The Interaction between Geometry and Physics*

Michael Atiyah

School of Mathematics
University of Edinburgh
James Clerk Maxwell Building
The King's Buildings
Mayfield Road
Edinburgh EH9 3JZ
Scotland
U.K.
m.atiyah@ed.ac.uk

Dedicated to Israel Moisevich Gel'fand on his 90*th birthday.*

Subject Classifications: 5802, 5302, 8102

1 Introduction

The theme of this conference is "The Unity of Mathematics," embodying the attitude of Gel'fand himself as demonstrated in the wide range of his many original works. I share this outlook and am happy to describe one of the most fascinating examples, representing the unity of mathematics *and physics*. The speakers were also encouraged to look to the future and not be afraid to speculate—again, characteristics of Gel'fand. In my case, this is perhaps an unnecessary and even dangerous injunction, since my friends feel that I am already too much inclined to wild speculation, and very rash enthusiasm should be dampened down instead of being whipped up. Nevertheless, I will indulge myself and try to peer into the future, offering many hostages to fortune.

Mathematics and physics have a long and fruitful history of interaction. In fact, it is only in recent times, with the increasing tendency to specialization in knowledge, that any clear distinction was drawn between the two. Even when I was a student in Cambridge around 1950, we studied "natural philosophy," which included physics and mechanics, as part of the mathematical tripos. Going further back, it is a moot point whether mathematicians or physicists should claim that Newton was one of theirs.

* This is a slightly extended version of the lecture delivered at Harvard and takes account of some of the comments made to me afterwards.

The great theoretical breakthroughs in physics at the end of the 19th century and the beginning of the 20th century: electromagnetism, general relativity, and quantum mechanics were all highly mathematical, and it is impossible to describe modern physics in nonmathematical terms. Michael Faraday was the last great physicist who was unskilled in mathematics. So much has mathematics pervaded physics that Eugene Wigner has, in a much-quoted phrase, referred to “the unreasonable effectiveness of mathematics in physics.”

The question whether this mathematical description of physics reflects “reality” or whether it is an imposition of the human mind is a perennial and fascinating problem for philosophers and scientists alike. Personally I am sure that new insights from neurophysiology, on how the human brain works, will shed much light on this age-old question and probably alter the very terms in which it is formulated.

Turning from the broad sweep of history to more contemporary events, one can, however, see some sharp oscillations in the synergy between mathematics and physics. The period (after the 1939–1945 war) of the great accelerators with their plethora of new particles, and the struggles of theorists with the infinities that plagued quantum field theories were far away from the concerns of most mathematicians. True, there were always mathematicians trying desperately to lay foundations, far behind the front line, and physicists themselves displayed great virtuosity in handling the techniques of Feynman diagrams as well as the symmetries of Lie groups that gradually brought order to the scene. But all this owed little to the broad mathematical community, unless one includes converts such as Freeman Dyson, and in turn it had little impact on mathematical research.

All this changed abruptly in the middle 1970s after the emergence of gauge theories, with their differential-geometric background, as the favored framework for the quantum field theory of elementary particles. Not only was there now a common language but it was soon discovered that some of the most delicate questions on both sides were closely related. These related the “anomalies” of quantum field theory to the index theory of elliptic differential operators.

Suddenly a new bridge was opened up or, to use a different metaphor, two groups digging tunnels from two ends suddenly joined up and found that the join fitted as beautifully as if it had been engineered by Brunel himself. This time the mathematicians were not building foundations; they were in the forefront where the action was.

I remember vividly those heady days and, in particular, a meeting I had in 1975 with the physicists at MIT, including Roman Jackiw and the young Edward Witten (who impressed me even then). I recall Jackiw asking whether this new interaction between the two sides was a short love affair or a long-term relationship!

Well, here we are, celebrating the silver wedding anniversary of this now firmly established marriage. The past 25 years has seen a really spectacular flowering, with tremendous impact both ways. The younger generation of theoretical physicists has rapidly mastered much of 20th century mathematics in the fields of algebraic geometry, differential geometry and topology. Many of them can manipulate spectral sequences with as much panache as the brightest graduate student in topology, and we mathematicians are constantly being asked the most searching and recondite questions in geometry and topology which stretch our knowledge to its limits.

On an occasion like this it seems appropriate to take a broad view and so I will try to survey rapidly the impact that the new physics has had on geometry (in the broad sense). This has usually taken the form of predictions, with great precision and detail, of some unexpected results or formulae in geometry. These predictions rarely come with any formal proof, though sometimes proofs can, with effort, be extracted from the physics. More often mathematicians are reduced to verifying these unexpected formulae by indirect and less conceptual methods.

