The Language of Nature

David Garfinkle

In 1960 Eugene Wigner published an article titled "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" [1]. He gave several examples of areas of mathematics that had developed independently of physics, but that nonetheless proved to be essential in the formulation of twentieth century physics. Wigner concluded that "The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve."

Wigner's view stands in contrast to the view of Galileo in which he used the image of nature and its laws as a book and asserted that that book was written in the language of mathematics [2]. In Galileo's view, when we develop physical theories we are discovering the underlying mathematical order of nature. I will argue that Galileo's view is essentially correct, but that the developments that led to Wigner's view reveal two deep truths: one about the nature of mathematics and the other about the nature of physics.

To begin, it is helpful to recall the development of non-Euclidean geometry [3]. Though expressed in an axiomatic form by Euclid, geometry can be thought of as something akin to a physical theory, with the properties of its lines revealing empirically discovered properties of stretched strings, or the lines of sight of surveying and thus the paths of light rays. However, of the axioms used by Euclid, one seemed less natural than the others: the parallel postulate, which is the statement that given a line and a point not on the line there is exactly one line through the given point that is parallel to the given line (Euclid did not put the parallel postulate in this form; but what he used is equivalent to this). Because the parallel postulate seemed unnatural, there were efforts to derive it from the other axioms of Euclidean geometry, but all such efforts were unsuccessful.

Finally in the 19th century Gauss, Bolyai, and Lobachevsky went in the opposite direction, producing a geometry that used the other axioms of Euclid, but in which

D. Garfinkle (✉)
Department of Physics, Oakland University, Rochester, MI 48309, USA
e-mail: garfinkl@oakland.edu

A. Aguirre et al. (eds.), *Trick or Truth?*, The Frontiers Collection,
DOI 10.1007/978-3-319-27495-9_11

the parallel postulate was replaced by the statement that through the given point there was more than one line parallel to the given line. In modern language, we would say that Gauss, Bolyai, and Lobachevsky had developed the geometry of the hyperbolic plane: a two dimensional space of constant negative curvature.

The development of non-Euclidean geometry marked the divergence of mathematics from physics. Euclidean geometry and non-Euclidean geometry gave different answers from each other, but they were both "right" in the sense that both were consistent systems derived by logic from a set of axioms. Though one could argue about which geometry better modeled the physical world, as long as one uses only consistency as the criterion, the two geometries were on an equal footing. No longer was mathematics to be about finding the "right" axioms (i.e. the ones that seem to come from nature). Instead the mathematician could choose any set of axioms he (or she) liked as long as they were consistent, and then whatever followed from those axioms would be mathematics.

Partly in response to this new found insight, the study of mathematics exploded: more and more systems of greater and greater abstraction and complexity were developed, and the features of these systems were worked out in meticulous detail. This trend continued through the 20th century, with mathematics becoming ever more formal and abstract, and the number of different systems studied by mathematicians ever increasing.

The 20th century also saw the development of new and revolutionary theories of physics: relativity and quantum mechanics. And, miraculously, when the new theories of physics needed new mathematics, lo and behold that new mathematics had *already* been developed: differential geometry for the general theory of relativity, group theory, complex vector spaces and their operators for quantum mechanics. Furthermore, this extraordinarily useful mathematics had been developed by mathematicians who had merely been pursuing abstract subjects with total disregard for any physical meaning or application those subjects might have.

It is that extraordinary coincidence, that the new physics needed new mathematics, *and there it was*, that seemed so miraculous to Wigner. But I will now argue that this development is not as miraculous as it seems. First, let us consider the nature of the new mathematics. Non-Euclidean geometry and the formal and axiomatic approach to mathematics gave the mathematicians enormous freedom. Any formal system of axioms could be considered mathematics: simply develop a system of axioms (as long as they are consistent with each other) and then these axioms along with all theorems derived from them form a mathematical system. But what were the mathematicians to do with all of this newfound freedom? If anything can be mathematics, then how is a mathematician to choose what to work on?

Wigner claimed that mathematicians chose systems and problems solely to exhibit their cleverness in making formal arguments. Well, perhaps that claim has some element of truth to it, and it is not surprising that Wigner, who was a close friend of John von Neumann since childhood, would have that view of mathematics and mathematicians. But just as "choose any set of axioms you like" does not give sufficient guidance in the development of new mathematics, so "exhibit your mathematical cleverness" is also too amorphous a criterion to be a guide in the development of new mathematics.

So how did the mathematicians decide what would be the new mathematics? They did it by a process of abstraction and generalization from the old mathematics: group theory to generalize the properties of the symmetries of objects in Euclidean geometry; differential geometry to generalize curved surfaces in three dimensional Euclidean space; operators and vector spaces to generalize both the properties of the usual vectors of three dimensional Euclidean space and the properties of linear differential equations and their solutions; analysis and topology to abstract and generalize a set of useful techniques for proving the convergence of series, especially those used to finally put calculus on a rigorous footing. This process of abstraction and generalization generated a great number of new mathematical objects and led to another pursuit of modern mathematicians: classification. For each new type of object (group, vector space, manifold, Lie Algebra, etc.) one would aim to produce a complete classification of all possible objects of that type. The process of abstraction and generalization can be thought of as exploring the realm of possible new mathematical systems by starting at the old systems and working one's way outward. The process of complete classification means that in a certain sense this exploration process is very thorough: that any new mathematics that is "sufficiently close" to the old mathematics will be found.

