# A Ricci Curvature Tensor by any Other Name

Mark Colyvan*

**Abstract**

There is something right about the view of mathematics as "the language of science". Thinking of mathematics as a language is useful in appreciating the significance of, and the difficulties encountered arriving at, a good notational system. Good notation is far from trivial. The development of differential geometry, for example, with its Ricci curvature tensors and the like, is intimately connected with the notation employed. But thinking of mathematics as *merely* language is to ignore the other roles mathematics can play in science. I will consider the role good notation can play in prompting new ideas and new developments in mathematics and science. I will look at the recent work on mathematical explanation and argue that there are genuinely mathematical explanations of empirical facts and, moreover, the transparency of some of these explanations is dependent upon good mathematical notation.

## 1 The Language of Science

One often hears the claim that mathematics is "the language of science". This, I take it, is meant as a compliment to mathematics. But mathematics is not the language of science in the way the French is the language of love. The latter is surely conventional, perhaps driven by aesthetic preferences for "amour" over "love" and the like. In any case, mathematics, if it is the language of science, is not like this. It's not as though science could have been written some other way but that it looks or sounds sexier when expressed mathematically. The point of the slogan is to emphasise that a great deal of science—especially physics, but many other branches as well—is typically highly mathematical and could not even be formulated without mathematics.

*Sydney Centre for the Foundations of Science, University of Sydney, A14 Main Quadrangle, Sydney, NSW, 2006, Australia. *Email:* mark.colyvan@sydney.edu.au.

It is worth taking seriously the idea that mathematics is the language of science, for there is undoubtedly something right about it. But in doing so we need to guard against a temptation to view mathematics as *merely* the language of science—a medium for expressing scientific ideas but not contributing to the scientific enterprise in any other way. We need to remember that mathematics plays a variety of roles in many diverse branches of science. In this paper I will consider some of these roles and show how the language of mathematics (i.e., mathematical notation) contributes to these roles. In particular, I will show how good, clear, and economical mathematical notation does not only assist in the communication of ideas, it also facilitates new ideas and understanding.

Understanding the role of mathematical notation in both mathematical and scientific practice is important for a number of reasons. One reason is that a great deal of the current work on the philosophy of applied mathematics has focussed on whether mathematics is genuinely indispensable to science and the upshot of this for the old mathematical realism–anti-realism debate. According to one line of thought, originating with W.V. Quine [29], the indispensable role mathematics plays in our best scientific theories gives us good reason to believe that there are mathematical objects [10]. The debate over this issue has turned to the details of mathematics in applications and, in particular, to the ways in which mathematics might be thought to be indispensable to science [17]. Prominent in these discussions is the intriguing idea that mathematics can explain empirical facts.[1] And this, in turn, has prompted a return to Eugene Wigner's [33] problem of explaining the alleged unreasonable effectiveness of mathematics [22].[2] All in all, the focus in the philosophy of applied mathematics has been almost exclusively on the relationship between mathematics and science. In this paper I want to take seriously the language part of the "language of science" slogan. In particular, I will argue that a good philosophical understanding of mathematical notation leads to some surprises. Good mathematical notation can facilitate the applications of mathematics both in mathematics and elsewhere in science. Moreover, I will (tentatively) suggest that mathematical notation can engender mathematical explanations. Mathematics may well be the language of science, but this language is an active participant in our scientific enterprise. And it is no dead language—it is continually evolving.

It is also interesting to note that mathematicians and historians of math-

[1] See, for example, [2, 3, 4, 12, 14, 25, 30], for the debate over the role of mathematics in empirical explanations.

[2] See, for example, [1, 8, 11, 13, 21, 31, 34] for some of the recent work on this topic.

ematics have long recognised the importance of good notation in mathematics.[3] But since the fall from grace of formalism as a philosophy of mathematics, mathematical notation has received little attention in philosophical circles.[4] Formalists saw the subject of mathematics as nothing but the notation and its manipulation. While Gödel's incompleteness theorems were supposed to have put paid to that view—at least to Hilbert's well-developed version of formalism—it is worth reflecting upon what was right about formalism and its attention to mathematical notation. Notation *is* very important to mathematics and a great deal of mathematical activity is concerned with the manipulation of symbols. Formalist went too far in suggesting that mathematics was nothing more than notation and its manipulation, but let's not forget the formalists' insight.