What is surprising, beyond the wide scope of the results in question, is how successful the program has been. Despite the absence of any firm foundations, physical intuition and skillful use of techniques, has not yet led to false conclusions. I am tempted to reverse Wigner's dictum and wonder at "the unexpected effectiveness of physics in mathematics."

2 The background

It may be helpful to start by reviewing rapidly the parts of geometry and of physics which have been involved in this new interaction.

Let me begin with geometry. As indicated above the differential geometry of bundles, involving connections and curvatures, is basic. The link with physics goes back essentially to Hermann Weyl's attempt to interpret Maxwell's equations geometrically, and the later improvement by Kaluza. This was just the abelian case of $U(1)$-bundles, but the nonabelian case, involving general Lie groups G, is much more sophisticated and its full mathematical development came much later.

Another key component goes back to the pioneering work of Hodge with his theory of harmonic forms, in particular the refined theory of Kähler manifolds with application to algebraic geometry.

It was Witten who pointed out that Hodge theory should be viewed as supersymmetric quantum mechanics, thus providing an important bridge between key concepts on the two sides. Moreover, when extended to quantum field theories, it showed mathematicians that physicists were trying to make sense of Hodge theory in infinite dimensions. Their success in this venture depended on subtle ideas of physics, going back to Dirac, which mathematicians had to absorb.

In the first few decades after the 1939–1945 war, topology was taking center stage in geometry. New concepts and techniques led to a good understanding of global topological problems in differential and algebraic geometry, culminating in Hirzebruch's famous generalization of the Riemann–Roch theorem. This involved a skilled use of algebraic machinery centering around the theory of characteristic classes and the remarkable polynomials originally introduced by J. A. Todd.

All of this was motivated by internal mathematical questions derived in the main from classical algebraic geometry, now augmented by the powerful machinery of sheaf cohomology of Leray, Cartan and Serre. Any suggestion that this might have relevance to physics would have been met with disbelief.

In fact, the most remarkable fact about the new geometry–physics interface is that topology lies at its heart. In retrospect the roots of this can be seen to go back to Dirac.

His argument explaining the quantization of electric charge—where all particles have an electric charge which is an integer multiple of the charge of the electron—is essentially topological. In modern terms, he argued that a charged particle, moving in the background field of a point magnetic monopole, had a quantum-mechanical wave function which was a section of a complex line bundle (defined outside the monopole). Thus, while classical forces can be expressed purely locally by differential geometric formulae, quantum mechanics forces a global topological view and the *integer* topological invariants correspond to *quantized* charges.

The full implications of this link between quantum theory and topology only emerged when string theory appeared in physics with its Kaluza–Klein requirement for extra dimensions above the four of space–time. The geometry and topology of the extra dimensions provided a strong link with the mainstream development of contemporary geometry.

The role of Lie groups and symmetry in physics has been clear for some time and this already has geometric and topological implications, but the higher dimensions of Kaluza–Klein, as mediated by string theory, involve manifolds which are not necessarily homogeneous spaces of Lie groups. This means that algebra alone is not the answer, and that the full power of modern geometry, including Hodge theory and sheaf cohomology, is required.

As mentioned briefly in Section 1, a key connection between the geometry and the physics came from "anomalies" and their relation to index problems. These were a natural extension of Hirzebruch's work on the Riemann–Roch theorem and the famous Todd polynomials, and variants of them now appeared as having important physical significance. In fact, Hirzebruch's work had been generalized by Grothendieck with his introduction of K-theory and, in its topological version, this turned out to be a very refined tool for investigating anomalies in physics. Some of these are purely global, having no local integral formulation, and K-theory detects such torsion invariants. These, and other clues, indicate that K-theory plays a fundamental role in quantum physics but the deeper meaning of this remains obscure.

Finally, I should say a word about the role of spinors. Ever since they arose in physics with the work of Dirac they have played a fundamental part, providing the fermions of the theory. In mathematics spinors are well understood algebraically (going back to Hamilton and Clifford) and their role in the representation theory of the orthogonal group provides the link with physics. However, in global geometry, spinors are much less understood. The Dirac operator can be defined on spinor fields and its square is similar to the Hodge–Laplace operator. Its index is given by the topological formula referred to above in connection with anomalies. However, while the geometric significance of differential forms (as integrands) is clear, the geometric meaning of spinor fields is still mysterious. The only case where they can be interpreted geometrically is for complex Kähler manifolds where holomorphic function theory essentially extracts the "square root of the geometry" that is needed. Gauss is reputed to have said that the true metaphysics of $\sqrt{-1}$ is not simple. The same could be said for spinors, which are also a mysterious kind of square root. Perhaps this remains the deepest mystery on the geometry–physics frontier.

