Einstein and geometry

Michael Atiyah

School of Mathematics, The University of Edinburgh, Edinburgh EH9, 3JZ, UK

Einstein initiated and stressed the role of geometry in fundamental physics. Fifty years after his death the links between geometry and physics have been significantly extended with benefits to both sides.

Keywords: Einstein, geometry, quantum mechanics, topology, string theory.

1. General relativity

EINSTEIN is generally recognized as the greatest physicist of the 20th century and perhaps the greatest physicist since Newton, though Faraday and Clerk Maxwell are close competitors. Einstein is a case where popular acclaim and scientific standing are in agreement. But unlike Newton, Einstein was not a mathematician. He used mathematics in an essential way but he did not create it and he relied on his colleagues for technical help. It is all the more remarkable that his ideas have triggered great advances in geometry, even in parts of the subject apparently far removed from physics.

I will attempt to describe and explain how this has come about. But first I should make some general remarks about the relation between physics and mathematics. The conventional view is that mathematicians have developed machinery for studying numbers (which might represent physical quantities) and the way in which those relate to each other in the form of equations. Physicists then use this language and embody their conclusions in 'laws' described by equations. Thus Newton's gravitational theory is described by the inverse square law of mutual attraction, while the fundamental laws of electro-magnetism are encoded in Maxwell's equations.

While this orthodox view is formally correct, it hides some essential features. In physics the starting points are the concepts: particles, forces, space, time, motion, interaction. Objects are seen to move around and act on one another. The secondary part of the story is the taking of measurements by the experimental scientist. Numbers are written down, tabulated, compared.

The earliest part of mathematics to be studied in depth was geometry, in the hands of the Greeks. The basic concepts here are: points, lines, angles, triangles, circles and their mutual relation. Numbers, giving distances and areas come shortly thereafter, but equations did not enter the picture until the work of Descartes in the 17th century.

The connection between physics and geometry starts at the conceptual stage in a fully 3-dimensional picture of the world, and has nothing to do with any reference frame which one may choose in which to take measurements. It is not easy to move from physics to geometry without choosing (x, y, z) coordinates and writing equations, but it is more fundamental. Descartes' introduction of coordinates may have been an essential step in the formalization of mathematical physics but it was also an abdication: it gave up on trying to understand physics geometrically.

Newton understood this, which is why he presented his *Principia* in geometric form, but this was too difficult for posterity who followed the ideas of Descartes and Leibniz.

This brief philosophical review is essential if we want to understand how Einstein's ideas came to influence geometry. As we all know, Einstein's monumental contribution was the replacement of the Newtonian theory of gravity by what is called General Relativity. This theory has two essential features, the first is to move from 3-dimensional geometry to 4-dimensional geometry by incorporating time as a fourth variable. This is the content of Special Relativity, but the second key step is to interpret gravitation as the curvature of this 4-dimensional space–time geometry.

Standard text-books make great play with the technical details, introducing coordinates, writing equations and then showing that the resulting physics is independent of the choice of coordinates. To a geometer this is perverse. The fundamental link is from physics to geometry, from force to curvature and the algebraic machinery that encodes this is secondary. God created the universe without writing down equations!

2. Electro-magnetism

As far as gravity is concerned, Einstein's General Relativity is a beautiful and complete theory. But as Einstein realized it has to be extended to account for other physical forces, the most notable being electro-magnetism. It is perhaps no accident that the first and most significant step in this direction was taken by a mathematician – Hermann Weyl. He showed that, by adding a fifth dimension, electro-magnetism could also be interpreted as curvature. His idea was that the size of a particle could alter as it passed through an electro-magnetic field. In analogy with railways it was called a gauge theory, and this name has stuck through subsequent evolutions of the theory.

e-mail: m.atiyah@ed.ac.uk

Unfortunately for Weyl, Einstein immediately objected on physical grounds that this would have meant different atoms of, say hydrogen, would have different sizes depending on their past history, in contradiction with observation. Given this devastating critique, it is remarkable but fortunate that Weyl's paper was still published, with Einstein's objection as an appendix. Clearly the beauty of the idea attracted the editor, despite the fatal flaw. In fact, beauty often wins such contests, because with the advent of quantum mechanics, with its complex wave functions, it was pointed out by Kaluza and Klein that Weyl's gauge theory could be salvaged if one interpreted the variable as a phase rather than a length. A pure phase shift by itself is not physically observable and so Weyl's theory avoids the Einstein objection.

Quantum mechanics

While quantum mechanics thus came to the rescue of Weyl's gauge theory and so continued the Einstein programme of geometrizing physics, it also seemed to demolish the whole idea. While quantum mechanics is a very subtle and beautiful mathematical theory, it strays very far from geometry and is conceptually difficult to comprehend. In fact, as is well known, Einstein never fully accepted quantum mechanics as the final word. He disliked its philosophical basis with its need for probability and uncertainty.

