Michael Atiyah: Geometry and Physics

Nigel Hitchin
Mathematical Institute, Woodstock Road, Oxford, OX2 6GG
hitchin@maths.ox.ac.uk

December 31, 2020

1 Introduction

Over the years, since I was a student, I experienced the many ways in which Michael Atiyah brought mathematics to life by importing new ideas, reformulating them in his own way, communicating them in his inimitable style and using these fresh insights to advance his own research. While in Oxford he would take every opportunity to travel to the USA and listen to what was currently interesting to him, then explain the results with his added contributions at his regular Monday seminar, hoping that some of these themes would be taken up by his audience.

He was at heart a geometer, but in the mid 1970s he became convinced that theoretical physics was by far the most promising source of new ideas. From that point on he became a facilitator of interactions between mathematicians and physicists, attacking mathematical challenges posed by physicists, using physical ideas to prove pure mathematical results, and feeding the physicist community with the parts of modern mathematics he regarded as important but were unfamiliar to them.

Here is what he wrote in a retrospective view [24], producing two lists of the historical influences in both directions. It gives an impression of the weighting he applied to areas of the two disciplines.

Mathematics ahead of physics: curved space, Lie groups, higher dimensions, fibre bundles, spinors, exterior algebra, non-commutative algebra, Hilbert space, special holonomy.

Physics ahead of mathematics: infinite-dimensional representations, Maxwell theory, Dirac theory, supersymmetry, quantum cohomology, conformal field theory, quantum field theory in 3 and 4 dimensions.

He concludes the article by attacking the conventional view of the relationship – the market force model where we mathematicians alter our production in the light of the changing needs of physics. Instead he sees mathematicians as gardeners breeding new species, and physicists as the modern version of the 19th century collectors who searched the world for exotic specimens to reinvigorate our gardens. For Michael Atiyah, I hope to show that both models are valid.

Although Michael had been exposed to physics lectures as a student in Cambridge it was only to expand his general knowledge – he also attended lectures in architecture and archaeology. His peer Roger Penrose, also studying algebraic geometry at the time, took little interest in physics then although he would famously go on to win a Nobel Prize in Physics! Their subsequent paths diverged and Atiyah's contributions to topology and geometry had already won him a Fields Medal in 1966 before having any substantial contact with physicists.

In what follows I want to discuss some of the examples of the interaction which is apparent in his work, in roughly chronological order, giving on the way an outline of the mathematical issues for a general audience. Michael Atiyah was a bridge-builder, and I come from one side of that bridge, so others may interpret things differently.

First here is a timeline of his mathematical development before he succumbed to the influence of physics:

- 1929 Born 22 April
- 1949 -1955 BA and PhD Cambridge
- 1952 - 1958 Algebraic geometry: this period saw him introducing the new approaches coming from France. His work included results on ruled surfaces expressed in terms of vector bundles on algebraic curves, a classification of bundles on elliptic curves, and a new way of looking at characteristic classes in holomorphic terms.
- 1959 - 1974 K-theory: during this time he became effectively a topologist, collaborating with Hirzebruch in constructing K-theory, a cohomology theory of vector bundles which provided the framework to resolve efficiently many outstanding problems in algebraic topology
- 1963 - 1975 Atiyah-Singer index theorem: here collaborations with Singer and also Bott drew on both topology and algebraic geometry together with analysis and differential geometry and is discussed in the next section.

- 1966 Fields Medal

2 The index theorem

A precursor to the interaction took place at the time of a later phase in the development of the Atiyah-Singer *index theorem* around 1971. The theorem is Atiyah's most notable achievement and, together with Singer, he received the Abel Prize for this work in 2004. The index theorem is a huge generalization of the classical Riemann-Roch theorem of the 19th century. In modern terminology, the motivation for that theorem was to seek an estimate for the dimension of the vector space of holomorphic sections of a line bundle L over a Riemann surface, or algebraic curve C, in terms of its degree. If the degree d is large enough the answer is $d+1-g$ where g is the genus of C. In general, the dimension may change as L varies, fixing the degree, and the Riemann-Roch formula is

$$\dim H^0(C,L) - \dim H^1(C,L) = d+1-g$$

where $H^0(C,L)$ is the space of sections and $H^1(C,L)$ the first cohomology group of the sheaf of sections. The latter is more difficult to conceptualize, a fact which became important later on in Atiyah's interaction with Penrose, but at this point the formula can be seen as the statement that the difference of two dimensions is not only invariant under continuous deformation of the line bundle and the curve, but there is an explicit formula. The right hand side involves topological invariants – the degree d, or Chern class, of the line bundle and the genus g, the basic invariant of the surface.