## 2 The Value of Good Notation

### 2.1 The Natural Numbers

To get a feel for the value of a good notation system, let's start with a couple of examples. First, consider the Arabic notation for the natural numbers: $1, 2, 3, 4, 5\ldots$. The familiarity of this notation system makes it easy to overlook just how powerful and extraordinary it is. James Robert Brown [6, p. 85] notes that one reason for the power of this notation lies in the fact that the most important property of the natural numbers—their recursiveness—is reflected in the notation.[5] He suggest that "[t]his is the mathematical equivalent of poetry's onomatopoeia" [6, p. 86]. There is a sense in which the structure of the natural number notation behaves like the structure of the natural numbers themselves.

Brown goes on to suggest that the recursive nature of the natural numbers would need to be known before recursive notation could be introduced. But this doesn't seem right. It is surely at least possible to conceive of a society without the concepts of infinity or of recursion. Yet, they still might entertain the idea of a large but finite number system—enough numbers to make the Arabic notation worthwhile. One could even imagine a member of

[3] For example, see, [9], for the history of mathematical notation.

[4] One notable exception is James Robert Brown [6] who devotes a whole chapter of his introductory philosophy of mathematics text to the topic. Other than this, there are only a handful of other papers on the topic [32, 15].

[5] The notation also makes it easy to assess many inequalities. For example, it is obvious that 373559 is greater than 4749; all one needs to do is note that the first string is longer than the second. Notice that this is not the case in Roman notation where, for instance, x > viii.

such a society realising that their notation allows for the representation of numbers larger than any they currently use. This, in turn, might lead to the discovery of the natural numbers in all their recursive glory. In short, Brown is too quick to dismiss the idea that the notation can help reveal hitherto unknown mathematical facts. At least, there is nothing in Brown's discussion to rule this out (and, based on things he's said elsewhere, I suspect he would be open to the possibility of such notation-driven discovery).

Another important feature of good notation is that it can facilitate computation. The Arabic notation delivers all the standard algorithms for addition, subtraction, multiplication, and division. It does this, of course, by appreciating how the recursive structure of the natural numbers is represented in the Arabic notation. This is far from trivial. Try multiplying two large numbers directly in Roman numerals and you will soon see the superiority of Arabic notation.[6] It is not always made clear that such algorithms depend on the notation. This is something that the formalists were well aware of: you can reduce a great deal of mathematical activity to manipulations of symbols. Be that as it may, there does seem to be calculational power in at least some notational systems.

## 2.2 Elementary Differential Calculus

Now consider the notation of elementary calculus. It is well known that Leibniz and Newton had different notation for the derivatives of functions. Leibniz used $dy/dx$ for the first derivative and $d^2y/dx^2$ for the second derivative, whereas Newton used the dot (or prime) notation: $\dot{f}$ and $\ddot{f}$ (or $f'$ and $f''$) respectively. Newton's notation is more economical but it does not generalise so well to higher dimensions, where one needs to be explicit about which independent variable we are differentiating with respect to. Indeed, Newton was mostly interested in differentiating with respect to time (and the dot notation is still commonly used for time derivatives). Leibniz's notation generalises very well to higher-order partial derivatives, because even in the case of one independent variable, the notation is explicit about differentiating with respect to the variable in question.

I am not claiming any victory for Leibniz over Newton here; it's just that this looks like a case where good notation can facilitate new mathematics (in this case multi-variate calculus) by making the transition to the general

[6] Brown [6, pp. 86–93] gives other instructive examples of computational power arising from notation. For example, he considers knot theory, where the notation allows for the determination of whether two tangles are the same. See especially the result invoking Conway notation and continued fractions.

case seamless. This case is different from the last in that here we see that notation might be good for one purpose but not good for another: Newton wins on the economy stakes but Leibniz wins on the generalisability stakes. Also note again how good notation might help suggest new mathematics. In Leibniz notation, it is very natural to consider the question of whether the mixed partial derivatives $\partial^2 u/\partial x\partial y$ is the same as the other mixed partial derivative $\partial^2 u/\partial y\partial x$. Of course these mixed partial derivatives are equal at any point where the function $u$ has continuous mixed partial derivatives. For present purposes, I simply note that questions such as the relationship between the mixed partial derivatives naturally arise from the notation itself. Good notation, it seems, can prompt the user to keep track of distinctions the inventor of the notation may not have even noticed.

