and unmediated way, and so whose nature must be grasped by anyone hoping to understand the story.

It is not the strategy of this paper, however, to defend the causal approach to scientific explanation by upholding a Galilean view. Rather, I will assume a contrary and rather deflationary thesis about the role of mathematics in science—the representational view—and show how, on that view, the examples of mathematically driven scientific explanation cited above ought to be interpreted. I assume the representational view partly out of an inkling that it may be correct, though I will not argue for such a conclusion here, and partly as a matter of rhetorical strategy, since it allows me to demonstrate that even if, as the representational view implies, there is no prospect whatsoever of mathematical properties playing a role in causation, mathematically driven explanations may nevertheless be understood as wholly causal.

According to the representational view there is either no mathematics in the natural world or mathematics exists in nature in an entirely passive, hence non-explanatory, way. (As an example of the latter possibility, consider the thesis that numbers are sets of sets of physical objects; it follows that they have a physical aspect, but they make up a kind of abstract superstructure that does not participate in the causal and thus the explanatory economy as such.) The role of mathematics in science, and more specifically in explanation, is solely to represent the world's non-mathematical explanatory structure—to represent causes, laws and the like. A knowledge of mathematics is necessary to understand our human book of science, then, but it is not the content but rather the language that is mathematical. God does not write in mathematical characters—not when she is telling explanatory stories, at least—but we humans, attempting to understand God's ways, represent her great narrative using representational tools that make use of mathematical structures to encode the non-mathematical explanatory facts.

Such a view is suggested by two recent theories of the role of mathematics in science, the mapping account of Pincock (2007) and the inferential account of Bueno and Colyvan (2011). According to both theories, mathematics plays a role in explanation by representing the non-mathematical facts that do the explaining, in particular, facts about causal structure.[1]

Can the representational view capture the way in which my example explananda—hexagonal convection cells, elephant seal homozygosity, constant gas pressure—seem to depend on certain mathematical facts? Can they gloss the sense in which the bridges are untraversable *because of* Euler's theorem? The best sense that a representationalist can make of such talk is, I think, that the "because" is figurative: a state of affairs obtains because some non-mathematical fact obtains, and that non-mathematical fact is represented by the mathematical fact, which in a fit of metaphor we proffer as the reason.

1. Other work by Pincock and Colyvan—for example, Pincock (2015)—suggests that these authors may not hold that the mapping and inferential accounts (respectively) exhaust the role of mathematics in science. I take a certain view of the scientific role of mathematics from these authors, then, and to obtain what I call the representational view, I append *And that is all*.

There is something non-mathematical about the bridges of Königsberg that renders them untraversable; that non-mathematical fact is represented by Euler's theorem and so—eliding, conflating, metonymizing—we say that the failure of any attempt at traversal is "because of" Euler's theorem itself.

If that were all that the representationalist had to say about mathematical explanations in science, this paper would be short and uneventful. But there is another striking aspect of these explanations besides the "because of", that on the one hand poses a greater challenge to representationalism, and on the other hand to which the representationalist can give a far more interesting reply. This will be my topic for the remainder of the paper.

The challenge turns from the "because of", with its implication of an apparent explanatory relation between a mathematical and a physical fact, to the role of mathematical thinking in understanding. When I try to understand what is going on with the bridges or the elephant seals, it seems that thinking mathematically gives me a kind of direct insight into the relevant explanatory structure. Where does that insight come from?

One answer that the representationalist is well positioned to provide is: resemblance. On Pincock's mapping view, for example, the mathematical structures that feature in explanations are isomorphic to explanatorily relevant structures in the physical world. Grasp the mathematical structure and you grasp the physical structure, at an abstract level at least.

This answer is a good one, but it does not go far enough. Often I gain the majority of my explanatory insight from seeing a mathematical derivation or proof. In many cases, such proofs do relatively little to help me to grasp those aspects of mathematical structure that mirror explanatory structure. The isomorphism between the layout of the city of Königsberg and the corresponding graph (figure 1) is obvious; the Euler proof does not make it any clearer—indeed, it simply presupposes it. Much the same can be said for my other paradigm cases.[2]

2. The treatment in the main text is a little quick, in a way that will become clearer when

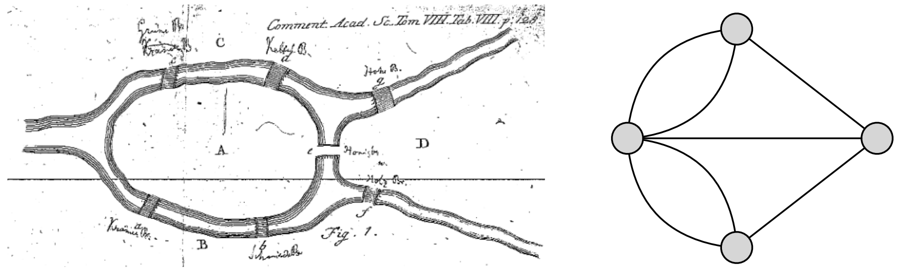

Figure 1: Left: Euler's sketch of the bridge layout. Right: a graph representing the bridges. (Source of Euler figure: MAA Euler Archive, http://eulerarchive.maa.org.)