3 Dimensional hierarchy

Although string theory may require higher dimensions at a fundamental level, at normal energy scales we operate in a space–time of four dimensions, and the extra Kaluza–Klein dimensions merely determine the kinds of fields and particles that we have to deal with.

Since four-dimensional theories present many serious problems it is useful to study simpler "toy models" in low dimensions. We can then think of a dimensional hierarchy where the theory gets more complicated as we increase the dimensions. In general, we write $D = d + 1$ for the space–time dimension, d being the space dimension.

For $d = 0$, we just have quantum mechanics and the associated mathematics of (finite-dimensional) manifolds, Lie groups, etc. For $d = 1$, we get the first level of quantum field theory, which involves things like loop spaces and loop groups. Much of this can now be treated by mathematically rigorous methods, but it is still a large and sophisticated area. For $d = 2$, the quantum field theory becomes more serious and rigour, for the most part, has to be left behind. This is even more so with the case $d = 3$ of the real world.

The increasing complexity as D increases is reflected by (and perhaps due to) the increasing complexity of the Riemannian curvature. Thus for $D = 1$ there is no curvature, for $D = 2$ we have only the scalar curvature, and for $D = 3$ we have the Ricci curvature, while only for $D = 4$ do we have the full Riemann curvature tensor. At the level of the Einstein equations, for classical relativity, this is related to the increasing difficulty of geometric structures in dimensions $D \leq 4$. For $D = 2$, we have the classical theory of Riemann surfaces (or surfaces of constant curvature). For $D = 3$ the theory of 3-manifolds is already much deeper as is made clear by the work of Thurston (and more recently Perelman). For $D = 4$ the situation is vastly different, as has been shown by Donaldson, using ideas coming from physics as we shall discuss later.

In the subsequent sections, we shall review the ways in which physics has impacted on mathematics, organizing it according to this dimensional hierarchy. However, before proceeding, we should make one general remark which applies throughout. Typically, in the applications, there are formulae which depend on some integer parameter such as a degree. From the physics point of view, what naturally emerges is something like a generating function involving a sum over all values of the parameter. Traditionally this is not the way geometers would have looked for the answer, and one of the remarkable insights arising from the physics is that the generating functions are very natural objects, sometimes being solutions of differential equations.

4 Space–time dimension 2

4.1 Rigidity theorems

The space V of solutions of an elliptic differential operator on a compact manifold M is finite dimensional. If a compact group G acts on the manifold, preserving the operator, then V becomes a representation of G. In fact, all representations of compact Lie groups arise in this way. In very special circumstances we may have *rigidity*, meaning that the representation on V is trivial. For example, if V is the space of harmonic forms, of degree p, and if G is connected, then by Hodge theory the action is trivial. A different example arises if we take the Dirac operator D and spin manifold M (of dimension $4k$), then D and its adjoint D^* have solution spaces V^+, V^- and the index of D is defined as

$$\text{index } D = \dim V^+ - \dim V^-.$$

If G acts on M preserving the metric, then it commutes with D and so V^+, V^- become representations of G and the index becomes a virtual representation or character. The rigidity theorem (of Atiyah and Hirzebruch) says that this character is trivial, i.e., a constant. (In fact, more is true—it is zero for a nontrivial action.)

Arguments from quantum field theory (for maps of space–time into M) led to the discovery of a whole sequence of rigidity theorems for the Dirac operator coupled to certain bundles. Moreover, the generating function turns out to be a modular form, something predicted by the relativistic invariance of the quantum field theory.

This discovery stimulated a whole new branch of topology, called “elliptic cohomology” with fascinating connections to number theory as explained by Michael Hopkins [6].

This subject is an application of physics to differential topology, but the remaining subjects of this section will be concerned with algebraic geometry.

4.2 Moduli spaces of bundles

The Jacobian of an algebraic curve classifies all holomorphic line bundles over it with degree zero. It can also be described as the moduli space of flat $U(1)$-bundles. Its study was a major feature of 19th century mathematics in the context of theta functions. It has a natural generalization to vector bundles of higher rank, the study of which emerged in the middle of the 20th century and is much more involved. In particular, not much was even known about the topology of these moduli spaces.