There is a more analytical approach. If s is a C^∞ section of L then on a local open set U_α it is defined by a function f_α. The section s is holomorphic if f_α satisfies the Cauchy-Riemann equation $\partial f_\alpha/\partial\bar{z}{=}0$. Another choice of representative function f_β on U_β is equal to $g_{\alpha\beta}f_\alpha$ where $g_{\alpha\beta}$ is holomorphic. Then $\partial f_\beta/\partial\bar{z} = g_{\alpha\beta}\partial f_\alpha/\partial\bar{z}$ which means that $(\partial f_\alpha/\partial\bar{z})d\bar{z}$ is a well defined section of the tensor product $L\otimes\bar{K}$ where $\bar{K}$ is the line bundle of differential forms expressed locally as $a(z,\bar{z})d\bar{z}$. We then have a differential operator

$$\bar{\partial} : C^\infty(L) \to C^\infty(L\otimes\bar{K})$$

and in this context the classical Riemann-Roch formula is

$$\dim\ker\bar{\partial} - \dim\operatorname{coker}\bar{\partial} = d+1-g.$$

The index theorem replaces this by the general case of a compact manifold M with two vector bundles V_+, V_- and an elliptic operator D transforming sections of V_+ to sections of V_- and then

$$\dim\ker D - \dim\operatorname{coker} D = \operatorname{ind} D$$

where the index $\operatorname{ind} D$ is expressed as a specific polynomial in characteristic cohomology classes related to the tangent bundle of M and the vector bundles V_+, V_-, evaluated on the fundamental class of M.

A key moment in the formulation of the theorem, even before its proof, was a visit in 1962 of Singer to Oxford where he offered the Dirac operator as the candidate for using the index theorem to explain the integrality of the particular combination of characteristic classes given by Hirzebruch's $\hat{A}$-polynomial. Appropriating the Dirac operator could have been the first contact with physics but, as Atiyah explains in [23]:

> Several decades later, and with all the interaction now taking place between physicists and geometers, it may seem incredible that we did not get there earlier and more directly. There are several explanations. First, at that time physics and mathematics had grown rather far apart in these areas. Second, physics dealt with Minkowski space and not with Riemannian manifolds, so any relation we might have noticed would have seemed purely formal. In fact as a student of Hodge I should have known better, since Hodge's theory of harmonic forms, while finding its main application in algebraic geometry, was explicitly based on analogy and extension of Maxwell's theory. Even more surprising is that Hodge and Dirac were both professors in the mathematics department at Cambridge at the same time and knew each other well, and yet it never occurred to Hodge to use the Dirac operator in geometry. Part of the difficulty lies of course in the mysterious nature of spinors. Unlike differential forms they have no easy geometrical interpretation. Even now, at the end of the century, and with some spectacular progress involving spinors and the Seiberg-Witten equations, we are still in the dark in some fundamental sense. What, geometrically is a spinor and why are they important?

The last point is probably the most relevant. If you reads Hodge's book, written in the 1930s, you will see that even defining a manifold and differential forms takes up a lot of space. To imagine that spinors could be incorporated into that picture is asking a lot.

The 1960s saw the index theorem develop and multiply in many ways. The first proof was largely topological using cobordism theory. The second proof was motivated more by algebraic geometry and Grothendieck's approach to the Riemann-Roch theorem.

Finally in 1971 came a new method from V.K.Patodi using the heat equation and this had more resonance with what physicists were doing. It was analytical and involved the spectrum rather then just the nullspace. This was a fundamental change of direction – algebraic geometry, the initial inspiration for the theorem, only encounters the null spaces of operators but physicists are always more interested in the eigenvalues. Singer's physicist colleagues at MIT were now able to communicate with him having a better common language and this was a new opening between the two fields.