The role of proof (or derivation) in these explanations is better described in this way: by following the proof, I see how the mathematical facts necessitate, and so explain, certain physical facts. The Galilean is in a superb position to give this gloss; the representationalist not at all.

To sum up, then, representationalism can go a certain distance in making sense of the role of mathematics in my paradigm explanations. It can to some extent explain away "because" talk, and it can to some extent explain how grasping mathematical structure helps us to grasp explanatory structure. But it does not make very good sense, apparently, of the way in which grasping mathematical proofs helps us to understand physical phenomena. In this respect, the Galilean approach is far superior. That is the challenge to representationalism that I hope to meet in what follows, accounting in representationalist terms for the power of mathematical proof to provide us, in the paradigms above and in other such cases, with causal understanding.

I present my approach to causal explanation later in this paper. In the main text, I have taken the aspects of the city layout represented by the Königsberg graph to be the relevant explanatory structure. In fact, the explanatory structure is more abstract than this; it is the fact about the city layout represented by the graph's having more than two odd-edged nodes. The critique holds, however: whatever the proof does, it goes well beyond helping us to see more clearly that both the city plan and the graph have this property.

3. Explanatory Relevance as Causal Difference-making

Scientific explanation, according to the approach I will adopt in this paper without argument, is a matter of finding causal difference-makers. Much of what I want to say could be framed in terms of any sophisticated difference-making account, but I will—naturally—rely here on my own "kairetic account" (Strevens 2004, 2008).

The raw material of explanation, according to the kairetic account, is a fundamental-level causal structure revealed by physics. For simplicity's sake, assume that we live in a classical world constituted entirely of fundamental-level particles that interact by way of Newtonian forces. Then the fundamental-level causal structure is the network of force, that is, the totality of forces exerted, at each instant of time, by particles on one another, whether gravitational, electromagnetic, or something else. This web of force, together with other relevant facts about the particles—their positions, their velocities, their inertial masses, and so on—determines each particle's movements, and so determines everything that happens in the material world. That, at least, is the Newtonian picture, which I assume here; modern physics of course requires some revisions.[3]

The web of force is vast and dense, titanic and tangled; it is beyond the power of human science to represent any significant part of it explicitly and exactly. Were scientific explanation to require us to provide an exhaustive inventory of the forces acting on a particle at any given time—an inventory that would include, for a particle of non-zero mass, the gravitational influence of every other massive particle in the universe—we would have no prospect whatsoever of constructing complete explanations.

We aspiring explainers avoid such impossible demands, because we are interested in explaining mainly high-level events and states of affairs, and

3. I will not countenance here the possibility that modern physics will show the world to be devoid of causality, or the milder but still alarming possibility that causality might only "emerge" at levels higher than that of fundamental physics.

explanation requires that we identify only the aspects of the causal web that make a difference to whether or not those events occurred or those states of affairs obtained—which difference-makers are far more sparse than causal influences.

To fill out this picture, consider event explanation in particular. With a rebel yell, Sylvie hurls a cannonball at the legislature's prize stained-glass window; it shatters. What explains the shattering? In asking this question, I am interested in why the window shattered rather than not shattering. The explainers I have in mind are the ball's hitting the window, Sylvie's throwing the ball, the window's composition—and not much more. I could have asked a different question: why did the window shatter in exactly the way that it did, with this shard traveling in this direction at this velocity and so on? To answer such a question I would have to take into account many more casual influences—many more Newtonian forces—that acted on the shattering. Sylvie's yell, for example, caused the window to vibrate a little, which accounts in part for the exact trajectories of the myriad shards.

The contrast between these two questions—the question of why the window broke, and the question of why the window broke in precisely the way that it did—illustrates the difference between a high-level level event such as the breaking and the low-level or "concrete" event that realizes the breaking, that is, the window's breaking in precisely such and such a manner, specified down to the most minute details of each molecule's trajectory. Because explanation is about finding difference-makers, an answer to the latter question must cite pretty much every causal influence on the window, while an answer to the former question ignores elements of the causal story whose only impact is on *how* the window broke, and focuses instead on those elements that made a difference to *whether* the window broke. Sylvie's insurrectionary cry made a difference to the precise realization of the window's shattering, and helps to explain that concrete event, but it made no difference to whether not the window shattered; it thus plays no part in explaining the high-level event of the shattering.

Science's explanatory agenda is focused almost exclusively on high-level events as opposed to their concrete realizers. Biologists want to explain why humans evolved large brains, but they are not (on the whole) interested in accounting for the appearance of every last milligram of brain tissue, except insofar as it casts light on the bigger question. Planetary scientists would like to explain the formation of the solar system, but they certainly have no interest in explaining the ultimate resting place of individual pebbles. Economists are interested in explaining why the recent financial crisis occurred, but they are not (on the whole) interested in explaining the exact dollar amount of Lehman Brothers' liabilities. In each case, then, the would-be explainers must decide which elements of the causal web, the densely reticulated network of influence responsible for all physical change, were significant enough to make a difference to whether or not the phenomena of interest occurred—to the fact that human brains grew, that the solar system took on its characteristic configuration, that between 2007 and 2008 the global financial system warped and fractured.