Again quantum field theory in two dimensions has led to beautiful formulae relating to the cohomology of these moduli spaces. Rigorous mathematical proofs of these results are now available, inspired by the physics.

4.3 Moduli spaces of curves

Somewhat analogous, but deeper than the moduli spaces of Section 4.2, is the moduli space of curves of genus g. The classical theory of the elliptic modular function deals with the case $g = 1$, but for higher genus the moduli space remained rather unknown.

As in Section 4.1, physics has again produced remarkable formulae for the cohomology of these moduli spaces. This time the physics is related to gravity rather than gauge theory and is important for string theory.

4.4 Quantum cohomology

Classical geometry led to many enumerative problems, the simplest of which was to count how many points were common to a number of subvarieties of a given algebraic variety. This led to *intersection theory*, which, in the hands of Lefschetz, was developed as an aspect of homology theory. Subsequently, this was viewed as the ring structure of cohomology theory.

A deeper class of enumerative problems arises when we want to count not *points* but *curves*. How many curves of given type (degree, genus, singularity structure) lie on a given algebraic variety. This was a difficult unsolved problem even for curves in the plane.

In quantum field theory, holomorphic curves appear as "instantons" of a two-dimensional field theory, measuring important nonperturbative features of the theory.

Taking into account instantons of genus zero leads in particular to a ring associated with the target manifold which depends on a parameter t and in the classical limit $t \to 0$ reduces to the cohomology ring. This new ring is called the quantum cohomology ring and it encodes information about numbers of rational curves.

The quantum cohomology rings of various varieties have been calculated, thus leading to explicit enumerative formulae. An important point to mention is that the quantum cohomology ring only has a grading into odd and even parts, not an integer grading like the classical cohomology.

4.5 Mirror symmetry

This subject, which has now grown into a large industry, is related to Section 4.4 and is one of the ways in which the enumerative problems have been solved.

It was discovered by physicists that certain algebraic varieties come in pairs M and M^*, called mirror pairs. The most interesting case is when M and M^* are three-dimensional complex algebraic varieties with vanishing first Chern class (Calabi–Yau manifolds). The remarkable thing about mirror symmetry is that M and M^* have quite different topologies. In fact, the ranks of the odd and even Betti numbers switch, so that

$$\chi(M^*) = -\chi(M),$$

where χ is the Euler number.

For physicists, M and M^* give rise to the same two-dimensional quantum field theory, but quantum invariants involving instanton calculations on M can be calculated by classical invariants involving periods of integrals on M^*. This is what leads to effectively computable formulae and gives the theory its power.

The mathematical study of Mirror Symmetry has now progressed quite far involving symplectic as well as complex geometry. Recent work is formulated in the language of derived categories and surprisingly such extremely abstract mathematical techniques appear in return to be relevant to the physics of string theory.

5 Three-dimensional space–time

The most striking application of quantum field theory in three dimensions was undoubtedly Witten's interpretation of the polynomial knot invariants discovered by Vaughan Jones. It was already clear, from the work of Jones, that his invariants were essentially new and very powerful. Old conjectures were quickly disposed of. What Witten did was to show how the Jones invariants could be easily understood (and generalized) in terms of the quantum field theory defined by the *Chern–Simons Lagrangian*. One immediate benefit of this was that it worked for any oriented 3-manifold, not just S^3. In particular, taking the empty knot one obtained numerical invariants for compact 3-manifolds.

These developments have stimulated a great deal of work by geometers. In particular, there are combinatorial treatments which are fully rigorous and mimic much of the physics.

In three dimensions, we are in the odd situation of having two completely different theories. One the one hand there are the quantum invariants just discussed, while on the other hand there is the deep work of Thurston on geometric structures, including the important special case of hyperbolic 3-manifolds. It has been a long-standing and embarrassing situation that there was little or no connection between these two theories. For example, given an explicit compact hyperbolic 3-manifold, how does one compute its quantum invariants? Some answers were available for the simpler structures (positive curvature or fibrations) but not for the hyperbolic case.

Recently conjectures have been made proposing a general link, with hyperbolic volumes appearing as limits of Jones invariants. In particular Gukov [5] has attempted to establish a basis for this link using Chern–Simons theory for the noncompact groups $SL(2, \mathbb{C})$ which ties in to three-dimensional gravity. This follows earlier work of Witten and others. It looks very promising and one might hope to connect it ultimately to the recent work of Perelman [9].

Perhaps, looking at current research and peering into the future, I can make a few further comments.

