How not to factor a miracle

Derek K. Wise

Mathematics is a bit like Zen, in that its greatest masters are likely to deny there being any succinct expression of what it is. It may seem ironic that the one subject which demands absolute precision in definitions would itself defy definition, but the fact is, we are still figuring out what mathematics is. And the main way to figure out what mathematics really is is to do mathematics.

Mastering any subject takes years of dedication, but with mathematics it takes years before one even begins to see what it is that one has spent so long mastering. I say “begins to see” because so far, I have no reason to suspect this process terminates. Neither do wiser and more experienced mathematicians I have talked to.

In this spirit, for example, The Princeton Companion to Mathematics [PCM], wisely renounces any tidy answer to the question “What is mathematics?” Rather, the book’s approach is to give the reader a feel for the whole of mathematics by providing 1000 pages of expositions of topics within mathematics, all written by top experts in their own subfields.

Unfortunately, while mathematicians are often reluctant to define mathematics, others are not. In 1960, despite having made his own mathematically significant contributions, physicist Eugene Wigner defined mathematics as “the science of skillful operations with concepts and rules invented just for this purpose” [W]. This rather negative characterization of mathematics may have been partly tongue-in-cheek, but he took it seriously enough to build upon it an argument that mathematics is “unreasonably effective” in the natural sciences—an argument which has unfortunately been unreasonably influential among scientists.

What weight we attach to Wigner’s claim, and the view of mathematics it promotes, has not just metaphysical implications, but also practical implications for the progress of mathematics and physics. In fact, my purpose here is not so much to refute 55-year-old views of a great physicist (whose contributions to physics and mathematics I admire) as it is to challenge a worldview that these ideas helped promote. If the effectiveness of mathematics in physics is a ‘miracle,’ then this miracle may well run out. In this case, we are justified in keeping the two subjects ‘separate’ and hoping our luck continues. If, on the other hand, they are deeply and rationally related, then this surely has consequences for how we should do research at the interface.

Is physics unreasonably well described in the language invented for that purpose?

The essence of Wigner’s claim—based an unsupported characterization of mathematics as an elaborate mental game that has nothing a priori to do with physics—is that mathematics is somehow nonetheless amazingly effective in describing physics. Were physics and mathematics disjoint endeavors, one empirical, one cognitive and largely arbitrary, then it would indeed be astounding, even ‘miraculous,’ to find them interacting as fruitfully and precisely as they do. But this viewpoint disregards both history and the mathematical biases set up by our place in the natural world.

First, historically, the distinction between mathematics and physics is relatively recent, with the movement toward ‘pure’ mathematics progressing gradually since the 18th century. The historical counter-argument to Wigner’s thesis has been taken up by mathematics historians [GG] in detail, and that is not my purpose here. However, it is worth pointing out that, for example, calculus and a large part of differential equations—still among the physicist’s most important tools—were designed precisely for physical applications. It is at least anachronistic, but bordering on absurd, to claim these are unduly effective in the very subject they owe their existence to.

In fact, the joint development of physics and mathematics predates the recognition of physics as a separate science. I must agree with Einstein that “mathematics generally, and particularly geometry, owes its existence to the need which was felt of learning something about the relations of real things to one another,” and that “we may in fact regard [geometry] as the most ancient branch of physics” [E2].

Second, at a deep level, mathematics—or rather our limited view of it—is founded on our perception of the natural world. To get an idea of how much our view of mathematics depends on physics, we need not imagine a world with drastically different physics: we can simply consider our own world at different scales. If we were very large mathematicians, or equivalently
if Newton’s constant $G$, which regulates the strength of gravity, were large, so that Einstein’s famous curving of space and time were evident at the scale of everyday life, then we would at least have a very different history of geometry.