Next, consider the standard notation for the Laplace operator in two dimensions:

$$\Delta u(x,y) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{1}$$

This notation is economical, allowing us to write Laplace's equation in two dimensions as: $\Delta u(x,y) = 0$, and if the dimension is understood, we need only write: $\Delta u = 0$. The economy of this notation is even more apparent when we go beyond two dimensions. The generalised $n$-dimensional Laplace operator is defined as

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} + \dots \frac{\partial^2 u}{\partial x_n^2}. \tag{2}$$

It is worth pointing out that the notation itself, insofar as it ignores the dimension, might even be thought to suggest the generalised $n$-dimensional version of the operator. Even if the generalisation were not already in play, the notation would remind one that there is nothing special about the two-dimensional case. We thus see that good notation can be economical as well as suggest (or at least lend itself to) further mathematical developments. Since mathematical developments are very often generalisations to more abstract structures, we can restate this last virtue of good notation as that of *facilitating more abstract, generalised mathematical theories*.[7] This brings us very naturally to our next example, *differential geometry*.

[7] There is also the related issue of which concepts are the most fruitful ones to focus on. Lakatos's most significant contribution to the philosophy of mathematics was to argue (convincingly) that mathematical concepts are very often refined in the light of mathematical results and counterexamples [24]. So here we might reasonably ask why the Laplacean operator is so useful in both pure and applied mathematics and physics. I set

## 2.3 Differential Geometry

Differential geometry is the study of manifolds in a coordinate-free setting. The notation here is crucial and there are several different notational systems in circulation—each with its own virtues and limitations. (Indeed, such is the proliferation of alternative notations in the subject that there's a joke that differential geometry is the study of that which is invariant under change of notation.) One immediately striking feature of the notation in differential geometry is that it is so compact. For example, one of the key items in differential geometry is the Riemann curvature tensor. In 4-dimensional manifolds, as we have in general relativity, the Riemann curvature tensor has 256 components. This reduces to 20 independent components once various symmetries and identities are exploited. But, remarkably, the tensor can be written using simply four indices: $R^j_{ikl}$. Just as in the case of the notation for the Laplacian operator, much is being tracked implicitly and there are, for example, various summation conventions in place throughout differential geometry. But still it is striking how much is captured in the various versions of the notation.

Another feature of the notation worthy of comment is that since so many of the key items in differential geometry are built up out of features of the metric for the manifold in question, it is relatively straightforward to keep track of what is intrinsic and what is extrinsic to the manifold. Here an intrinsic item is one which can be determined from within the manifold in question: distances, angles and so forth, but not normal vectors. This feature of the notation is key in proving one of the first major results of differential geometry: Gauss's Theorema Egregium. Finally, the notation, in a sense, mirrors the key structural features of interest. One can thus get a long way by just manipulating the notation—learning rules about how to manipulate the indices and the like. The formalist idea of what mathematics is about does not seem too far off the mark here. But this is no accident. You can go a long way by symbol manipulation in differential geometry precisely because the notation is so constructed to faithfully track the key features of interest in the manifolds in question.

In order to get a better feel for what I'm suggesting here, let's consider some of the main objects in differential geometry and their notation. We have the Christoffel symbols (of the second kind): $\Gamma^i_{jk}$. These are intrinsic, fully determined by metric coefficients, and measure (some aspects of) the

---

this interesting issue aside and assume that we have already decided which mathematical concepts are useful and we are merely concerned with the question of representing them in our notational system.

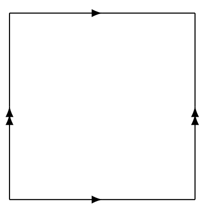
curvature of the surface.[8] I have already mentioned the Riemenn curvature tensor, $R^j_{ikl}$. This is an intrinsic invariant and can be defined in terms of Christoffel symbols. And, of course, there's the Ricci curvature tensor: $\mathrm{Ric}_{\mu\kappa}$. This employs only two indices and yet it fully describes the way in which volume in the manifold deviates from Euclidian volume. The Ricci curvature tensor is especially important in general relativity, where it is a key component of the Einstein field equation [16].

As I mentioned before, an important early result in differential geometry is Gauss' Theorema Egregium. This tells us that the Gaussian curvature of a manifold, although defined extrinsically (in terms of normal vectors and the like) can be expressed in terms of a combination of a Riemann curvature tensor and a metric tensor, both of which are intrinsic. The Gaussian curvature is thus also intrinsic. This result is remarkable for a number of reasons but prime among these is that it allows us to make sense of a curved $n$-dimensional manifold, without requiring a larger space, an $(n + 1)$-dimensional space, through which the $n$-dimensional space curves. The $n$-dimensional space is curved, and that's that [19, 28].

## 2.4 Topology

Next consider an example from topology. Topology is the study of features of spatial structures preserved under continuous transformations, or as the slogan goes: "topology is rubber-sheet geometry". There is a particularly powerful piece of diagramatic notation that's a kind of cardboard cut-out plan—except that the cardboard is flexible and the construction sometimes cannot be physically carried out [27]. The idea is simply to draw a square and identify the edges as required, via fold-and-paste instructions.