The role of a theory of explanation is to provide a criterion for difference-making that captures this practice—that classifies as difference-makers just those aspects of the causal web that are counted as such by scientific explainers. An obvious choice is a simple counterfactual criterion: a causal influence on an event is an explanatory difference-maker for the event just in case, had it not been present, the event would not have occurred. Had Sylvie thrown the cannonball without a word, the window would still have broken, so her vocal accompaniment is not a difference-maker for the breaking. But had she not thrown the cannonball at all, the window would have remained intact; thus, her throwing is a difference-maker. As is well known from the literature on singular causation, however, the counterfactual criterion fails to capture our judgments of difference-making, both everyday and scientific: Bruno might have been standing by to break the window in case Sylvie failed; in these circumstances, it is not longer true that had Sylvie refrained from throwing, the window would not have broken (Lewis 1973). The counterfactual criterion counts her throw as a non-difference-maker in Bruno's presence, but we want to say that, since Sylvie did in fact throw the ball and her ball broke the window, her throw was a decisive difference-maker for the breaking.

In the light of these and other problems for the counterfactual approach (Strevens 2008, chap. 2), I have proposed an alternative criterion for difference-making. The "kairetic criterion" begins by supposing the existence of a complete representation of the relevant parts of the causal web, that is, a complete representation of the causal influences on the event to be explained. This representation takes the form of a deductive argument in which effects are deduced from their causes along with causal laws. In the case of the window, for example, the trajectory of each shard of glass will be deduced from the relevant physical laws and initial conditions—the trajectory and makeup of the incoming cannonball, the molecular constitution and structure of the window and its connection to its frame, and all other relevant environmental circumstances, including in principle the gravitational influence of the distant stars. Such a deduction shows how the breaking in all of its particularity came about as the aggregate result of innumerable causal influences; it is a representation of the complete causal history of the breaking.

A few comments on this canonical representation of the causal process leading to the breaking. First, it is of course quite beyond the powers of real scientists, even very well-funded and determined real scientists, to construct such a representation. The canonical representation's role is to help to lay down a definitive criterion for causal difference-making; in practice scientists will decide what is likely to satisfy the criterion using a range of far more tractable heuristics. To give a simple example, the gravitational influence of other stars almost never makes a difference to medium-sized terrestrial events such as window breakings; the stars can therefore from the outset be ignored.

Second, there is much more to say about the structure in virtue of which the canonical representation represents a causal process. I have said some of it in Strevens (2008) chap. 3; in the current paper, however, the details are of little importance. I will simply assume that there is some set of conditions in virtue of which a sound deductive argument represents a causal process or, as I will say, in virtue of which it qualifies as a veridical causal model.

Third, in assuming that the explanandum can be deduced from its causal antecedents, I am supposing that the process in question is deterministic. In the stochastic case, what is wanted is rather the deduction of the event's probability, as suggested by Railton (1978). Again, I put aside the details; for expository purposes, then, assume determinism.

On with the determination of difference-makers. The idea behind the kairetic account is simple: remove as much detail as you can from the canonical representation without breaking it, that is, without doing something that makes it no longer a veridical causal model for the event to be explained. The "removal" consists in replacing descriptions of pieces of the causal web with other descriptions that are strictly more abstract, in the sense that they are entailed by (without entailing) the descriptions they replace and that they describe the same subject matter or a subset of that subject matter.[4]

In the case of the broken window, for example, much of the structure of the cannonball can be summarized without undermining the veridicality or the causality of the canonical model. What matters for the deduction is that the ball has a certain approximate mass, size, speed, and hardness. The molecule by molecule specification of the ball's makeup that appears in the canonical representation can be replaced, then, by something that takes up only a few sentences. Likewise, the fact of Sylvie's war cry can be removed altogether, by replacing the exact specification of her vocalization with a blanket statement that all ambient sound was within a certain broad range (a range that includes

4. The removal operation is constrained additionally by a requirement that the representation should remain "cohesive", which ensures that abstraction does not proceed by adding arbitrary disjuncts. Cohesion is relevant to some aspects of the following discussion, but for reasons of length I will put it aside.

almost any ordinary noises but excludes potential window-breakers such as sonic booms).

When this process of abstraction has proceeded as far as possible, what is left is a description of the causal process leading to the explanandum that says as little about the process as possible, while still comprising a veridical causal model for the event's production. The properties of the process spelled out by such a description are difference-making properties—they are difference-makers for the event. The approximate mass, size, speed, and hardness of the cannonball make a difference to the window's breaking, then, but further details about the ball do not. Nothing about Sylvie's yell makes a difference except its not exceeding a certain threshold. These difference-makers are what explain the window's breaking; aspects of the causal web that do not make a difference in this sense, though they may have affected the event to be explained—determining that this shard went here, that one there—are explanatorily irrelevant.