Patodi, building on the work of Singer and H.McKean, used the heat kernels for the second order operators DD^* and D^*D where D^* is the formal adjoint. The cokernel of D is isomorphic to the kernel of D^* and since D^*D is non-negative it has the same nullspace as D. Hence the index can be written

$$\operatorname{ind} D = \dim \ker D^*D - \dim \ker DD^*.$$

The eigenspaces of DD^* and D^*D for nonzero eigenvalues are interchanged by D so they have the same non-zero spectrum. Then the index is formally the difference

$$\operatorname{tr} \exp(-tD^*D) - \operatorname{tr} \exp(-tDD^*)$$

but there is an asymptotic expression as t tends to zero which involves integrals of locally-defined expressions, and when appropriately identified these turn out to be differential forms defining the characteristic classes for the right hand side of the index theorem.

There is then a symmetry between the nonzero part of the spectra of the two operators, but a lack of symmetry for the zero eigenvalue. On an even-dimensional manifold there are two bundles of spinors, V_+, V_- which in physical terms are the spinors of positive and negative *chirality*, the sign dependent on the orientation. The Dirac operator interchanges them and the lack of symmetry for the zero eigenvalue is called the chiral *anomaly* – equivalently the index. An anomaly is a lack of invariance under a natural transformation, in this case change of orientation or parity.

Even if the index is zero there may not be a natural symmetry between the two spaces, a feature which becomes relevant when the operators belong to a family, for example when the Dirac operator is coupled to a connection, meaning extending D to act on $V_+ \otimes E$ where E has a connection, or gauge potential, and the parameter space is a family of connections modulo gauge equivalence as in [10].

This situation was in a way foreseen in the adaptation of the bare index theorem stated above to the index theorem for families which was dealt with in one of the later papers in the series of proofs [4]. Here the index is the difference of two vector spaces, rather than two dimensions, varying over the parameter space. Though the dimensions may jump, there is nevertheless a well-defined class in the K-theory of the parameter space, K-theory being specifically created by Atiyah and Hirzebruch to cater for this scenario. If the original model for the index theorem for families was Grothendieck's Riemann-Roch theorem concerning the higher direct images of a sheaf under a proper map of algebraic varieties, the analysis required to apply this to elliptic operators became appropriate to the more general discussion of anomalies later.

Another consequence of Patodi's approach to proving the index theorem was the introduction of the eta-invariant. This was another measure of the asymmetry of the spectrum and also related to anomalies. The eta-invariant concerns the Dirac and related operators on an odd-dimensional manifold. Here there is only one spinor bundle V and D is self-adjoint and has a real spectrum, discrete but extending from $-\infty$ to $+\infty$. The lack of symmetry to be measured here is between the positive and negative eigenvalues and the device which can detect it – the *eta-invariant* –is the analytic continuation to $s = 0$ of

$$\eta(s) = \frac{1}{2}\left(\sum_{\lambda\neq 0} \frac{\operatorname{sgn}\lambda}{|\lambda|^s} + \dim\ker D\right).$$

The original motivation was to develop an index theorem for even-dimensional manifolds with boundary, for example to compare the index for square integrable solutions to the polynomial in characteristic classes which gave the index in the compact situation. The difference is a function on the boundary with one local term involving curvature (the second fundamental form) and the other the global eta-invariant. It depends on the metric so is not a topological invariant but its derivative in a family is the integral of a local curvature term.

The eta-invariant appears as a global anomaly for a diffeomorphism $f : X \to X$ of a manifold. This entails forming the mapping torus M, by taking $X \times [0, 1]$ and identifying the ends. The global anomaly is then defined to be $e^{2\pi i\eta}$ and appears for the physicists as the phase of a partition function in the functional integral.

Although these developments appeared somewhat later than the 1970s they provide enough evidence that the driving force in index theory was moving away from algebraic geometry and was more in tune with the requirements of theoretical physics, even if the authors were not aware of it.

In 1983, the tables were turned when the physicist Alvarez-Gaumé produced a new proof of the index theorem based on supersymmetry [3]. It follows the Patodi approach but in a new formalism. For example, the symmetry of non-zero eigenvalues becomes

> The properties of supersymmetry however, imply that this supersymmetric index depends only on the zero energy states due to the fact that all non-zero energy states appear in bose–fermi pairs.