In the first place, while quantum field theory gives (at least heuristically) a very satisfying explanation of the Jones theory and most of its properties, it fails in one important respect. It does not explain why the coefficients of the Jones polynomials are *integers*. In Witten's description, the values of the Jones polynomials at certain roots of unity are expectation values and the physics gives no indication of their arithmetic nature.

A really fundamental treatment should provide such an explanation, while preserving the elegance of the quantum field theory approach.

After my lecture, I was reminded that recent work of Khovanov [7] does give a direct explanation for the integer coefficients in the Jones polynomial. Khovanov constructs, from a knot, certain homology groups as invariants and the Jones coefficients appear as Euler characteristics. While this explains their integrality it does not explain what relations these Khovanov homology groups have to the physics.

There is a somewhat parallel situation with respect to the Casson invariant of a homology 3-space. On the one hand Witten has shown that it is given by a variant of Chern–Simons theory. On the other hand it can also be interpreted as the Euler characteristic of the Floer cohomology groups, which are the Hilbert spaces of the Donaldson quantum field theory in four dimensions (as we shall discuss in the next section). This might suggest that the Khovanov homology groups should simply be interpreted as the Hilbert spaces of some four-dimensional quantum field theory. No such theory appears at present to be known.

Speculating in another direction, I note that the Jones polynomial is naturally a character of the circle, the integers being the multiplicities of the irreducible representations. One may ask where the circle comes from. Now the knots studied by Jones are traditional ones in $\mathbb{R}^3$ and we have an S^2 at ∞, on which $SO(3)$ acts. Moreover, the equivariant K-theory of S^2 is given by the character ring of the circle

$$K_{SO(3)}(S^2) \cong R(S^1).$$

Here S^1 appears as the isotropy group of the action (and is unique up to conjugation). Since K-theory appears to play a special role in quantum field theory it is tempting to interpret the Jones polynomial as an element of the equivariant K-theory of S^2, where we think of S^2 as any large 2-space enclosing the knot. This idea receives some encouragement from the fact that the Jones polynomial for links (generalizing knots) involves integer series in $t^{\frac{1}{2}}$, which corresponds to characters of the double cover of our original S^1. But this is natural if we replace $SO(3)$ by $\mathrm{Spin}(3)$, as a physicist would do.

Such equivariant K-groups have appeared in connection not with knots in $\mathbb{R}^3$ but in connection with finite configurations of distinct points in $\mathbb{R}^3$ [2] and this might provide some link. This idea is reinforced by the further speculation made in [3] relating to Hecke algebras, which provide the original Jones approach to knot invariants.

As explained in [2] the 2-sphere involved there is naturally the complex 2-sphere, which occurs as the base of the light-cone in Minkowski space. The connection with quantum theory that is postulated is in the spirit of the ideas of Roger Penrose as mentioned later in Section 9.

6 Four-dimensional space–time

I have already alluded several times to Donaldson theory, on which I shall now elaborate.

For any compact oriented 4-manifold X, any compact Lie group G and any positive integer k, Donaldson studies the moduli space M of k-instantons. These are anti-self-dual connections for the G-bundle (with topology fixed by k). For this he has first to choose a Riemannian metric (or rather a conformal structure), but he then computes some intersection numbers on M and shows these are independent of the metric. In this way, Donaldson defines invariants of X, which are just polynomials on the second homology of X.

As is now well known, these Donaldson invariants proved spectacularly successful in distinguishing between 4-manifolds and they opened up the whole subject of smooth 4-manifolds just as Freedman had closed the subject of topological 4-manifolds.

While the idea of using instantons came from physics, Donaldson was just using the classical equations of Yang–Mills theory. But Witten subsequently explained that Donaldson's theory could be interpreted as a suitable quantum field theory in four dimensions. Moreover, this was just a slight variant on a standard theory known as $N = 2$ supersymmetric Yang–Mills.

This physical interpretation of the Donaldson theory was interesting for physicists but it was not clear what the mathematical benefit was. However, a few years later, the benefit became abundantly clear. As part of some very general ideas of duality in quantum fields theories, Seiberg and Witten produced a quite different theory which was expected to be equivalent to Donaldson theory. This has now been essentially confirmed by mathematicians, though a rigorous proof of the equivalence is not yet complete. Moreover, the Seiberg–Witten equations are technically easier to handle and so they have proved more powerful in many cases. In particular, they have led to a proof of the old conjecture of René Thom about the genus of surfaces embedded in $\mathbb{CP}_2$.