Observe that the kairetic account envisages two kinds of causal relation. One kind is causal influence, which is revealed by the correct fundamental-level theory of the world and serves as the raw material of causal explanation. The other kind is causal difference-making, an explanatory relation that links various properties of the web of influence to high-level events and other explananda. Difference-making relations are built from causal influence according to a specification that varies with the phenomenon to be explained.

The cases of mathematically driven understanding presented above, you will note, involve high-level difference-making, in which many prima facie causally significant features of the setup turn out not to be difference-makers: the development of particular convection cells, the shape of particular containers, the twists and turns taken in an attempt to travel an Eulerian path around Königsberg. That is an important clue to what mathematics is doing for us, as you will shortly come to see.

My goal is to show that mathematically driven explanations in science are causal, in the manner prescribed by the kairetic or some other difference-making account. My working assumption is that the role of mathematics in science, including explanation, is purely representational, standing in for inherently non-mathematical features of nature. If mathematics is an aid to scientific explanation, then, its assistance had better be indirect, arriving in virtue of something that it does as a representer of causal structure (though not necessarily representation simpliciter). To see what that something might be, I turn to the topic of understanding.

4. Mathematics and Causal Understanding

According to what I have elsewhere dubbed the "simple view" of understanding, to understand a phenomenon is to grasp a correct explanation for that phenomenon (Strevens 2013). Combining the simple view with the kairetic account of scientific explanation yields the following thesis: to understand a material event is to grasp the difference-making structure in which that event is embedded and in virtue of which it occurred. The explanation, on this approach, is "out there": it is a collection of causal facts—causal difference-making relations, to be precise—waiting to be discovered by science. Understanding is the cognitive achievement realized by epistemically connecting to the explanation, to these facts, in the right way.

The philosophy of understanding is much concerned with what counts as the "right way". Is it a matter of having deep knowledge of the relevant domain or is it rather a matter of possessing some ability that goes beyond mere knowledge? In this paper I will keep my distance from these debates, assuming that at a minimum, in order to grasp a causal difference-making structure a seeker of understanding must grasp both the nature of the difference-makers for the explanandum and the way in which they make the difference that they do.

What role can mathematics play in all of this? Difference-making structures are not inherently mathematical—that, at least, is my representationalist working assumption—but mathematics might nevertheless help us to get a grip on such a structure, attaining the kind of epistemic connection to the difference-makers and their difference-making that constitutes understanding. Here's how.

A part of grasping an explanation is to apprehend clearly the topology of the relevant relations of causal difference-making. Mathematics is often central to this task, transparently and concisely representing the structure to be grasped, whether by way of a directed graph (the formal equivalent of boxes and arrows), a set of differential equations, a stochastic dynamical equation, or in some other way.

In performing this function, mathematics does just what the representationalist claims: it provides compact, precise, in many cases tailor-made symbolic systems to represent relationships in the world. As I argued in section 2, however, this representational role, vital though it may be, does not cast much light on the importance of mathematical derivation or proof in causal understanding. A system of definitions seems to be sufficient to undergird a system of representation; theorems derived from those definitions add no representational power and in many cases make the representation no more effective than it was before. The graph representing the bridges of Königsberg presents the essential structure of the problem just as plainly and perspicuously to someone who does not know of Euler's proof as to someone who does. Yet understanding the proof seems absolutely central to understanding why Kant failed to complete an Euler walk around the bridges on some fine day in May.

To appreciate the function of proof, we need to turn to another facet of causal understanding. Knowing the difference-makers is not enough, I submit, to grasp an explanation; you must understand *why* they are difference-makers, or in other words, you must grasp the reasons for their difference-making status.

Go back to the cannonball through the window, to begin with a very simple case. What does it take to understand why the window shattered? You must, at the very least, grasp the fact that the ball's hitting the window caused it to shatter—that the striking was a causal difference-maker for the shattering. But there is more to explanation and understanding than this. It is also important to see in virtue of what aspects of the situation the ball caused the shattering, that is, to see how and why it was a difference-maker. In part this is a matter of grasping (at the appropriate level of abstraction) the structure of the underlying causal process: the transfer of momentum to parts of the window; the stress thereby placed on the bonds holding the window together; the catastrophic failure of the bonds due to their inelasticity.

Equally, it is a matter of seeing that these elements of the causal web were sufficient in themselves to bring about the breaking, that they and nothing else (aside from their own causes, such as Sylvie's throwing) were the difference-makers for the breaking. This insight comes most directly and also most deeply through an application of the kairetic criterion, that is, through seeing that it is possible to abstract away from all other properties of the web while still deriving the fact of the shattering. And mathematical proof is the royal road to this goal: a proof, once fully understood, shows us with unrivaled immediacy what is and is not required for a derivation. In so doing, it shows us why difference-makers satisfy the criterion for difference-making—why they are difference-makers.

The proof, in short, because it is not part of the difference-making structure, is not a part of the explanation. Its role is not to explain but to help us to grasp what explains—to see the difference-makers for what they are—and so to help us to understand.