Mathematics would be much more radically different if instead we were very small mathematicians, or equivalently if Planck’s constant $\hbar$, which regulates the level of ‘quantum fuzziness,’ were large. This would also have affected our ideas of geometry, since there would be no stationary objects with well-defined positions to serve as reference points for measurements. But we would likely have much less interest in counting things, since distinct collections of objects become much less sensible at quantum scales. We would have little motivation to invent number theory, set theory, or combinatorics. These subjects, which are fundamental parts of mathematics as we know it, owe their existence to our place within the physical world.

Moreover, as ‘quantum mathematicians’ we would presumably not prove theorems according to the familiar rules of logic, which are deeply tied to set theory, but according to rules of quantum logic which reflect the fuzzy, indistinct nature of propositions about fundamentally quantum mechanical systems.

More radical still, we can consider the possibility that $G$ and $\hbar$ are both large, corresponding to a realm of experience where effects of ‘quantum gravity’ should be evident. Here the foundations of mathematics would surely differ even more radically from our usual ones, in ways that we can only speculate depending on which of many approaches to quantum gravity, if any, one believes.

While it is interesting to consider how our foundations of mathematics are determined by physics, it is in fact recognized from a perspective entirely within pure mathematics that mathematics itself could be built on alternative foundations. This goes beyond the philosophy of mathematics to what we might call the mathematics of mathematics, in that the types of foundations of mathematics—different ‘mathematical worlds’ with their own notions of logic—can be classified mathematically. For example, the potential ‘mathematical foundations’ with many of the same formal properties as the usual set-based foundations are classified by topos theory. This motivates attempts to use topos theory in the foundations of physics.

If our experience of the physical world were different, then mathematics would look quite different to us as well, but it would still be mathematics. Presumably it would be even better than standard mathematics at describing phenomena within the appropriate ‘scale’.

### Appropriate abstraction

Outside of mathematical or philosophical discussions, when people call a thing “abstract,” they may only mean difficult to understand, or perhaps unnecessarily cerebral. But true abstraction is just the opposite: a path to clarity. It means stripping away all inessential details, arriving at an idealization whose life purpose is to exemplify the properties or concepts under consideration, and nothing else. Russell used the word in this sense when he said the language of physics is necessarily abstract, in order to “say as little as the physicist means to say.”

The raises a key question: How does one arrive at a good mathematical abstraction? It is perhaps not surprising that we can find useful mathematical abstractions of the simplest objects in our experience; ultimately, our ability to find good mathematical abstractions of concepts in physics, especially as compared to other sciences, must rest on physics being much simpler than other sciences.

Much more surprising—and I gather this is the part Wigner also finds incredible—is the ability to ‘stay’ at an abstract mathematical level, not referring back to direct physical experience, and still somehow manage to arrive at new mathematical constructs which later turn out to have applications to physics.

So, let us ask: Where do new mathematical concepts come from? Sadly, Wigner is again rather dismissive on this point, claiming that most advanced mathematical concepts are defined just so that the mathematician can “demonstrate his ingenuity and sense of formal beauty.” If we are not to be limited by the way we have been trained to think, we need to develop new ways of thinking and being creative. Of course, the reason we do not train graduate students in making good definitions is that the’s no single process that leads to the right definition, nor for that matter a single criterion that makes a definition ‘right.’ Some definitions, once one recognizes a need for them, are rather obvious. In other cases, it
may take years for a community of mathematicians to agree on the best definition for a concept.

To see what makes mathematicians think a definition is good, it is best to study examples, and it is worth considering at least one in some detail. One of my current favorite examples is the notion of quantum groupoid. Groupoids are a modern mathematical way to study the notion of symmetry—arguably the most important concept in mathematical physics. Intuitively, quantum groupoids result from taking the concept of groupoid, based on sets and hence classical logic, and importing it into the world of vector spaces and hence quantum logic.

But realizing this intuitive idea and settling on a definition of quantum groupoid took time and effort. To be honest, the definition that finally emerged looks quite complicated and perhaps esoteric from the perspective of an outsider showing up on the scene after the dust had settled. It is much more complicated than what I would try, or what researchers did in fact try first.