I should emphasize that the equivalence between the Donaldson and Seiberg–Witten theories is one between generating functions. Each theory has its instantons, but there is no simple relation between instantons of separate degrees, only between the total sums over all degrees. This should be compared with the classical Poisson summation formula which expresses a sum over one lattice in terms of the sum of Fourier transforms over the dual lattice. Thus these dualities of quantum field theories should be viewed as some kind of nonlinear analogues of the Fourier transfom. I shall return to this theme at the end of my lecture.

One surprising feature of the Seiberg–Witten theory, and the classical equations they lead to, is that they deal with a $U(1)$ theory coupled nonlinearly to spinors. Thus spinors appear explicitly here, while they do not appear in the $SU(2)$ Donaldson theory. This only increases the mystery of spinors and emphasizes my earlier remarks about our lack of any deep understanding of them.

At present it is not clear whether Donaldson theory, with various refinements, will explain all geometric phenomena in four dimensions. It may do so, but it is also possible that it may take another 100 years to fully understand the geometry of four dimensions, just as it has taken a century to move from Riemann surfaces to an equivalent understanding of three dimensions. If so, this may accompany a similar period for a proper understanding of the physics of space–time, a topic to which I will return in the last section.

7 Topological quantum theories

As I have quickly outlined, there are a large number of important areas where quantum theories yield topological results. All of these are, in fact, topological field theories. They are especially simple theories in which the only output is topological. The Hamiltonian of such a theory is zero, so there is no continuous dynamics. However, the theory has nontrivial content related to topological phenomena. This makes this area much simpler and hence more tractable mathematically. In [1] I gave an axiomatic description of a topological quantum field theory, analogous to the classical axiomatization of homology by Eilenberg and Steenrod. The key part of such a theory is its construction by some explicit method which could in principle (as with homology) be either combinatorial or analytic.

It might appear that, for a real physicist, such purely topological theories could have no serious interest. But this is wrong for two reasons. In the first place the complexity of a really physical quantum field theory can rise from one of two sources. First there is the analytical study of small fluctuations, but this is to a great extent based on standard examples and perturbation theory. Then there are nonperturbative phenomena and these are illustrated very well by purely topological theories.

But not only can topological theories play the role of toy models to study nonperturbative effects, they can also arise from a physical theory in some limiting regime. As a simple illustration consider Hodge theory, or supersymmetric quantum mechanics. The full theory requires us to know all the eigenvalues of the Hodge Laplacian. But under rescaling we can consider the limit when all eigenvalues get very large, so that only the zero eigenvalues survive. This recovers the homology (as the harmonic forms).

Although I have spoken only about quantum field theories, the connection between geometry and physics also extends to string theories. In particular, Witten has shown that Chern–Simons theory for $U(N)$, in its perturbative form, is a topological string theory for open strings on T^*S^3, the cotangent bundle of S^3, with the 0-section as a brane of multiplicity N where the string must end.

More recently Vafa and others have argued that this theory is dual to the (topological) theory of closed strings on a rank 2 vector bundle over $\mathbb{CP}_1$. In this duality, one switches from a perturbative expansion valid for *large level*, of Chern–Simons $U(N)$ gauge theory, to a perturbative expansion of the closed string theory for *large* N. The geometry behind this duality is best understood in M-theory terms involving a suitable 7-manifold of G_2-holonomy [4].

This duality of Vafa leads to explicit formulae for every genus of the world-sheet and these have now been verified by mathematical computations using fixed-point methods on moduli spaces [8]. Interestingly, in the end, everything boils down to purely combinatorial formulae. On the one hand string theory arises from the Feynman diagrams of perturbation theory, while on the other hand we have combinatorial data associated with degenerate algebraic curves. Riemann surfaces and analysis interpolate between these two different combinatorial schemes, but with sufficient skill a direct computation can be made. However, this is not very enlightening.

In his lecture at this conference, Vafa [10] discusses many aspects of topological theories in much greater detail. I refer to his text for more information.

8 The significance for mathematics

It should be clear from my rapid survey that quantum theory, in its modern form, has had profound consequences for mathematics and in particular for geometry. But it is hard to grasp the real significance of all this and to predict what its future will be.

While physics can inject new ideas and techniques into mathematics, it cannot in the end provide a foundation for it. Even if, one day, we can develop a completely rigorous quantum field theory or string theory, it would be bizarre if this had to be the pillar on which mathematics, or large parts of it, rested.