Let's start with a familiar example: a torus.

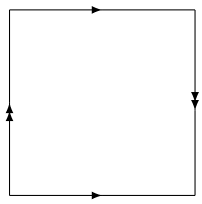


Here the arrows mean that we identify the right edge with the left edge, with the same orientation, and we also identify the top edge with the bottom edge,

[8] Even their names—"Christoffel symbols"—places the notation at the fore, as though there were no mathematical object or concept behind the notation. This is already unusual, and gives one a clue as to how important the notation is here.

again with the same orientation. That is, cut out a square sheet of rubber and glue the right edge to the left to form a cylinder, then glue the two circular ends of the cylinder (what were the top and bottom edges of the original square) together to form a torus.

Now we use this notation to represent the Klein bottle.



This is the same as the torus except that the second identification is performed with the reverse orientation. The latter requires a twist. But this cannot be physically performed because we've run out of dimensions. The twist needs to be in a fourth dimension. The Klein bottle is a three dimensional object but it is not constructible in three dimensions and has no inside or outside. Möbius strip.

The Klein bottle, as presented above, is a lovely demonstration of the power of good notation. The notation makes it plain that there are such things as three-dimensional objects that require more than three spatial dimensions for their construction. The notation also shows that there are three dimensional objects with no inside or outside. Before seeing and understanding the relevant notation, one might be sorely tempted to dismiss the idea of such objects as nonsense. But even if we had not stumbled onto Klein bottles already, the notation just developed would lead us to them. We can even deduce the properties of objects like Klein bottles by studying the details of the notation. (For example, from the notation alone it can be seen that the Klein bottle has no inside or outside.) Arguably, the notation also helps us to see why we cannot build such objects: because we are stuck in three spatial dimensions yet we need to engineer a twist in a fourth dimension.

Another interesting issue raised by this particular piece of notation involves the difference between algebraic methods and geometric methods. For a long time it has been the orthodoxy in mathematics to see geometric intuitions as unreliable. Instead, we should trust only purely algebraic methods. The latter are considered to be the epitome of rigour. Pictures and geometric intuitions are at best pedagogical devices, or so this line of thought goes. This view has been challenged recently [6, 18] and some have even suggested that the distinction between algebraic methods and geometric ones is not as sharp as it might first seem. This example lends support to the latter thesis. This algebraic topology notation is something of a halfway house between pure algebra and pure geometry. It is both notation and a kind of blueprint for construction of the objects in question. The first seems to belong to algebra, while the second is clearly geometric. But whichever way you look at it, we have a powerful piece of notation here that does some genuine mathematical work for us.

## 2.5 The Point at Infinity

One final example of how clever notation can reveal something mathematically interesting. Here we show that one can add infinity to the complex plane, to deliver the extended plane, without leading to trouble. This is done by a construction called *the stereographic projection*. But this construction can also be considered to be an alternative notation for the complex plane. Here's how it goes. Consider the complex plane laid out horizontally and with each point represented with the usual Cartesian 2-coordinates $(x, y)$. Now add a third, vertical dimension, the $z$ dimension, and with this three space represented with the usual Cartesian 3-coordinates $(x, y, z)$. Consider a unit sphere with its centre at the origin $(0, 0, 0)$: $x^2 + y^2 + z^2 = 1$. Now we take the line generated by joining a point in the the $x$–$y$ plane with the north pole of the hemisphere $(0, 0, 1)$. In particular, we are interested in where this line intersect our unit sphere. For example, the line joining the point $(1, 0)$ in the $x$–$y$-plane and the north pole of our sphere intersects the sphere at $(1, 0, 0)$. The line joining the point $(0, 0)$ to the north pole, intersects the sphere at the south pole $(0, 0, -1)$. If we now identify each point of the Cartesian plane with the corresponding point of intersection of the sphere, we see that all the points inside the unit circle $x^2 + y^2 = 1$ are mapped to unique points on the part of the sphere below the $x$–$y$-plane, all the points on the unit circle get mapped to their 3-dimensional counterpart (i.e., $(x, y)$ gets mapped to $(x, y, 0)$), and all the points outside the unit circle get mapped to unique points on the portion of the sphere above the $x$–$y$-plane. Moreover, the further from the origin the point is, the closer its corresponding point of intersection is to the north pole of the sphere. Via this conformal mapping—we have created an alternate representation for each point in the real plane.