Thus the new mathematics was related to the old mathematics, which was in turn related to the old physics. But why was the new mathematics just what was needed for the new physics? Here the answer has to do with the fact that old physical theories are limiting cases of new physical theories. To cite some well-known examples: Newtonian mechanics is the limit of special relativistic mechanics in the case of small velocity; special relativity is the limit of general relativity in the case of weak gravity (and thus Newtonian gravity is the limit of general relativity in the case of small velocity and weak gravity); classical mechanics is the limit of quantum mechanics for large systems (or more specifically actions large compared to Planck's constant).

But why do we discover the limiting cases first? Because of limited data and Ockham's razor. We, and the things we move and throw, are slow moving objects (compared to the speed of light) and so the most easily accessible data for us on the motion of objects is well described by Newtonian mechanics. But Ockham's razor enjoins us to prefer simple theories to complicated ones, and so Newtonian mechanics was developed before special relativity. We (and the things we move and throw) are large objects, compared to the size of atoms, and so classical mechanics is a good description of the data that was accessible to our experiments before the end of the 19th century, and therefore classical mechanics developed before quantum mechanics.

When new data becomes available, data no longer well described by the old theory, we must develop a new theory. But what new mathematics is needed for the new theory? If the old theory is a limiting case of the new theory, then it is likely that the mathematics of the new theory has something in common with the mathematics of the old theory. In particular, the new mathematics is likely to be some sort of generalization of the old mathematics. But generalization (along with abstraction) is precisely the business of contemporary mathematicians, as it was for the mathematicians of the 19th and 20th century ever since the development of non-Euclidean geometry. Furthermore, since a thorough classification of the possible newly discovered mathematical objects is also the business of mathematicians, it is likely that the particular generalization of the old mathematics needed by the new physics would be one that the mathematicians would discover, even before it was needed. Thus, we should not be surprised that when a new physical theory is developed, the new mathematics that is needed for that new theory is already at hand.

One of the things that struck Wigner as particularly miraculous is the central role of complex numbers in quantum mechanics. Complex numbers were discovered hundreds of years before quantum mechanics as a mathematical trick for finding roots of polynomial equations. But why should a trick for finding roots lie at the heart of our deepest theories of nature? It seems to me that the answer lies in the mathematical concept of a field. Roughly speaking, a field is a number system that has the additive and multiplicative properties of the real numbers. However, the mathematical classification of fields shows that the field concept is very restrictive: there are very few fields, even fewer that contain the real numbers, and of those fields that contain the real numbers the simplest one is the complex numbers. Thus one can think of the complex numbers as the minimal extension of the real numbers. It is therefore not so surprising that when studying the algebraic properties of real numbers, mathematicians stumbled across this minimal extension. And it is also not so surprising that for a system of nature based on the complex numbers, there are limiting cases that use only the real numbers and that physicists found those limiting cases first.

In future developments in theoretical physics, there is likely to be the need for some new mathematics. I will now argue that when that occurs, the effectiveness of the new mathematics will not seem miraculous. Here it seems to me that the situation is somewhat obscured by the present practice of string theory, in which much exotic mathematics is used and in which each new development is hailed as a breakthrough, if not a revolution. However, mathematics is effective in physics (let alone unreasonably effective) when it is an essential part of a theory that has been confirmed by experiment, not when it is part of the formalism of a highly speculative theory that has not been so confirmed. I predict that eventually the current mania for string theory will die down and that in its wake much of the mathematics that it uses will be regarded as not having a role in the description of nature. In the meantime, the mathematicians will go on to develop many new areas of mathematics, so that the subset of mathematics that is used effectively in physics will become an ever smaller fraction of the total amount of mathematics. At that point, it is unlikely to appear miraculous that the particular tiny subset of mathematics used in the description of nature is appropriate for that purpose. Thus it seems to me that the seemingly unreasonable effectiveness of mathematics was a one-time historical phenomenon of the twentieth century: it occurred shortly after physics and mathematics went their separate ways. Thus in that short time (1) mathematicians had developed some mathematics without regard for any physical application it might have, and (2) the new mathematics was developed by generalizing parts of mathematics that were closely related to physics, so that (3) when new physical theories needed new mathematics there was a good chance that this closely related mathematics would be just what was needed. As we get further and further from the 19th century parting of the ways of physics and mathematics, it is likely that these two fields will have less and less in common, and what little overlap there is will seem fortuitous rather than miraculous.

In summary the “unreasonable” effectiveness of mathematics is not so unreasonable after all. Nature and nature’s laws are mathematical, just as Galileo and Newton taught us. It is our job as theoretical physicists to discover those laws. When new laws are discovered, they may use mathematics that has not been used in physics before. If so, we should not be surprised if that mathematics has already been invented by mathematicians for their own purposes. New physics has old physics as a limiting case. Thus it is not too surprising that the mathematics needed for the new physics is something that can be found by abstracting and generalizing the mathematics used in the old physics. Therefore the mathematicians, who develop new mathematics by abstracting and generalizing old mathematics, may develop just what we need even before we need it. This is indeed, as Wigner said, a wonderful gift; but perhaps we do understand it after all.

References

1. E. Wigner, Communications in Pure and Applied Mathematics, 13 (1960).
2. Galileo Galilei The Assayer.
3. For a short summary of the history and properties of hyperbolic geometry see e.g. Chapter 1 of S. Weinberg Gravitation and Cosmology, Principles and Applications of the General Theory of Relativity (1972) (John Wiley and Sons, New York).