More importantly it approaches in a systematic way the polynomials in characteristic classes which give the index. For Patodi himself there were "miraculous cancellations" which Atiyah and Singer replaced by using invariant theory to determine the individual terms to be assembled into the final formula. Here it arises by

> computing the standard partition function at temperature β^{-1} for a free gas of fermions moving in one dimension, with masses determined by the eigenvalues of the curvature.

In Getzler's more mathematical treatment [36] it is the heat kernel of the one-dimensional harmonic oscillator which picks out the particular polynomial in the curvature to yield Hirzebruch's $\hat{A}$-genus. Explaining why this complicated expression should give an integer was the initial motivation for Atiyah and Singer back in 1962, so here was physics providing a coherent structure yielding a precise known formula.

3 Twistor theory

I was Michael Atiyah's research assistant at the Institute for Advanced Study, where he was a permanent member, from 1971-73. Most of his work at this time revolved around applying the heat equation approach to the index theorem, but at some point in early 1973 we both went over to Princeton University to hear a talk by Roger Penrose on black holes and singularities. After the seminar, the two of them spoke together at some length: I imagined they were reminiscing about their student days in Cambridge, but it turned out they were both planning to take up positions in Oxford – Penrose to the Rouse Ball Chair of Mathematics and Atiyah to a Royal Society Professorship in the Mathematical Institute, so this was more likely the topic of conversation.

On his arrival, Penrose quickly set up a group of young students and postdocs focusing on his notion of twistor theory, which he initiated in 1967 [47]. This regarded compactified, complexified Minkowski space as a 4-dimensional complex projective quadric. Interpreting this as the classical Klein quadric, it parametrizes lines in a complex projective 3-space, which is the (projective) *twistor space*. Introducing complex coordinates was philosophically justified by arguing that quantum mechanics required complex numbers, so any unified theory would have to use them in gravity, the essential point though was that twistors were more fundamental.

It is easier to describe the approach without the compactification, which means removing a projective line from complex projective 3-space $\mathbf{C}\mathrm{P}^3$. The one-parameter planes through this line at infinity give a projection $\mathbf{C}\mathrm{P}^3 \backslash \mathbf{C}\mathrm{P}^1 \to \mathbf{C}\mathrm{P}^1$ which expresses this space as a vector bundle

$$\mathcal{O}(1) \oplus \mathcal{O}(1) \to \mathbf{C}\mathrm{P}^1.$$

Here $\mathcal{O}(1)$ denotes the line bundle of degree one over $\mathbf{C}\mathrm{P}^1$ whose holomorphic sections are linear forms $a\zeta + b$ in a parameter ζ on $\mathbf{C} \subset \mathbf{C} \cup \{\infty\} = \mathbf{C}\mathrm{P}^1$. (Here Riemann-Roch gives $d + 1 - g = 1 + 1 - 0 = 2$). A projective line in $\mathbf{C}\mathrm{P}^3 \backslash \mathbf{C}\mathrm{P}^1$ is now a holomorphic section of $\mathcal{O}(1) \oplus \mathcal{O}(1)$ and so is given by $s(\zeta) = (a\zeta + b, c\zeta + d)$ with $(a, b, c, d) \in \mathbf{C}^4$.

We are supposed to regard this as complexified Minkowski space rather than just a 4-dimensional vector space and the key idea is to define the light cone from each point. A point $x \in \mathbf{C}^4$ corresponds to a line P_x in twistor space and the light cone through x is defined to be the set of lines meeting P_x. So take the zero section $s(\zeta) = (0, 0)$ as P_x then the light cone consists of the sections $(a\zeta + b, c\zeta + d)$ which vanish for some value of ζ. This condition is $ad - bc = 0$. If $(a, b, c, d) = (z + t, x + iy, x - iy, t - z)$ then this reads $t^2 - x^2 - y^2 - z^2 = 0$ the usual Minkowski light cone.

One of the early achievements of twistor theory was a description of solutions of zero rest mass field equations in terms of contour integrals. These equations include the Dirac equation, Maxwell's equation and the wave equation, all conformally invariant equations.