So, what makes this definition is ‘good’ or even ‘correct’ in some sense? If you simply handed me the definition of quantum groupoid, told me some impressive theorems that could be proved about them, I would think it very unlikely that it would ever find use in physics. But consider this evidence in favor of the quality of the definition:

- The definition naturally relates mathematical structures with many known applications, also in physics, including ‘groupoids’ and ‘quantum groups.’
- More naive attempts at the definition were refined by the need to include natural examples.
- Despite being rather complicated, equivalent definitions were discovered independently by researchers with quite different starting points, from within different subfields of mathematics.

Notice that all of this evidence (and more could be listed) is purely mathematical, and none of it is based on our ability to “demonstrate [our] ingenuity and sense of formal beauty.” Rather, the evidence gives a strong sense that this is a robust mathematical concept. This together with the tight relationship to concepts with known physical applications which one may naturally want to combine makes me think physical applications for quantum groupoids may not be so far off after all.

Ways of ignoring

Of course, the simpler a physical system is, the more likely we can find an appropriate mathematical abstraction of it, and in practice we do not model the entire system we are interested in but some simplification of it. The idea that we can study a complicated system by isolating ‘parts’ of it and focusing on those is known as reductionism.

At a basic level, reductionism means ignoring nearly everything in the universe. But there different ways of ignoring: it is one thing to ignore most of the things in the universe, and quite another to ignore most of the properties of the things in the universe. For clarity, I will call the first of these reductionism and the second co-reductionism, though this is not a standard name. These both have important roles to play in fundamental physics as we know it, so it is worth taking some time to understand them in detail.

Mathematically, the difference between these two ways of ignoring corresponds to the difference between sub-objects and quotient objects. A sub-object is essentially what it sounds like: a smaller piece of some other object. A particular subset of set of all people is the set containing only Eugene Wigner, and this can be represented as an arrow from the subset into the whole set, indicating how the part is included in the whole:

We engage in this kind of reductionism if we want to write a biography of Wigner, making a study of him as an individual. Of course, we care not just about the Wigner here, but also his place in the set of all people—that is where the arrow comes in. However, our main subject is Wigner, and we ignore all other people at least insofar as they do not interact with Wigner.

Quotient objects represent a less familiar way of thinking, but are easy to understand by examples. A particular quotient of the set of all people is the set of genders. This can be represented as arrow going from the whole set to the quotient set:

The arrow simply represents assignment of the appropriate gender to each person. Here we are not forgetting any people, but rather forgetting nearly everything about all of them.
Mathematicians have a nice way to think of this: the arrow is thought of as a process in which we “identify” (i.e. declare to be identical) any two source elements that map to the same target element. In the example, we imagine that Wigner and I, by virtue of being both male, are identified: we become the same person according to a vastly oversimplified model where the only distinguishing feature of a person is gender.

It is important to distinguish between ‘sub-’ and ‘quotient’ constructions. For example, \( \varphi \) and \( \sigma \) do not refer to particular people, so it would be wrong to think of \( \{ \varphi, \sigma \} \) as a ‘subset’ of the set of people; it is a quotient. (Confusing this quotient with a subset may be the closest I can come to a mathematically rigorous definition of sexism.)

**Fundamental physics**

We of course use much more sophisticated instances of reductionism and co-reductionism constantly, usually a mixture of both, in anything we might study. Often, out of necessity, we ignore what cannot be justifiably ignored, and settle for idealized models drastically simpler than the systems we are really interested in. While this remains an essential and very useful tool in all sciences, what is amazing in the natural sciences is that reductionist strategies work at a deeper level than simply pragmatic idealization.

Reductionism owes its success to the fact that matter really does appear to be composed of indivisible parts. Beginning with the atomic theory, and continuing with subatomic particles, finally to the list of about 17 observed particles that high-energy physicists now take to be fundamental, we now have an astonishingly precise picture of matter—a reductionist’s dream, with a finite set of building blocks and neat rules on how to combine them.