What is useful about this apparently-cumbersome representation, is that every point on the plane is represented by an ordered triple (points on the surface of the sphere) and every point of the surface of the sphere, except one—the north pole—has a corresponding point on the $x$–$y$-plane. But now we just add the north pole and stipulate that it is the point at infinity. The motivation for this should be plain to see. It should also be clear that the north pole represents infinity in all directions on the $x$–$y$-plane. Using this construction (or alternative notation) we have shown that it is sensible to talk of the extended real plane (i.e., the complex plane with a single point at infinity added). That this can be done is not at all apparent in the standard notation. Again, we have found an interesting extension of our mathematics, and this extension was facilitated by the alternative notation employed in the stereographic projection.

# 3 Notation and Explanation

So far we have seen that good notation can enjoy the virtues of economy, calculational power, and facilitating advances in mathematics. But let's push things a little further. Let's pursue the suggestion, hinted at in the topology example, that good notation can make mathematical explanations more perspicuous.

Mathematicians talk of explanatory and non-explanatory proofs in mathematics [20]. The explanatory proofs tell you why the theorem in question holds while the non-explanatory proof merely tell you that it holds. The explanation is something delivered by the proof and, of course, there can be explanatory and non-explanatory proofs of the same theorem. Exactly what this explanation amounts to is hard to say. Indeed, it is hard to say what an explanation is, in general. In empirical science and in everyday life, very often a request for an explanation is a request for details of the causal history. Why did the window break? Because it was hit by a rock. Why did the cathode emit X-rays? Because it was bombarded by high energy electrons. In mathematics, however, whatever an explanation amounts to, it can't be anything to do with causal history. Explanations of why the Poincaré conjecture is true, cannot have anything to do with causal histories. Apart from anything else, if mathematical objects such as functions and derivatives exist, they do not have causal powers and they do not enter into causal networks. Causal accounts of explanation, whatever virtues they have elsewhere in science, simply miss the mark in mathematics. So what is a mathematical explanation?

One promising line on mathematical explanation is that it is unification.[9] According to this account of explanation, we explain some phenomenon

[9]See [23, 26] for more on this account of explanation.

when we incorporate it into a wider class of phenomenon. To give a classic example: Newton's theory of gravitation unified, planetary motion, terrestrial projectiles and the tides, and in doing this it explained these phenomenon. It is important to note that on this view, Newtonian gravitation explains the tides, even though the reason for the tides is explained in terms of a somewhat mysterious force acting at a distance: gravitational attraction of massive bodies. Explanations do not need to bottom out in self-evident truths, and they don't have to eliminate all mystery. In mathematics, we can see how this might work. Seeing connections between apparently unrelated parts of mathematics is indeed enlightening, and arguably explanatory. For example, Euler's famous formula, $e^{i\theta} = \cos\theta + i\sin\theta$ reveals deep connections between trigonometry and complex analysis. Representing systems of simultaneous equations via linear algebra enables results about when solutions can be found in terms of the properties of the corresponding matrices (determinants, eigenvalues and so forth).

That's explanation within mathematics. Some have gone further to suggest that mathematics can explain empirical facts.[10] For example, it has been argued that the reason one cannot square the circle using only compass and straightedge is that $\pi$ is transcendental. Here we have mathematics explaining an important fact about history: every attempt to devise a method for squaring the circle failed and they all failed for the same reason. In another example, Alan Baker [2] suggests that the reason North American cicadas have life cycles of 13 and 17 years involves facts about prime numbers. It turns out that having a life cycle of 17 years (or some other prime number in the vicinity) is a very effective way of avoiding predators, since only those predators with the same prime life cycles or annual predators will coincide with the cicadas at their most vulnerable stage. Other life cycles, 14 years, say, may coincide with predators with 14 year, 7 year, 2, year and 1 year cycles. Prime life cycles for such semelparous reproductive species increases survival rates. And a major part of the explanation of why this is so comes from number theory. This explanation is not purely mathematical—it also involves facts about predators and the the cicada life cycle—but it does have a mathematical component to the explanation. In other words, an explanation of the 13 and 17 year life cycles of North American cicadas would not be complete without an appeal to the fact that 13 and 17 are prime numbers.

Now we return to the main topic of this paper—mathematical notation.

[10] See, for example, [2, 4, 7, 10, 12, 14, 25, 30] for discussion of whether mathematics can explain empirical facts.

I will demonstrate how good mathematical notation can help engender explanations, both intra-mathematical explanations and extra mathematical explanations of empirical facts.