Here is the process as originally formulated by Penrose, working with a holomorphic function f of three variables: [48]:

- f is holomorphic in some domain $\mathcal{D}$
- there is a gauge freedom $\mathcal{G}$ whereby $f \mapsto f + h^- + h^+$ where $h^\pm$ is holomorphic on some extended domain $\mathcal{D}^\pm (\supset \mathcal{D})$ of twistor space in which the contour γ can be deformed to a point (to the "left" in $\mathcal{D}^-$ and to the "right" in $\mathcal{D}^+$)
- $\mathcal{G}$ depends upon the location of γ
- by invoking $\mathcal{G}$, then moving γ, then invoking new $\mathcal{G}$, moving γ we obtain a whole family of equivalent twistor functions all giving the same field.

Although 20 years earlier, Penrose and Atiyah were both graduate students in algebraic geometry, it was Atiyah who absorbed the new approach of sheaf cohomology coming from Paris, while Penrose, a student of Todd rather than Hodge, worked in a more classical vein. So when the two got together in Oxford, it was Atiyah who pointed out that this was just a sheaf cohomology group.

The simplest case is the wave equation. Penrose's original contour description [49] gives

$$\phi(x,y,z,t)=\int_{\gamma\subset\mathbf{C}\mathrm{P}^1} f((z+t)+(x+iy)\zeta,(x-iy)-(z-t)\zeta,\zeta)d\zeta$$

where $f(z_1,z_2,z_3)$ is holomorphic, immediately yielding a solution to

$$\frac{\partial^2\phi}{\partial t^2}-\frac{\partial^2\phi}{\partial x^2}-\frac{\partial^2\phi}{\partial y^2}-\frac{\partial^2\phi}{\partial z^2}=0.$$

This might be zero, for example if the function extends to the interior or the exterior of γ which is the gauge freedom in his description above.

The sheaf-theoretic viewpoint decomposes $\mathbf{C}\mathrm{P}^1=U_0\cup U_1$ as the complement of ∞ and 0, and represents the twistor space $Z=\mathcal{O}(1)\oplus\mathcal{O}(1)$ as the union of $\mathbf{C}^2\times U_0$ and $\mathbf{C}^2\times U_1$, both copies of $\mathbf{C}^3$ identified over $\mathbf{C}^2\times\mathbf{C}^*=\mathbf{C}^2\times U_0\cap U_1$. The function f is defined on this intersection and is a Čech representative of an element a in the sheaf cohomology group $H^1(Z,\mathcal{O}(-2))$. Then each point $x\in\mathbf{C}^4$ defines a twistor line P_x and we restrict a to give an element of $H^1(\mathrm{P}_x,\mathcal{O}(-2))$. This is one-dimensional (Riemann-Roch) and varying the line gives a scalar field ϕ which, as above, satisfies the wave equation. The space $H^1(\mathbf{C}\mathrm{P}^1,\mathcal{O}(-3))$ is 2-dimensional and gives a solution to the Dirac equation with a specific chirality and so forth. The opposite chirality appears in a slightly different format. This is because the twistor space itself depends on orientation – two orientations give dual copies of $\mathbf{C}\mathrm{P}^3$.

Atiyah also pointed out that the Klein correspondence between the quadric and the projective space should be seen as a double fibration: the 5-dimensional space of pairs – lines in $\mathbf{C}\mathrm{P}^3$ and points on them – which fibres over both spaces. Quite soon the twistor research group was using spectral sequences and the full range of techniques in sheaf cohomology to derive further results.

Penrose clearly was concerned with Minkowski signature, the complex conjugation on $\mathbf{C}^4$ defined by $(a,b,c,d)\mapsto(\bar{a},\bar{c},\bar{b},\bar{d})$ has fixed points $\mathbf{R}^4$ with Minkowski signature. This conjugation interchanges the twistor projective space with its dual projective space. Having listened to Penrose explaining twistor space in terms of the classical geometry they were both familiar with, Atiyah looked at it from the point of view of Euclidean signature, the Riemannian situation which featured in most of his work.

