Shaking the foundations of mathematics: There are many interpretations of mathematics. The one we use has proved itself. But what would happen to our view of the Universe if we experimented with another?

21 November 1992 by **JOHN BARROW**
Magazine issue 1848. **Subscribe and save**

Mathematics is the language scientists rely on to describe everything in the physical world, from the inner space of elementary particles to the outer space of distant galaxies. Most scientists believe deeply in the mathematical structure of nature; they see their task as simply discovering what this structure is. They view mathematics as an unambiguous business which leaves no scope for subjective contributions, no individuality of choice; it develops inexorably by a logical process in which old symbols are traded for new. But a closer look reveals that it is not quite so simple to pin down exactly what 'mathematics' is.

What is clear is that mathematics works spectacularly well - sometimes rather unexpectedly. In 1914, when Einstein was struggling to endow space and time with a curved geometry, and wanted to express the laws of nature in a form that would be found the same by all observers, no matter what their state of motion, the Hungarian-born mathematician Marcel Grossman introduced him to a little-known branch of 19th-century mathematics called tensor calculus that was tailor-made for his purpose. More recently, particle physicists have discovered that symmetry dictates the way elementary particles behave. Fractals - once no more than an idiosyncratic branch of mathematics - have revealed themselves in countless natural phenomena, from the clustering of galaxies to the structure of crystallising snowflakes.

Basic questions

But this approach - to say that mathematics is valid because it has helped to solve some problems - avoids some basic questions. Is mathematics really something we discover, or do we invent it? Clearly, there is more than one way of interpreting what mathematics is. Each interpretation would lead us to accept a different set of statements as being true, and this could have serious implications for the way we see the physical world. I would argue that we cannot explore the truly fundamental nature of the Universe without a clear understanding of what we mean by mathematics and where its limits lie. If scientists are to encode our observations of the world into some ultimate mathematical 'theory of everything' they must address the question of which mathematics to adopt.

This problem has never been considered an issue in the development of fundamental physics. It first exercised mathematicians during the early years of this century as they wrestled to come to grips with a host of unusual mathematical ideas in logic, set theory and the study of actual infinities which had no simple intuitive counterparts. Mathematicians also faced several bewildering problems which rocked their confidence. Logical paradoxes like that of the barber (A barber shaves all those individuals who do not shave themselves. Who shaves the barber?) or the dilemma of the sets of all sets (Is it a member of itself?) threatened to undermine the entire edifice of logic and mathematics, for who could foresee where the next paradox might surface? In the face of such dilemmas David Hilbert, the foremost mathematician of the day, proposed in 1900 that we cease worrying about the meaning of mathematics altogether. Instead, he argued for an approach that became known as formalism. This says that we should simply define mathematics as a tapestry of formulae that can be created from the collection of all possible consistent initial axioms by manipulating the symbols involved, according to specified rules: mathematics 'is' this vast embroidery of interwoven logical connections.

Hilbert and his followers did not even try to explain why mathematics is so uniquely suited to describing nature. For them mathematics did not have any meaning: formulae existed on pieces of paper but mathematical entities had no other claim to existence whatsoever. The formalist would no more offer an explanation for the mathematical character of physics than seek to explain why physical phenomena do not obey the rules of chess or blackjack. Hilbert thought this strategy would, by definition, rid mathematics of all its problem areas. Given any mathematical statement it would be possible to determine whether it was a true conclusion of a set of starting assumptions by working through the network of logical connections. Hilbert and his followers were confident that all known mathematics would fit into this straitjacket.

Then, totally unexpectedly, the enterprise collapsed overnight. In 1931, Kurt Godel, an unknown young mathematician at the University of Vienna, showed Hilbert's goal to be unattainable in systems large enough to include ordinary arithmetic. He showed that whatever set of consistent rules one adopts for manipulating mathematical symbols, there must always be some statement, framed in the language of these symbols, whose truth or falsity cannot be decided using those axioms and rules. Worse still, there is no way of telling whether or not the starting assumptions are logically consistent or not. Suddenly, mathematical truth had become larger than axioms and rules. If you try to solve the problem by adding a new rule or a new axiom you just create new undecidable statements. To understand mathematics fully you must go outside mathematics.