A historical perspective may help us get a glimpse of the future in this respect. Fourier analysis emerged, in the 18th century, from physics, specifically the study of heat conduction. But in due course it was absorbed into a purely mathematical theory, and was fundamental in the subsequent development of Linear Analysis. Later, in the 20th century, this theory was generalized to the noncommutative situation centering around group representation theory. In fact, one could say that noncommutative Fourier analysis has been one of the central theories of 20th century mathematics.

As I have mentioned earlier, the dualities of quantum field theory and string theory which lie behind some of the most striking applications of physics to geometry can be viewed as some kind of nonlinear Fourier Transform. In special finite-dimensional cases these are now understood mathematically, and are related to classical ideas of integral geometry. These include the Penrose Transform, the Mukai Transform, the Nahm Transform, and the inverse scattering transform in soliton theory. In fact, solitons are a prominent part of all these dualities. However, the full dualities of string theory (or QFT) are infinite dimensional and nonlinear.

All this suggests that a prominent theme of 21st century mathematics might be the development of a fully-fledged nonlinear Fourier Transform theory for function spaces. Of course, there are too many kinds of nonlinearity to be encompassed in any nontrivial way by a single theory. Clearly, physics appears to be singling out a type of nonlinearity for which a deep but tractable duality will hold. The key feature of this nonlinearity appears to be supersymmetry, which in some way extends the symmetry arising in group theory. In geometric terms this means that we deal not just with homogeneous spaces of Lie groups but also with Riemannian manifolds having special holonomy, such as Kähler manifolds, Calabi–Yau manifolds or G_2-manifolds. The Lie groups are still there, but only at the (integrable) infinitesimal level. These ideas are, in fact, not far removed from the original ideas of Lie, who moved on from finite-dimensional Lie groups to the infinite-dimensional structures occurring for example in complex manifolds. Lie himself was disappointed that his fundamental ideas did not appear to be given their due credit. The 20th century certainly rectified this omission, but perhaps the 21st century will take it even further.

9 The significance for physics

In the previous section, I tried to peer into the future to see what kind of mathematics might emerge from the current geometry–physics interface. Trying to forecast the physics is even harder, and I am less qualified, but perhaps an outsider can offer a different perspective.

As we know, the holy grail in current fundamental physics is how to combine Einstein's Theory of General Relativity with Quantum Theory. These two theories operate very effectively but at quite different scales, GR at cosmic distances and QM at subatomic scales.

The difficulty in combining the two theories is both conceptual and technical. As is well known, Einstein dreamed of a unified geometric theory, extending GR, and he never accepted the philosophical foundations of QM, with its uncertainty principle. In the long debate on this controversy between Einstein and Bohr the general verdict of the physics community was that Einstein lost and that his idea of a unified field theory was a hopeless pipe dream.

With the remarkable success of the standard model of elementary particles, incorporating geometrically the electromagnetic, the weak force and the strong force, Einstein's ideas were given new life. But the framework remained that of QM, and GR remained strictly outside the scope of the unification. Now, with string theory offering the hope of the ultimate unification it might appear that the old controversy between Einstein and Bohr has been resolved, with the honours more equally split. Unification is perhaps being achieved, but QM has persisted.

This is the orthodox view of string theorists and they have impressive evidence in their favor. The only fly in the ointment is that no one yet has any real idea of what their ultimate M-theory is. Perhaps in the coming years this will be clarified and we will learn to live with the mysterious world of 11 dimensions and its hidden supersymmetries. Perhaps only a few technical obstacles remain to complete the structure.

But it is at least worth exploring alternative scenarios. There are in particular two attractive ideas that have their devotees. The first (in historical precedence) is Roger Penrose's twistor theory. On the one hand this has, as a technical mathematical tool, proved its worth in a number of problems. It is also related to supersymmetry and duality. Links with string theory are being explored. But beyond these mathematical technicalities there lies a deeper philosophical idea. Penrose is an Einsteinian who believes that in the hoped-for marriage between GR and QM it is the latter that must give the most, adapting itself to the beauty of GR. Twistors are thought of as a first step to achieving this goal. Moreover, Penrose speculates that the mysterious role of complex numbers in QM should ultimately have a geometric origin in the natural complex structure of the base of the light-cone in Minkowski space. So far, it has to be conceded that the weight of evidence is not in Penrose's favor, but that does not mean that he may not ultimately be vindicated.

A completely different scenario is offered by Alain Connes' noncommutative geometry, a theory with a rich mathematical background and a promising future. Links with physics exist and new ones are being discovered. In a sense Connes takes off from the Heisenberg commutation relations, in a definitely non-Einsteinian direction. However, he tries to keep the geometric spirit by using the same concepts and terminology. It is certainly possible that the final version of M-theory may use Connes' framework for its formulation.