### 3.1 Analytic Geometry

Chief among René Descartes' many contributions to philosophy, mathematics, and science, is his combining geometry and algebra in *analytic geometry*. The idea is so familiar that it is hard to fully appreciate just how brilliant this innovation was. The idea is simply that we invoke a coordinate system for the plane, then two-dimensional geometric figures can be represented algebraically in terms of these coordinates. In the other direction, we can represent algebraic equations and inequalities geometrically. With this algebraic notation for geometric figures, we are able to use geometry to help visualise otherwise abstract algebraic problems, and to use the rigorous methods of algebra to solve geometric problems.

For example, from the relevant geometry we can see why the polynomial $x^2 - 1$ has 2 real roots: because the corresponding parabola $y = x^2 - 1$ has its vertex below the $x$-axis. We can also see why the polynomial $x^2 + 1$ has no real roots: because the corresponding parabola $y = x^2 + 1$ sits entirely above the $x$ axis.[11] It is, of course, possible to understand why $x^2 + 1$ has no real roots by considering the algebra alone, but with the connection to geometry in place, you can *see* why. This is just the tip of the iceberg. By invoking the power of analytic geometry we can visualise differentiation (as the function representing the slope of a tangent to a curve at a given point), integration as the area under a curve, and so on. These and many other applications are familiar. It is worth reflecting for a moment, however, on the important role the algebraic notation for geometric figure plays in all this. Indeed, all there is to analytic geometry is the merging of algebraic notation and geometry, and somehow both geometry and algebra are enriched by the merger.

One striking example of the power of algebraic methods in geometry is the Weierstrass function. It is tempting to believe that continuous functions can fail to be differentiable at only a countable number of points. That is, if a function is continuous on an interval, it may have at most a countable number of points at which it is not differentiable. Weierstrass demonstrated that this natural and intuitive suggestion is mistaken. He showed that there are continuous functions which are nowhere differentiable—they are nothing

[11] Although, of course, the fundamental theorem of algebra, guarantees that the polynomial in question has 2 complex roots: $\pm i$, where $i = \sqrt{-1}$.

but cusps. The function Weierstrass originally produced was

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x),$$

where $0 < a < 1$, $b$ is a positive odd integer and $ab > 1 + 3\pi/2$. Once you see the trick involved, you see that there are many other such functions. It is very hard to imagine discovering such functions without the power of algebraic notation at one's disposal. After all, pure geometric intuitions, if anything, suggest that are no such functions. But more importantly, the algebraic notation allows the explanation of why some continuous functions may fail to be differentiable on intervals.

My suggestion in this and in the following example is that the notation helps us understand what's going on, and in this sense, helps engender explanations. By this I don't mean to suggest that the notation is the reason for the phenomenon in question, just that the notation helps make the explanation accessible to us.[12] Next we consider an example where algebraic notation is used in a more essential way, and the explanation in question again relies on the algebraic representation of geometry.

### 3.2 Squaring the Circle

Squaring the circle is the well known problem of constructing a square of the same area as a given circle, using just a straightedge and compass.[13] That is, using only these two instruments we must construct a square of side $r\sqrt{\pi}$, where $r$ is the radius of the circle in question and $\pi$ is the ratio of the circumference to the diameter of a circle. It is well known that after centuries of attempts, the impossibility of this construction was finally proven in 1882 by Ferdinand von Lindemann. The proof, however, comes from abstract algebra, not geometry. A quick sketch of the connection here is worthwhile, because the algebraic notation is crucial to the proof in question [5].

---

[12]We can distinguish two senses of explanation: one an objective sense and the other a psychological sense. For example, a scientific explanation might not be understandable to someone not versed in the relevant science. That does not stop it from being an explanation in the objective sense. But its failure to enlighten the person in question prevents it from being a psychological explanation for that particular person. It is the latter we're concerned with here.

[13]This is one of three famous ancient geometric construction problems. The other two problems are trisecting an angle, and doubling the cube. The former is the problem of trisecting and arbitrary angle, the latter is that of creating a cube with twice the volume of a given cube.