At the other end of the spectrum, what I have called co-reductionism plays a key role in the other fundamental pillar of modern physics, general relativity. In general relativity we ignore almost all properties of matter, considering only its mass. Remarkably, this again seems to be more than a practical convenience: gravity really does not seem to care about other properties of matter.

What seems quite miraculous about this whole situation is that gravity, the one force that is so weak that one only notices it when there is a lot of matter around, is also the one force entirely unconcerned with any distinguishing properties of that matter: it seems to care only about mass and location of that mass. This is probably not actually unreasonable at all; I suspect there is a deep reason for it which we will eventually see mathematically.

Quantum gravity, the attempt to combine general relativity with quantum physics, faces a dilemma, since these theories involve orthogonal types of ‘ignoring.’ Most approaches to the problem can be classified into one of two scenarios. In one scenario, we try to make gravity more like the rest of physics, for example by studying elementary particles gravity (‘gravitons’) or frameworks to generalize them. In the other scenario either we attempt to apply principles of quantum physics, which in fact we know entirely from the world of particles, to the picture of spacetime.

Briefly, either we attempt to augment a theory arrived at by reductionist tactics to include a theory arrived at by co-reductionist tactics, or the other way around. Neither seems obviously bound to work. If our luck does run out, it won’t be a failure of reductionism, and not of mathematics. In fact, the precise and structural ways of thinking provided by mathematics will be our best hope to surpass this obstacle.

**Factoring a miracle**

I have argued that what is truly surprising in physics is not that mathematics—being after all deeply related to physics both historically and in determining foundations—is effective in physics, but rather that reductionist strategies work so well, for apparently purely physical reasons. But there is also a deeper sense of reductionism lying at the heart of Wigner’s argument, and the worldview we are led to if we take it seriously.

On the surface, Wigner’s claim that mathematics is unreasonably effective in physics seems related to Einstein’s famous statement about the world being amazingly comprehensible. But taken in context, it is clear Einstein does not wonder at the comprehensibility of the world in some metaphysical sense, but rather that the world is comprehensible to us:

> That the totality of sensory experience is such that it can be organized through thinking . . . is a fact that we can only marvel at, but which we will never be able to comprehend. We can say: the eternally incomprehensible thing about the world is its comprehensibility.

While it is impressive that one can get so far in principle with extracting mathematical structure from the physical world, the shocking thing is that we have the mental capacity to do this in practice. The real miracle is the level of complexity of the world relative to our own intelligence. This is indeed cause
to marvel. It is one thing for the universe to be sensible in some precise way, and quite another for some entity within the universe to make sense of it.

The desire to ‘factor out’ the human element no doubt stems from reductionist tendencies, where in this case the component we attempt to ignore is ourselves. I see no sensible way to do this.
References

[A] P. W. Anderson, More is different, *Science* **177** (1972) 393–396.

[E] Albert Einstein, Physik und Realität, *J. Franklin Inst.* **221** (1936) 313–347.

[E2] Albert Einstein, Geometry and Experience, http://www-history.mcs.st-and.ac.uk/Extras/Einstein_geometry.html

[D] Freeman Dyson, Missed Opportunities, *Bull. Amer. Math. Soc.* **78** No. 5 (1972) 635–652.

[I] Chris J. Isham, Topos methods in the foundations of physics, in *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, ed. H. Halvorson, Cambridge University Press, Cambridge, 2011.

[PCM] *The Princeton Companion to Mathematics*, eds. Timothy Gowers, June Barrow-Green, Imre Leader, Princeton University Press, New Jersey, 2008.

[GG] Ivor Grattan-Guinness, Solving Wigner’s mystery: The reasonable (though perhaps limited) effectiveness of mathematics in the natural sciences, *The Mathematical Intelligencer* **30** (2008) no. 3, 7–17.

[W] Eugene Wigner, The unreasonable effectiveness of mathematics in the natural sciences, *Comm. Pure Appl. Math.* **13** (1960) 1–14.