* * *

Let me now return to the examples of mathematically driven understanding that I presented above: hexagonal Rayleigh-Bénard convection cells, genetic uniformity in elephant seals, the irrelevance of container shape to gas pressure, and the bridges of Königsberg. In each of these cases, I suggest, the value of mathematical proof lies in its helping us to grasp which aspects of the great causal web are difference-makers for the relevant explanandum and why—and complementarily, helping us to grasp which aspects of the web are not difference-makers and why. What makes these particular examples especially striking, and the underlying mathematical proofs especially valuable, is that there are many important-looking parts of the causal story that turn out, perhaps contrary to initial expectations, to be non-difference-makers. The mathematics shows us why, in spite of their substantial causal footprint, they make no difference in the end to the phenomenon to be understood.

Consider the elephant seals. Large numbers of seal alleles went extinct in a short time, but the extinction had nothing to do with the intrinsic nature or developmental role of those alleles. They simply suffered from bad luck—and given the small size of the seal population in the early twentieth century, it was almost inevitable that bad luck would strike again and again, eviscerating the gene pool even if the species as a whole endured. The mathematics reveals, then, that the extinction of so many alleles was due to a haphazard mix of causal processes—mostly to do with mating and sex (though also including death by accident and disease)—whose usual aleatory effect on the makeup of the gene pool was powerfully amplified by the small size of the population, wiping out almost all the elephant seals' genetic diversity.

The mass extinction of seal alleles has a causal explanation, then—a highly selective description of the operation of the relevant part of the causal web, that is, the ecology of the Northern elephant seal over several decades. To see that this is the correct explanation, however—to see that in spite of its high-level of abstraction, its omission of so much that seems important, it contains all the explanatorily relevant factors, all the difference-makers—mathematical thinking is invaluable. It is the mathematics that enables you to see both how cited factors such as mating choice and sex, not normally regarded as indiscriminate extinguishers of biological diversity, erased so many alleles, and why as a consequence many uncited factors better known for their selective power, above all the various genes' phenotypic consequences, were not difference-makers at all.

Or consider gas pressure. The essence of the explanation for a gas's uniform pressure on all surfaces of its container is causal; it embraces both the causal process by which the gas spreads itself evenly throughout the container, creating a uniform density, and the process by which a gas in a state of uniform density creates the same pressure on all surfaces. As in the elephant seal case, however, the explanation has very little to say about these causal processes. It barely mentions the physics of molecular collision at all, and the container walls themselves figure in the story only in the most abstract way. The walls' shape, in particular—the geometry of the container as a whole—is conspicuous only by its omission from the explanation. Mathematics helps us to grasp this explanation by showing us why the details of collision and container shape make no difference—in effect, by showing us that uniform pressure can be derived from a description of a few abstract properties of the gas however the details are filled out.

* * *

Now let me tackle the tantalizing Königsberg case. Here, it is tempting to say, mathematics takes over from causal explanation altogether, yielding a bona fide example of the explanation of a physical fact—Kant's failure to complete an Euler walk around the Königsberg bridges on May Day, 1781—that lies entirely beyond causation. I assimilate it, nevertheless, to the other examples in this paper. The explanation of Kant's failure takes the form of a highly abstract description of the relevant piece of the causal web—that is, of his day's wanderings—that extracts just the difference-making features of the web. The role of the mathematics is not strictly speaking explanatory at all; rather, it helps us to understand why a certain ultra-abstract description of Kant's movements that day constitutes a correct explanation, that is, a description which includes all the difference-makers and therefore omits only those properties of the web that made no difference to the event to be explained.

To see this, start with a different bridge-traversal task: say, the task of visiting each of the four Königsberg land masses (two islands and the two banks of the river) exactly once—or in more abstract terms, the task of visiting each node in the corresponding graph exactly once, which in graph theory is called a Hamiltonian walk. Such a journey is possible in the Königsberg setup, but it is also possible to go wrong, choosing to traverse a bridge that takes you back to a land mass you have already visited before the walk is complete. Suppose that Kant attempts a Hamiltonian walk. He chooses a good starting point (in this case, all starting points are equally good); he travels to another node (so far, so good); but then he makes a bad decision and travels back to his starting point without visiting the other two nodes in the graph. Why did his attempt fail? He made a wrong turn. A brief explanation would simply lay out the facts that make the turn a bad one and then note that he made it all the same. The same is true for the case where he fails because he chooses a bad starting point, say the middle node in the graph shown in figure 2.

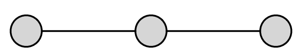

Figure 2: To complete a Hamiltonian walk of this graph, begin at one end or the other but not in the middle

A great deal is left out of these explanations. They omit everything about Königsberg except the barest facts as to the layout of its bridges and everything about Kant's means of locomotion that is not relevant to his conforming to the rules for making a graph-theoretic walk. Also omitted, most importantly, is any specification of Kant's travels after the point at which he makes a bad decision (either choosing a wrong turn or a wrong starting point). If a fatal error has already been made, these facts make no difference to his failing to complete a Hamiltonian walk, because they can be deleted from the causal story without undermining its entailment of failure.