With the wind taken out of their sails, the formalists drifted away from their goal. They continued to develop mathematical formalisms, but they now knew that consistency and completeness were unachievable. A more limited formalism was taken up again in 1939 by a group of French mathematicians writing under the pseudonym 'Nicolas Bourbaki'. They tried to make sense of the decidable part of mathematics by emphasising the presence of common structures in different branches of mathematics - geometry, arithmetic, algebra and so forth. They saw mathematics as a living, growing structure that was created by human mathematicians. But this approach never appealed to scientists and today Bourbaki's project attracts strong criticism from many mathematicians because of its separation of mathematics from problem solving and the contemplation of nature. Over the past 20 years it has also proved a controversial aspect of the style in which mathematics is taught.

Human invention

But Bourbaki's emphasis on the human creation of mathematics is accepted by many other users of mathematics in subjects such as economics, sociology, anthropology and psychology, where there is much emphasis on human activities. To these 'inventionists', mathematics appears to be just another of these activities - it is simply what mathematicians do. Mathematical entities like sets or triangles would not exist if there were no mathematicians. We invent mathematics, we do not discover it. A mathematical description of the world then looks less impressive - it is only those physical phenomena best suited to mathematical description that we have been able to uncover.

If mathematics is just a useful human invention, we would expect significant cultural differences within the subject. But though there are discernible styles in the presentation of mathematics and in the type of mathematics investigated in different cultures, mathematicians from totally different economic, cultural and political backgrounds at different times throughout history have discovered the same mathematical theorems. This unusual phenomenon distinguishes creative mathematics from music or the arts. Pythagoras's theorem was independently discovered many times by different thinkers (although Pythagoras probably was not one of them), but Shakespeare's Hamlet or Beethoven's fifth symphony are unique.

This seems to imply that the foundation of mathematics lies outside the human mind and is not totally fashioned by our human way of thinking. Our own existence and that of any conscious observers requires the existence of some order in nature. The study of this order constitutes part of what we call mathematics. But the most telling objection to the inventionist view of mathematics is the evolution of our minds. If mathematics in some sense comes out of our minds or is copied from the structure of the natural phenomena that we witness, where does this formative mathematical structure come from?

We could argue instead that the mathematical structure of the world is impressed on the human mind by an evolutionary process that rewards faithful representations of reality with survival - unfaithful images of reality have a low survival value. Our eyes code information about the real nature of light and our ears something of the real properties of sound, which is why they have evolved as effective receptors of light and sound. In the same way our minds bear some image of the ordered structure of nature. If that order did not exist, or if we were unable to store accurate representations of it, then we would not exist either. How much of the structure of the world and its correct encoding by our minds is necessary for our evolution and survival is far from understood. Yet the most dramatic examples of the effectiveness of mathematics arise in those areas of human inquiry, like the physics of elementary particles or the evolution of the Universe as a whole, which are most divorced from the human scale.

The idea behind this criticism of inventionism is that there is an order in the physical Universe that is independent of any 'observers'. A faithful representation of part of it must be possessed by any 'observer' of the Universe. This interpretation of mathematics - mathematical Platonism - is part of a long tradition of 'natural philosophy' which says the world is in some deep sense mathematical. Mathematical concepts are not invented; they exist and are discovered by mathematicians. Mathematics would exist even if mathematicians did not. This idea was summed up by James Jeans in 1930 when he suggested that 'the Great Architect of the Universe now begins to appear as a pure mathematician'. And indeed, if the entire material Universe can be described by mathematics, as modern cosmology assumes, this implies that there must exist some immaterial logic that is larger than the material Universe.

A working philosophy

Formalism and inventionism are uneasy about the unreasonable effectiveness of mathematics in the description of nature, but Platonists can use it as crucial evidence. Many scientists and mathematicians work within its framework: they carry out their daily work as if the Platonic viewpoint were true, even though they might be loath to defend it too strongly at the weekend.