First, we catalogue the legitimate, basic ruler and straightedge constructions (drawing a line through two existing points, constructing a circle with centre at one existing point and running through another existing point, and so on). We then provide notation for the basic geometric objects (lines, points and arcs of circle) and note that we can represent these objects in the Cartesian plane, in the usual way. We then show that the permissible geometric constructions give rise to a small set of algebraic operations on line lengths: addition subtraction, division, multiplication and taking the square root. The idea here is that if two line segments of lengths, $a$ and $b$ are given, we can construct line segments of length $ab$, $a+b$, $a-b$, $a/b$ and $\sqrt{a}$. What is crucial is that these are the *only* algebraic operations the geometric constructions licence. Note that what we have done is again forge a link between the geometric constructions and algebra. This allows us to apply algebraic methods to the problem. We have thus transformed the problem from one of geometry—that of constructing a given length, $\sqrt{\pi}$—to one of abstract algebra—determining whether $\sqrt{\pi}$ can be obtained from by successive applications of the algebraic operations just listed. In particular, the notation just described turns this into a problem of determining whether $\sqrt{\pi}$ can be the root of a polynomial with powers 0, 1, or an even integer and with rational coefficients. This is where the Lindemann result comes in. Lindemann proved that $\pi$ (and therefore $\sqrt{\pi}$) is transcendental. That is, $\pi$ (and $\sqrt{\pi}$) is not the root of any polynomial with rational coefficients. The construction in question is thus impossible.[14]

Although the fact that $\pi$ is transcendental is the key to the impossibility result, it is important to see how the problem needed to be set up as an algebraic problem. This involved the introduction of algebraic notation for the geometric objects and operations, and noting that the geometric operations give rise to some familiar algebraic systems. Again we see good mathematical notation playing a key role in delivering a mathematical explanation. But notice that the explanation goes beyond pure mathematics. We have also explained why all attempts to square the circle failed and why anyone who claims to have squared the circle is not taken seriously in mathematical circles.

[14]The other classic geometric constructions mentioned earlier—trisecting an angle and doubling the cube—also turn out to be impossible. They are impossible for different reasons: respectively, $\cos(\pi/9)$ is not the root of any of the polynomials just described, nor is $\sqrt[3]{2}$. In each case, the same notation outlined here, is used and is important in delivering the impossibility result in question.

## 4 Shakespeare's Mistake

We have seen that good notation can enjoy the virtues of economy and calculational power, and can facilitate advances in mathematics. More controversially, good notation may also contribute to mathematical explanations and perhaps even explanations beyond mathematics. This already takes us well beyond the standard view (insofar as there is such a thing) of mathematical notation. The standard view I have in mind here, suggests that its the mathematical objects that matter, not the notation we use for them: as Shakespeare put it in *Romeo and Juliet*:

> What's in a name? that which we call a rose
> By any other name would smell as sweet.

Following Shakespeare's lead, it might be tempting to suggest that a mathematical object by any other notation would be just as useful. But we have seen that this is not so. Sometimes the notation encode properties of the objects in question and in such cases alternative notations would be less revealing.[15] Perhaps because of this encoding, good notation can keep track of distinctions we may not have initially noticed, force us to investigate such distinctions, and to see possibilities for future research not previously anticipated. In short, properties of the notation are important in mathematics. So Shakespeare was wrong, at least about mathematical notation, but I'd suggest he was wrong about natural language as well. A rose by any other name may well smell as sweet, but sometimes the name chosen is important; smell isn't all we're interested in. In any case, it is far from clear that mathematics would be served equally well by alternative notations. Getting the notation right features prominently in mathematical practice. And there is a good reason for this: good notation does serious work in mathematics. A Ricci curvature tensor, for instance, by any other notation may not live up to its full potential. And it is for this reason that mathematicians spend so much time developing and arguing about notation.

Mathematics may well be the language of science, but we must not underestimate the power of this language, for both expressive purposes and in facilitating new ideas and explanations. Good mathematical notation, like any powerful language helps us to think and to gain understanding.

[15]It is worth noting that this is true of natural language as well: onomatopoeias reveal something about the sounds they name, and other words such as "computer' and "amplifier" tell you something about the objects they name.

Shakespeare, of all people, should have known better.[16]

## References

[1] Azzouni, J. 2000. 'Applying Mathematics: An Attempt to Design a Philosophical Problem', *The Monist*, 83: 209–227.

[2] Baker, A. 2005. 'Are There Genuine Mathematical Explanations of Physical Phenomena?', *Mind*, 114: 223–238.

[3] Baker, A. 2009. 'Mathematical Explanation in Science', *British Journal for the Philosophy of Science*, 60: 611–633.

[4] Batterman, R.W. 2010. 'On the Explanatory Role of Mathematics in Empirical Science', *The British Journal for the Philosophy of Science*, 61(1): 1–25.

[5] Bold, B. 1982. 'The Problem of Squaring the Circle', Chapter 6 in *Famous Problems of Geometry and How to Solve Them*. New York: Dover, pp. 39–48.

[6] Brown, J.R. 2008. *The Philosophy of Mathematics: A Contemporary Introduction to the World of Proofs and Pictures*, 2nd edition. London: Routledge.