In the case of a bad choice of starting point, then, there is no description at all of the movement from land mass to land mass (that is, from node to node); the explanation is over almost as soon as it begins, with the description of the problem, the initial bad choice of starting point, and a certain fact about the bridges: from that starting point, no Hamiltonian path can be traced. Yet, I claim, like any causal difference-making explanation, this one is a description of the relevant causal process in its entirety. It does not describe everything about that process—it leaves out the non-difference-making properties—but what it describes is present in the explanation only because it is a feature of the causal process.

Indeed, in its omission of any aspect of Kant's route after the initial choice of starting point, the explanation is not so different from, say, the explanation of genetic homogeneity in elephant seals. There, too, there is no attempt to trace a particular causal trajectory. What matters instead is a rather abstract feature of process, that it contains many events that act like random samplers of genes, and that the intensity of the sampling is such as to very likely exclude, over a certain length of time, almost every allele from the gene pool. Likewise, what matters about Kant's walk is that it is a journey carried out under a certain set of constraints (formally equivalent to a walk around a graph), that it began from a certain point, and that under these constraints, no journey beginning from that point can complete a Hamiltonian walk. The actual route taken is not a difference-maker.

From there, it is one short step to the explanation of Kant's inability to complete an Euler walk: here all possible starting points are "bad", so the identity of Kant's actual starting point is also not a difference-maker. What is left in the explanation is only generic information: the structure of the bridges and land masses and the aspects of Kant's journeying that make it formally equivalent to a walk around a graph. It is a description of a causal process—a description adequate to entail that the causal process ended the way it did, in Euler-walk failure—yet it has nothing to say about the specifics of the process, because none of those specifics is a causal difference-maker. Euler's theorem helps you to understand why.[5]

The case is very similar to another well-known example in the philosophy of explanation first brought into the conversation by Sober (1983) and then discussed extensively by (among others) Strevens (2008), namely, the explanation why a ball released on the inside lip of an ordinary hemispherical salad bowl will end up, not too long later, sitting motionless at the bottom of the bowl. The explanation identifies certain important features of the relevant causal web, that is, of the causal process by which the ball finds its way to the bowl's bottom: the downwardly directed gravitational field, the convex shape of the bowl, the features in virtue of which the ball loses energy as it rolls around. But it has nothing to say about the ball's actual route to the bottom—nothing about the starting point (that is, the point on the rim of the bowl where the ball was released), nothing specific about the manner of release, and nothing about the path traced in the course of coming to rest at the foreordained point.

The only philosophically important difference between the ball/bowl explanation and the bridges explanation is that mathematics plays a far more important role in helping us to grasp why the specified properties of the bridges setup are difference-makers and the omitted properties are not. In the case of the bowl, simple physical intuition makes manifest the irrelevance of the release point and subsequent route; in the case of the bridges, we need Euler's proof to see why Kant's choice of route makes no difference to the end result.

To sum up: ordinary causal explanations such as the cannonball and the window, equilibrium explanations such as the ball in the bowl, statistical

5. Note that the most general version of the theorem is needed to determine correctly all the difference-makers. Consider a weaker version (of no mathematical interest) that applies only to systems with an odd number of bridges. Armed only with such a theorem, you would be unable to grasp the non-difference-making status of the fact that the number of bridges is odd.

explanations such as elephant seal homozygosity and uniform gaseous pressure, and what some have taken to be purely mathematical explanations such the famous Königsberg bridges case, are all descriptions of the causal processes leading to their respective explananda, couched at a level of description where only difference-makers appear in the explanatory story. Sometimes the difference-makers entail that the system takes a particular causal trajectory, but often not—often the trajectory is specified only at a very qualitative or diffuse level, and sometimes not at all.

Mathematics has more than one role to play in the practice of explaining, but its truly marvelous uses tend to involve the application of theorems to demonstrate the explanatory power—the difference-making power—of certain abstract properties of the causal web, and even more so the lack of difference-making power of other salient properties of the web. Deployed in this way, the mathematics is not a part of the difference-making structure itself; nor does it represent that structure. Rather, it illuminates the fact that it is this structure rather than some other that makes the difference; it allows us to grasp the reasons for difference-making and non-difference-making, so bringing us epistemically closer to the explanatory facts—and thus making a contribution, if not to explanatory structure itself, then to our grasp of that structure and so to our understanding of the phenomenon to be explained.

5. Explanation Beyond Causation?

What lessons can be drawn about causal explanation? Does the spectacular use of mathematics in cases such as the elephant seals or the Königsberg bridges show that scientific explanation goes beyond causation? Even if everything I have said so far is correct, it might be maintained that the Königsberg explanation, though it has causal content, is too abstract to constitute a causal explanation. Let me consider, and repudiate, some arguments to that effect.

I begin with a recapitulation. My view that the Königsberg explanation is a causal explanation is not based on the weak and inconclusive observation that the explanatory model has some causal content. It is based on the observation that the model's sole purpose is to pick out the properties of the web of causal influence that, by acting causally, made a difference to whether or not the explanandum occurred. The model is, in other words, exclusively concerned with detailing all relevant facets of the causation of the phenomenon to be explained. It aims to do that and nothing else. If that's not a causal explanation, what is?