Perhaps I can end by indulging in some wild speculation of my own, not I hope totally unrelated to the other ideas above.

I start, further back, by asking some philosophical or metaphysical questions. If we end up with a coherent and consistent unified theory of the universe, involving extremely complicated mathematics, do we believe that this represents "reality"? Do we believe that the laws of nature are laid down using the elaborate algebraic machinery that is now emerging in string theory? Or is it possible that nature's laws are much deeper, simple yet subtle, and that the mathematical description we use is simply the best we can do with the tools we have? In other words, perhaps we have not yet found the right language or framework to see the ultimate simplicity of nature.

To get a better idea of what I am trying to say, let us consider GR as a description of gravity. To a mathematician this theory is beautifully simple but yet subtle. Moreover, it is highly nonlinear so that it is extremely complicated in its detailed implications. This is no doubt why it appeals to both Einstein and Penrose as a model theory. Is it not possible that something having the same inherent simplicity (and nonlinearity) can explain all of nature?

While everyone might agree that this would be an ideal philosophical ambition, there appears to be the insuperable obstacle presented by QM. To get round this will require some conceptual leap, and such leaps have in the past only come when one is prepared to sacrifice some accepted dogma, such as Einstein did with the separation of space and time.

Let me, in such a speculative mood, raise one possibility of a dogma to be sacrificed. Ever since Newton, it has been a cardinal principle of physical sciences that we can predict the future from the present (given complete knowledge). This even holds in QM, where the state at time zero evolves by a Hamiltonian flow to give the state at future times. This assumption, which may have seemed rash to some, has abundantly proved its worth. But is it really true? Perhaps all we can say is that a knowledge of present *and past* enables us to predict the future? After all, this, in a sense, is true in the biological world where our DNA represents our past.

Of course, to explain the remarkable success of the standard dogma, the effect of the past would have to be minute and only noticeable at very short time-scales. But this is precisely where QM comes into play. So perhaps the uncertainty in QM is really a reflection of the fact that we (the observers) do not know our past. Perhaps the Hilbert Space state at the present time is determined by our past.

This metaphysical idea would have to be embodied in precise mathematical form consistent with GR. In particular, the fundamental equations would not be differential equations but integrodifferential equations, involving integration over the past. The nonlinearity of GR, together with the effect of past history, would be difficult to solve mathematically. But very good approximations might be obtained by using high precision mathematical tools of the kind appearing in string theory. The various dualities might appear from alternative ways of making the necessary approximations.

A theory on these lines would have satisfied Einstein and it seems at least worth exploring. The dream of all mathematical physicists is to find ultimate explanations which are inherently simple in mathematical form and yet can explain the fascinating diversity of nature. We should not settle for less.

References

[1] M. F. Atiyah, *Topological Quantum Field Theories*, Publications Mathématiques 68, Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France, 1989, 175–186.

[2] M. F. Atiyah, Configurations of points, *Philos. Trans. Roy. Soc. London Ser.* A, **359** (2001), 1375–1387.

[3] M. F. Atiyah and R. Bielawski, Nahm's equations, configuration spaces and flag manifolds, *Bull. Brazil Math. Soc. N. S.*, **33** (2002), 157–176.

[4] M. F. Atiyah, J. Maldacena, and C. Vafa, An M-theory flop as a large $\mathbb{N}$ duality, *J. Math. Phys.*, **6** (2002), 3209–3220.

[5] S. Gukov, Three-dimensional quantum gravity Chern-Simons theory and the A-polynomial, 2003, arXiv:hep-th/0306165.

[6] M. J. Hopkins, Algebraic topology and modular forms, in *Proceedings of the ICM, Beijing* 2002, Vol. 1, Higher Education Press, Beijing, 2002, 283–309; also in *Comm. Math. Phys.*, **255**-3 (2005), 577–627.

[7] M. Khovanov. A categorification of the Jones polynomial, *Duke Math. J.*, **101**-3 (2000), 359–426.

[8] K. Liu, Mathematical results inspired by physics, in L. I. Tatsien, ed., *Proceedings of the International Congress of Mathematicians, Beijing* 2002, Vol. 3, Higher Education Press, Beijing, 2003, 457–466.

[9] G. Perelman, The entropy formula for the Ricci flow and its geometric applications, 2002, arXiv:math.DG/0211159.

[10] C. Vafa, Unity of topological field theories, lecture given at *An International Conference on* "*The Unity of Mathematics*," Harvard University, Cambridge, MA, 2003.