[7] Brown, J.R. 2011. *Platonism, Naturalism, and Mathematical Knowledge*, London: Routledge.

[8] Bueno, O. and Colyvan, M. 2011. 'An Inferential Conception of the Application of Mathematics', *Noûs*, 45: forthcoming.

[9] Cajori, F. 1993. *A History of Mathematical Notation*, New York: Dover Reprints. (Original published in two volumes by Open Court in London in 1929.)

[10] Colyvan, M. 2001. *The Indispensability of Mathematics*, New York: Oxford University Press.

[11] Colyvan, M. 2001. 'The Miracle of Applied Mathematics', *Synthese*, 127(3): 265–278.

[12] Colyvan, M. 2002. 'Mathematics and Aesthetic Considerations in Science', *Mind*, 111: 69–74.

[16]Thanks to Jim Brown, Jeremy Butterfield, Matthew Spillane, and Zach Weber for very helpful discussions on the topic of this paper. I am also indebted to audiences at the 2009 Australasian Association of Philosophy Conference at the University of Melbourne, and at the 2011 European Philosophy of Science Association meeting at the University of Athens, where earlier versions of this paper were presented.

[13] Colyvan, M. 2009. 'Applying Inconsistent Mathematics', in O. Bueno and Ø. Linnebo (eds.), *New Waves in Philosophy of Mathematics*, Basingstoke UK: Palgrave MacMillan, pp. 160–172. Reprinted in M. Pitici (ed.), *The Best Writing on Mathematics 2010*, Princeton: Princeton University Press, 2011, pp. 346–357.

[14] Colyvan, M. 2012. *An Introduction to the Philosophy of Mathematics*, Cambridge: Cambridge University Press.

[15] De Cruz, H., and De Smedt, J. Forthcoming. 'Mathematical Symbols as Epistemic Actions', *Synthese*.

[16] d'Inverno, R. 1992. *Introducing Einstein's Relativity*. Oxford: Clarendon.

[17] Field, H.H. *Science Without Numbers: A Defence of Nominalism*, Oxford: Blackwell Publishers.

[18] Giaquinto, M. 2007. *Visual Thinking in Mathematics*, Oxford: Oxford University Press.

[19] Gowers, T. 2008. *The Princeton Companion to Mathematics*. Princeton: Princeton University Press.

[20] Gowers, T. and Neilson, M. 2009. 'Massively Collaborative Mathematics', *Nature*, 461 (15 October 2009), 879–881.

[21] Grattan-Guinness, I. 2008. 'Solving Wigner's Mystery: The Reasonable (Though Perhaps Limited) Effectiveness of Mathematics in the Natural Sciences', *The Mathematical Intelligencer*, 30: 7–17.

[22] Hamming, R.W. 1980. 'The Unreasonable Effectiveness of Mathematics', *The American Mathematical Monthly* 87(2): 81–90.

[23] Kitcher, P. 1981. 'Explanatory Unification, *Philosophy of Science*, 48, 507–531.

[24] Lakatos, I. 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge: Cambridge University Press.

[25] Lyon, A. and Colyvan, M. 2008. 'The Explanatory Power of Phase Spaces', *Philosophia Mathematica (3)*, 16: 227–243.

[26] Mancosu, P. and Hafner, J. 2008. 'Unification and Explanation: A Case Study From Real Algebraic Geometry', in P. Mancosu, (ed.), *The Philosophy of Mathematical Practice*. Oxford: Oxford University Press, pp. 151–178.

[27] Massey, W.S. 1989. *Algebraic Topology: An Introduction*, New York: Springer-Verlag

[28] Millman, R.S. and Parker, G.D. 1977. *Elements of Differential Geometry*, Englewood Cliff, NJ: Prentice-Hall.

[29] Quine, W.V. 1981, 'Success and Limits of Mathematization', in *Theories and Things*, Cambridge, MA: Harvard University Press, pp. 148–155.

[30] Steiner, M. 1978. 'Mathematics, Explanation, and Scientific Knowledge', *Noûs*, 12: 17–28.

[31] Steiner, M. 1998 *The Applicability of Mathematics as a Philosophical Problem*, Cambridge, MA: Harvard University Press.

[32] Weber, Z. Forthcoming. 'Figures, Formulae, and Functors', unpublished manuscript.

[33] Wigner, E.P. 1960 'The Unreasonable Effectiveness of Mathematics in the Natural Sciences,' *Communications on Pure and Applied Mathematics*, 13: 1–14.

[34] Wilson, M. 2000. 'The Unreasonable Uncooperativeness of Mathematics in the Physical Sciences', *The Monist*, 83: 296–314.