Objection number one: a genuine causal explanation not only lays out the causal difference-makers but also tracks the underlying causal process, whether it is a stroll around Königsberg or the trajectory taken by a ball on its way to the bottom of a salad bowl. Classify all scientific explanations, then, into two discrete categories, tracking and non-tracking. The tracking explanations not only cite causal structure but show how this structure guides an object or a system along a particular path that constitutes or results in the occurrence of the explanandum. The non-tracking explanations may cite causal structure, but they get to their explanatory endpoints not along specific paths but by other means, such as a demonstration that the endpoint is inevitable whatever path is taken. The non-tracking explanations are (according to the objection) non-causal.[6]

Such an explanatory dichotomy is, I think, indefensible. There is an enormous range of causal explanations saying more and less in various ways about the underlying causal web. The dimensions of abstraction are many, and explanations pack the space, forming a continuum of abstraction running from blow-by-blow causal tales that run their course like toppling dominoes to magical equilibrium explanations that pull the explanandum out of the causal hat in a single, utterly non-narrative, barely temporal move—and with, perhaps, a mathematical flourish. Sometimes an explanation begins narratively, like the explanation of Kant's failure to trace a Hamiltonian path

6. To make such a case for the non-causality of equilibrium explanations was Sober's aim in introducing the "ball in the bowl".

that begins with his bad decision as to a starting point, only to end quite non-narratively, with a proof that from that point on, failure was inevitable. Or it might be the other way around (if, say, the choice of starting point doesn't matter but later decisions do). Further, there are many degrees of abstraction on the way from simple narrative to magic hat. The elephant seal explanation tells a causal story of relentless extinction by random sampling, but the extinctions are characterized only at the most typological level. The gas pressure explanation is quite viscerally causal on the one hand—molecules colliding with one another and pounding on the walls of their container—yet on the other hand extraordinarily abstract, compressing heptillions of physical parameters, the positions and velocities of each of those molecules, into a few statistical aggregates. And these are only a handful of the possible routes to abstraction, each one tailor-made for a particular explanandum.

Consequently, I see no prospect whatsoever for a clear dividing line between causal tracking explanations and non-causal non-tracking explanations. The gulf between a conventional causal narrative and the Königsberg explanation is vast. But it ought not be characterized as one of causal versus non-causal character, in part because that is to suppose a dichotomy where there is a continuum of abstraction and in part because everywhere along the continuum the aim of explanation is the same: to find whatever properties of the causal web made a difference to the explanandum.

Objection number two draws the line between causal and non-causal descriptions of the web of influence in a different place, with fewer explanations on the non-causal side. The Königsberg explanation (observes the objector) is special even among very high-level, very abstract casual explanations: it deals in mathematical impossibility rather than physical or nomological impossibility. Does that difference in the guiding modality not constitute a discontinuity?

To put it another way, failure to complete an Euler walk of the Königsberg bridges is inevitable not only in universes that share our world's laws of nature.

If our physics were Newtonian, Kant could not complete the walk. Even if it were Aristotelian, he could not complete the walk. Were Kant descended from lizards rather than apes, he could not complete the walk; likewise if he were a silicon-based rather than a carbon-based life-form. The implementation of his psychology is equally beside the point: whether plotting his turns with neural matter, with digital processing, or using the immaterial thought stuff posited by dualist philosophers, he would be unable to pull off an Euler walk, for the very same reason in each case.

The explanation of Kant's failure, then, has enormous scope: it applies to many possible worlds other than our own provided that a few simple posits hold—namely, that the network of bridges has a certain structure and that the Kantian counterpart's movements are constrained so as to conform to the rules defining a graph-theoretic walk (movement is always from one node to another neighboring node along an arc).

Does that make the Königsberg explanation sui generis? It does not. Any explanatory model that abstracts to some degree from the fundamental physical laws accounts for its explanandum not only in the actual world but also in worlds whose laws differ from the actual laws solely with respect to features from which the model abstracts away. Since almost all explanatory models are abstract not only in what they say about particulars but also in what they say about the laws in virtue of which the particulars are causally connected, almost all explanatory models have a modal extent that reaches beyond the nomologically possible. The more they abstract, the wider the reach.

The Newtonian model for the cannonball's breaking the window, for example, abstracts from the exact value of the gravitational constant, implying a shattering for any value in the vicinity of the actual value—any value not so high that the cannonball thuds to the ground before it gets to the window or so low that it overshoots the window. The model thus applies to a range of broadly Newtonian theories of physics, differing in the value they assign to the constant.

More interestingly, I suggest that the simple kinetic theory of gases gives valid explanations in both classical and quantum worlds, and that the elephant seal explanation is valid for a great variety of possible biologies that depart considerably from the way things work here on Earth, in both cases because the explanatory models assume rather little about the physical underpinnings of the processes they describe. The great modal reach of the Königsberg model is, then, far from unusual. It is an exceptional case because it calls for so high a degree of explanatory abstraction, but its specialness is a matter of degree rather than of kind.