Realism of this sort about mathematical entities has a most extraordinary consequence. If there is a mathematical description of the evolution of the Universe in which conscious observers like ourselves arise, then intelligent observers must exist within the formalism of mathematics because they can be described with that mathematics. But where is this world of mathematical objects? How do we make contact with it? If mathematical entities exist beyond the physical world of particulars that we experience, we cannot treat mathematical knowledge like other forms of knowledge about the physical world that we acquire through our senses. We treat these as meaningful because the objects we learn about interact with us in a causal way. Mathematical entities do not affect us like this.

So the Platonic view of mathematics leads us into deep problems of metaphysics. Godel recognised this and maintained that there exists some immaterial reality with which we can have 'another kind of relation'. The British mathematician Roger Penrose, another Platonist, takes Godel's famous formula, (which encodes into arithmetic the logical paradox 'this sentence is false' and which demonstrates its own unprov-ability) as a mark of the non-algorithmic nature of human consciousness. This unprov-ability is something we 'see' is true but it cannot be proved within the logical system, or algorithm, being used, even though this algorithm is what some artificial intelligence enthusiasts seek to define as the workings of a mind. This is a surprising claim. It implies that anyone who cannot grasp the meaning and truth of what Godel is saying is in some sense not fully conscious. It would also mean that only those mathematical formalisms rich enough to contain arithmetic would lead necessarily to the existence of conscious observers within them. Less rich mathematical systems, like Euclidean geometry or arithmetic without the operation of subtraction, would not.

To find an alternative viewpoint, we have to go back again to the early years of this century, to the foment of uncertainty about logical paradoxes that spawned formalism. Another response to this turmoil was called 'constructivism' or 'intuitionism'. It first emerged in the 1880s, and according to the German mathematician Leopold Kronecker, one of its creators, it was the recognition that 'God made the integers, all else is the work of man'. What he meant by this after-dinner remark was that we should accept only the simplest possible, intuitively obvious mathematical notions - the whole numbers 1,2,3,4, . . . and the idea of counting - as a starting point and then derive everything else from these, step by step. The idea is to avoid manipulating counterintuitive entities like infinite sets - the collection of all the even numbers, for example - about which we could have no concrete experience.

This rather conservative approach to mathematics was taken up with fanatical devotion by the Dutch mathematician Luitzen Brouwer from 1907 until 1928, when he withdrew from mathematics, embittered and estranged from other mathematicians by his attempts to impose the constructivist dogma on them. His behaviour gave constructivism a rather bad name until 1967, when the American mathematician Errett Bishop pruned it of its philosophical questions at just the time when the computer was beginning to emerge as a practical tool for mathematicians and scientists. Computer language seemed a natural one for putting the constructivist picture of step-by-step mathematics into practice, and Bishop resurrected interest in constructivist philosophy among mathematicians.

For the constructivist, mathematics consists only of the collection of statements that can be constructed in a finite number of deductive steps from the natural numbers. The 'meaning' of a mathematical formula is then simply the finite chain of computations that have been used to construct it. This view sounds straightforward, but it has dramatic consequences. It creates a new category of mathematical statement. The status of any statement can now be true, false or undecided - rather as in Scottish law where courts can return a verdict of guilty, not guilty or not proven.

Pre-constructivist mathematicians had developed all manner of ways of proving formulae to be true that cannot be put into a finite number of constructive steps. One famous method, beloved of the ancient Greeks, was the reductio ad absurdum. In this method they began by assuming something was true. If, from that assumption they deduced something contradictory (like 2 = 1), they concluded that their original assumption must have been false. But this argument carries the hidden assumption that a statement is either true or false. So it is an invalid move according to the constructivists' rules, which say that a statement is only true if it is proved explicitly in a finite number of deductive steps. The constructivist approach outlaws the whole body of mathematical theorems which prove that something exists - because its nonexistence creates a logical contradiction - but do not construct an example of it.