While conceding its great practical success, Einstein remained opposed to quantum mechanics to his dying day. Increasingly he was regarded by the younger generation of physicists as being obstinate and out of touch. His continued search for a unified field theory only confirmed this widely held opinion.

3. Nuclear forces

Einstein and Weyl, who both went to the Institute for Advanced Study in Princeton as refugees from Germany in the 1930's, died in 1955, the year I myself went to Princeton as a fresh Ph D. This was also the year when Yang–Mills theory was born, the theory which developed in due course into the standard framework for understanding the 'weak' and 'strong' forces which operate on the nuclear scale and are believed, together with gravitation and electromagnetism, to provide all the fundamental forces of nature.

Yang–Mills theory can be roughly understood as the natural extension of Maxwell's theory in which the angular phase is replaced by a phase specified by rotation in a higher dimensional 'internal space'. This internal space is not part of our usual space–time but is additional to it, just as the Maxwell phase was interpreted as a fifth dimension. There is one fundamental difference between angles (rotation in a plane) and rotations in 3 or more dimensions. Two such rotations, about different axes, do not in general 'commute', that is to say that the result of performing the two rotations one after the other depends on the order in which they are performed. This is easy to verify by considering rotations of the earth. Consider for example a rotation A about the North Pole/South Pole axis of say 20° in a westward direction, and a rotation B around the axis through Chennai (and its antipode), which takes Bangkok to a position due North of Chennai (somewhere in Northern Kashmir). Performing first B and then A will take Bangkok, via Kashmir, to northern Iran. On the other hand, performing A first will take Bangkok approximately to Chennai, so that following this by B will leave it there. The results obviously differ – Chennai is not in Iran!

This non-commutativity of rotations has major consequences for Yang–Mills theory, making it a much more complicated and subtle theory than Maxwell theory. In particular it becomes non-linear, which has profound mathematical and physical consequences.

It is somewhat ironic that the ideas of Yang and Mills developed quite independently of Weyl and Einstein and that there was little interaction with them. No doubt the generation gap was too large. In addition, Yang–Mills theory was a quantum theory, still in its infancy, and the full geometrical implications were not yet apparent.

With the belated recognition that all four of the fundamental forces of nature were geometrical, one might have said that Einstein's dream of a unified field theory was finally realized, even if it came after Einstein's death. In fact, as has just been indicated, this was only partially true because of the presence of quantum theory. On the one hand the quantum aspects made the theory extremely difficult and sophisticated, taking further decades to unravel. On the other hand, Einstein's philosophical objections would remain. He would still be dissatisfied from beyond the grave. Nevertheless physicists now grudgingly acknowledge that Einstein's intuition was in part justified and that the revolution he introduced in General Relativity of geometrizing physics has proceeded much further. Perhaps the verdict would be that the final outcome of the long Einstein–Bohr arguments was a draw, with the big proviso that 'finality' has not yet been achieved.

4. String theory

At this point in the development, although geometry provided a common framework for all the forces, there was still no way to complete the unification by combining quantum theory and general relativity. Since quantum theory deals with the very small and general relativity with the very large, many physicists feel that, for all practical purposes, there is no need to attempt such an ultimate unification. Others however disagree, arguing that physicists should never give up on this ultimate search, and for these the hunt for this final unification is the 'holy grail'.

In the past thirty years a promising framework has appeared in which such a unification seems conceptually possible. This is 'string theory', based on the simple idea that point particles should be replaced by one-dimensional objects – strings, either open (with free ends) or closed (in circular form).

String theory, which is as yet unfinished and incomplete, involves yet more geometry beyond Yang–Mills. In the first place a string moving in time spans a surface which has its own geometry. For example the surface may acquire holes, a topological feature with profound implications, already known in mathematics. In the second place consistency of the physical theory requires that the string should be not just in ordinary 3-dimensional space, but in one of 9 dimensions (or 10 if one includes time). Both of these open up vast new (geometrical) territories which strengthen the link between geometry and physics. This works both ways, first large amounts of mathematics developed over previous centuries suddenly become relevant and available for physicists to use. Second, and perhaps more surprising, the ideas of physics including quantum field theory feed back into mathematics and lead to surprising developments. In fact the mathematical activity generated by this interaction with physics is, in my opinion, the most exciting development in mathematics of the past decades, and we seem still to be in the early stage.

5. M-theory

After the initial rapid development of string theory, a drawback appeared when it was realized that there were five different competing models of string theory. There seemed no reason why nature should prefer one to another.

But it was eventually discovered that all these five string theories were, in a subtle way, equivalent to each other. The best analogy is provided by the basic calculus of analytic functions, in which a function $f(z)$ can be expanded as a power series in z. As a simple example, the binomial theorem tells us that

$$(1 + z)^3 = 1 + 3z + 3z^2 + z^3.$$

However if we introduce a variable $u = 1–z$, then

$$(2–u)^3 = 8–12u + 6u^2–u^3,$$

represents the same function, so the two polynomials are really equivalent under a simple change of variable. String theories are like such power series expansions, but the equivalence between two of them is much subtler than a change of variable.