My response to both the second and the first objections, then, is to argue for a continuum (practically speaking, at least) of explanatory models in every relevant dimension, and to reject any attempt to draw a meaningful line across this continuum as invidious. Marc Lange (2013) has recently suggested a variant on the second objection that attempts to find a non-arbitrary line founded in gradations of nomic necessity.

The explanandum in question is that a double pendulum has at least four equilibrium configurations. Lange offers an explanation in the framework of Newtonian physics that he takes to be non-causal: the explanation depends on the fact that all force laws must conform to Newton's second law ($F = ma$) but on no further facts about the laws in virtue of which the pendulum experiences forces. Writing that "although these individual force laws are matters of natural necessity, Newton's second law is more necessary even than they", Lange suggests drawing the line between causal and non-casual explanations at the point that separates the force laws' physical necessity on the one hand, and the second law's higher grade of nomological necessity on the other. An explanation that depends only on this higher grade (or a grade higher still) is, he holds, non-causal. (Lange calls such explanations "distinctively mathematical", but that strikes me as a misnomer: $F = ma$ is no more mathematical than $F = GMm/r^2$; the higher necessity of $F = ma$ is nomological rather than mathematical necessity.)

Lange's view hinges on the proposition that there is something special about the line between the individual force laws and the second law. But what? It is not simply that the second law is more necessary: as I have shown above, in the space of valid scientific explanations, there is a continuum of modal strength running all the way from very particular contingent facts, to very particular facts about the actual laws of nature, to rather more abstract facts about the actual laws, and so on up to very abstract properties such as those that underwrite the kinetic theory in both classical and quantum worlds.

Why is this the particular line in modal space at which the causality "goes away"? Lange tells us, writing of the double pendulum explanation (p. 19):

> This is a non-causal explanation because it does not work by describing some aspect of the world's network of causal relations.... Newton's second law describes merely the framework within which any force must act; it does not describe (even abstractly) the particular forces acting on a given situation.

This, I think, is false. Newton's second law does describe, very abstractly, a property of the particular forces (and force laws): it says that they conform to Newton's second law. That is a fact about them. More generally, that a causal law operates (of necessity or otherwise) within a particular framework is a fact about that law. Thus it is a fact about the world's network of causal relations.

Two further remarks about Lange's view. First, it is inspired by (although I think does not strictly require) a metaphysics in which there are laws at different modal strata: say, force laws at the bottom stratum and then constraints on force laws, such as Newton's second law, at a higher stratum. The laws at each stratum impose non-causal constraints on the stratum below, while the laws at the bottom stratum are causal laws that determine the course of events in the natural world. Lange would say that the higher-level laws are not acting causally; I say that their action on the bottom-level laws is not causal, but their action on events most certainly—albeit indirectly—is.

Second, Lange treats the Königsberg bridges in a similar way to the double pendulum case (if only in passing). In the Königsberg case, however, the higher and therefore putatively non-causal grade of necessity is not a kind of nomological necessity; it is mathematical necessity. This picture is, I think, incompatible with representationalism, on which mathematics has no power to constrain what laws there can be. (The representationalist holds that our representations of the laws must conform to mathematical principles because the principles are built into our system of representation, not because they are built into the world.) I have assumed rather than argued for representationalism, so this cannot be regarded as a refutation of Lange's treatment of the bridges, but it does put his strategy outside the scope of this paper.

* * *

Is all scientific explanation causal? I have not argued for such a sweeping conclusion; what I have done is to remove an obstacle to maintaining such a view, and to argue more generally against any attempt to draw a line distinguishing "non-causal" from causal descriptions of the causal web.

Let me conclude by noting that there is an entirely different way that non-causal explanation might find its way into science: some scientific explanations might be constructed from non-causal raw material, say, from a kind of nondirectional nomological dependence rather than causal influence. Such explanations would describe difference-making aspects of the web of acausal nomological dependence; they would be non-causal from the bottom up. But whether there are any such things is a topic for another time.

References

Bueno, O. and M. Colyvan. (2011). An inferential conception of the application of mathematics. *Noûs* 45:345–374.

Lange, M. (2013). What makes a scientific explanation distinctively mathematical? *British Journal for the Philosophy of Science* 64:485–511.

Lewis, D. (1973). Causation. *Journal of Philosophy* 70:556–567.

Pincock, C. (2007). A role for mathematics in the physical sciences. *Noûs* 41:253–275.

———. (2015). Abstract explanations in science. *British Journal for the Philosophy of Science* 66:857–882.

Railton, P. (1978). A deductive-nomological model of probabilistic explanation. *Philosophy of Science* 45:206–226.

Sober, E. (1983). Equilibrium explanation. *Philosophical Studies* 43:201–210.

Strevens, M. (2004). The causal and unification approaches to explanation unified—causally. *Noûs* 38:154–176.

———. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.

———. (2013). No understanding without explanation. *Studies in History and Philosophy of Science* 44:510–515.