Constructing consequences

Constructivism would have interesting but largely unexplored consequences if adopted in physics, because many important physical theories like Einstein's general relativity or Niels Bohr's quantum mechanics use non-constructive reasoning at crucial points to deduce properties of the Universe. To most mathematicians such a strategy seems tantamount to fighting with one arm tied behind your back. The cosmological singularity theorems developed by Stephen Hawking and Roger Penrose between 1966 and 1972, for example, are non-constructive. They give sufficient conditions for the existence of a beginning to time by deducing a logical contradiction if no beginning exists. They do not explicitly construct the singular beginning, so they are not 'true' in constructivist mathematics.

Yet, looked at closely, constructivism does seem rather peculiar. It defines mathematics anthropocentrically as the sum total of all finite, step-by-step deductions from the bedrock of human intuition - the natural numbers. It implies that there is no mathematical existence before this construction takes place. Philosophically speaking, it puts humans back at the centre of the Universe, where they have not been since Copernicus and others began to regard them as external observers of the physical world: a view that heralded the scientific revolution of the 16th and 17th centuries and which has proved so successful for the development of modern science.

Besides, the notion that there exists a universal human intuition for the natural numbers does not have historical support. There are primitive cultures where counting proceeded only to 2, and none where any abstract notion of number seems to have emerged. The constructivist cannot tell if my intuition is the same as yours, or whether human intuition has evolved and will evolve further in the future. Mathematics created from human intuition is a time-dependent phenomenon that depends on the mathematician constructing it. It is a branch of psychology.

Constructivism leaves us with several unanswered questions. Why should we start with the natural numbers? What counts as a possible constructive step? Why are some constructions more useful and applicable to the real world than others? Most of all, why are some non-constructive concepts, such as the continuum of real numbers, so useful in the study of the physical world? Moreover, infinite sets of entities, which cannot be listed in a finite number of steps, have arisen somehow in human intuition.

Constructivism may nevertheless have something to teach us about the mathematical character of nature. Godel showed that there must always be some statements whose truth we can neither prove nor disprove. But what about all those statements whose truth we can decide by the traditional constructive and non-constructive methods of mathematics? How many of them could the constructivists prove? And what better constructivist is there than a computer? Is it possible in principle to build a computer which reads input, displays the current state of the machine and possesses a processor for determining a new state from its present one, and then use it to decide whether a given statement is true or false after a finite time? Contrary to the expectations of many mathematicians, the answer is no. In 1937 Alan Turing in Cambridge, and Emil Post and Alonzo Church in Princeton showed that there are statements whose truth requires an infinite time to decide. They are infinitely more complex than can be probed by step-by-step computation.

The idealised computer outlined above is called a Turing Machine; it is the essence of every computer. The operations it can perform are called 'computable', which means that we can fabricate a device from matter whose behaviour mimics that operation. Typical devices might be swinging pendulums or electrical impulses. Conversely, physical devices like these can be well described by computable mathematical operations. The fact that nature is well described by mathematics is equivalent to the fact that the simplest mathematical operations, such as addition, along with the more complicated operations used in science, are computable functions. If they were not, they could not be equivalent to any natural process and the usefulness of mathematics would be limited. Even though the world might be mathematical in character, we would not find mathematics a fruitful language to predict or describe its workings.

In recent years there has been a growing interest in the possibility that we might need to rethink our view of the laws of nature as great symmetries, replacing it with one based on the computational process. This perspective views the laws of nature as a form of 'software' running on the 'hardware' of elementary particles of matter and energy. It may be that the discontinuous character of computation is linked as closely to an underlying constructivist definition of mathematics as the particle physicists' view is wedded to the continuity of space-time and the Platonic philosophy of mathematics.

John D. Barrow is professor of astronomy at the University of Sussex and the author of Theories of Everything: The Quest for Ultimate Explanation (hbk, Oxford University Press, 1991; pbk, Vintage, 1992) and more recently Pi in the Sky: Counting, Thinking and Being (Oxford University Press, October 1992).