The five different string theories are now seen as different viewpoints of one underlying theory, which is not yet known but has been christened 'M-theory'. By analogy you might be given the power series expansion of say $(1 \quad –z)$ about several different values of z and you might be able to check that they were equivalent without recognizing that the function was a simple square root.

While a proper understanding of M-theory still eludes us, much is now known about it. In particular the various geometric results that have emerged from string theory become related in interesting but mysterious 'dualities' whose real meaning has yet to be discovered.

No one can predict what the future holds in store for M-theory. Are we nearly there, is the final understanding just round the corner? Will it come from a few more technical tricks or will it require some fundamental breakthrough? The biggest question of all and the one that Einstein would still be asking is: can M-theory be properly understood within the present framework of quantum mechanics or do we need to look for new foundations? I confess that I myself remain an Einsteinian and would be happy to see quantum mechanics replaced by something deeper. This remains, as in Einstein's day, a minority opinion but one shared for example by Roger Penrose.

6. Topology

I have alluded at various stages to the impact that these physical theories have had on geometry, without providing much detail. Let me now try to rectify this.

Classical physics, describing various forces, is closely linked (via Einstein and Maxwell) with notions of curvature in geometry. The connection between physics and geometry is therefore local: we can study the forces in a small piece of space-time and compare it with the local geometry. By contrast quantum physics is not related to geometry in this way. Its relation is 'global' and can only be seen in the whole picture (even if we are dealing with microscopic objects). The global aspect of geometry that is involved is 'Topology', such as the study of holes in surfaces or of knots in 3-dimensional space.

The first indication that quantum mechanics was related to topology was in the argument of Dirac which explained why the electric charge of any particle was an integer multiple of the charge of the electron. The integers came in essentially as 'winding numbers', counting the number of circuits made by a closed path. This number is topological because it does not depend on the detailed local geometry of the circuit, how long or wiggly it is, but only its overall global behaviour.

Winding numbers are related to the circle and hence to the angular phase of electro-magnetism. There are similar but more complicated topological properties associated with the higher dimensional phases of Yang–Mills theory so that the relation between quantum theory and topology carries over to the other forces. In addition, as pointed out earlier, moving strings generate surfaces which may have holes and these are topological in nature.

So the physics of string theory and M-theory is replete with topological information and many intricate and subtle aspects of the quantum theory are related to this underlying topology.

So what kinds of specific geometrical/topological results have emerged from the interaction with physics? In fact these are quite diverse and cover many types of problems. Here is a short list.

Knot invariants

The study of knots (closed pieces of string) is a standard but difficult branch of topology. The key problem is to find 'invariants' which will distinguish essentially different knots. An invariant is something (a set of numbers) which can be calculated from a picture of the knot, but is unaltered if we move the knot around to get a different picture. In the 1970's the world of topologists was astounded when the New Zealand mathematician, Vaughan Jones discovered a new type of invariant which helped to solve 100-year old problems. Shortly after, Edward Witten gave a physical explanation of the Jones invariants which cast new light on them and led to much further progress.

Donaldson invariants

Geometers have studied the topology of closed surfaces and their higher-dimensional analogues (manifolds) for a long time. But a remarkable breakthrough came in the early 1980's when Simon Donaldson found some totally new and unexpected invariants of 4-dimensional manifolds. These were based on the Yang–Mills equations of physics but it was not until later that Edward Witten again showed how to interpret Donaldson's invariants in terms of quantum field theory. Later still, using duality ideas from string theory, Witten and Seiberg made a significant improvement of Donaldson theory which led to solutions of old problems.

Counting curves

Classical algebraic geometers, ever since the time of Descartes, studied curves in the plane given by polynomial equations. It is a natural question to ask how many curves there are of a given type, passing through a given number of points. For example there is a unique straight line through any two points and a unique conic (ellipse, etc.) through five points. The question gets harder as the curves get more complicated and given by polynomials of higher degree. Quite remarkably, ideas from string theory have led to a complete solution of this problem.

These and other examples are now part of a broad area of 'quantum mathematics' – an evocative term which correctly conveys the origin of the ideas and results but is very loosely used and ill-defined. One of the big challenges for mathematicians at the present time is to see if one can understand these new mathematical theories without recourse to the physical background. Alternatively it may become necessary to incorporate or absorb various physical ideas into rigorous mathematics.

The converse process of providing rigorous mathematical treatment of quantum field theory, string theory, M-theory appears a very distant prospect. It will certainly have to wait till physicists have sorted themselves out and allowed the dust to clear.

6. Conclusion

Einstein would, I think, have been both surprised and gratified by the extent to which his geometrization of physics has progressed. The mathematical by-products would have surprised him even further. But the fact that his ideas were so fruitful would only encourage him in his fundamental beliefs. In particular he would still be encouraging us to dig beneath the mysteries of quantum mechanics. In another century we might find what Einstein was looking for.