He quickly observed that in this case $\mathbf{C}\mathrm{P}^3$ is the projective space $\mathrm{P}(\mathbf{C}^4)$ where $\mathbf{C}^4$ is regarded as a 2-dimensional quaternionic vector space. This gives it an antiholomorphic involution with no fixed points. Furthermore it fibres over the 4-sphere S^4 with fibres projective lines, the lines which are transformed into themselves by the involution acting as the antipodal map on $S^2 \cong \mathbf{C}\mathrm{P}^1$. In fact, as Atiyah pointed out, there was a more general picture. The bundle of complex structures on the tangent spaces of the 4-sphere is the complex manifold $\mathbf{C}\mathrm{P}^3$, the space for S^6 is a complex manifold with fibre $\mathbf{C}\mathrm{P}^3$ and so on. The Euclidean version of twistor space soon became important in another context.

4 Instantons

In 1977 Singer spent a sabbatical term in Oxford and explained to Michael a problem that had been put to him by physicists concerning self-dual solutions to the Yang-Mills equations. These equations described the critical points of the Yang-Mills functional on connections on a principal $SU(2)$-bundle over $\mathbf{R}^4$. The connection A is given locally by a 1-form with values in the Lie algebra $\mathfrak{g}$ defining covariant differentiation

$$\nabla_i = \frac{\partial}{\partial x_i} + A_i$$

and its curvature F_A given by $[\nabla_i, \nabla_j]dx_i \wedge dx_j$ is a globally defined $\mathfrak{g}$-valued 2-form. The Yang-Mills functional is the L^2 norm square of the curvature, which can be written using the Hodge star operator $*$ acting on 2-forms as

$$\int_{\mathbf{R}^4} |F_A|^2 dx = \int_{\mathbf{R}^4} \mathrm{tr}\, F_A \wedge *F_A$$

The original motivation was an attempt to describe quark confinement, which was not in the end realized, but from the mathematician's point of view the Wick rotation from Minkowski space to Euclidean space placed the problem in a familiar location. Since the Hodge star is conformally invariant in the middle dimension it meant that (with suitable justification which came later with Karen Uhlenbeck's work) it became a problem on the conformal compactification of $\mathbf{R}^4$ namely the sphere S^4. Since $*^2 = 1$ the curvature splits into two pieces $F_A^+ + F_A^-$ with $*F_A^+ = F_A^+$ and $*F_A^- = -F_A^-$. Then, using the volume form ν,

$$\int_{S^4} \mathrm{tr}\, F_A^2 = \int_{S^4} (|F_A^+|^2 - |F_A^-|^2)\nu$$

but the left hand side is a topological invariant, a characteristic class, which means the Yang-Mills functional, the integral of $|F_A^+|^2 + |F_A^-|^2$, is bounded below by $4\pi k$ for a nonnegative integer k and the bound is achieved if $F_A = *F_A$ which is *self-dual* or the opposite anti-self-dual $F_A = - * F_A$, the choice depending on orientation.

The challenge then, was to find all self-dual solutions for the simplest non-abelian group $SU(2)$. There are none for an abelian group since the curvature is then a harmonic 2-form and the 4-sphere has zero cohomology in degree 2.

One aspect of the mathematics/physics relationship is that physicists are much better at producing examples to prove a point. The mathematician's response that physicists are satisfied with examples rather than general theorems may well be justified but examples are crucially important for getting things moving. In this case Belavin, Polyakov, Schwartz and Tyupkin had in 1975 exhibited a nonsingular spherically symmetric solution with $k = 1$. This was followed by more examples [39] which yielded solutions depending on $5k + 4$ parameters which can locally be written as

$$A_i = -\frac{1}{2} e_i \cdot d \log \rho \qquad \rho = 1 + \sum_1^k \frac{m_i^2}{|x - x_i|^2}. \tag{1}$$

Here the function ρ is a linear superposition of $1/r^2$ potentials – fundamental solutions of Laplace's equation in four dimensions. The points $x_i \in \mathbf{R}^4$ could be interpreted as locations of particles and indeed these were called "pseudoparticle" solutions. The notation for the A_i here is essentially describing the connection on the spinor bundle of $\mathbf{R}^4$ outside the points x_i. The singularities of ρ at these points mean that a local gauge transformation in a neighbourhood of the points extends the formula to describe a smooth connection on a bundle with topological invariant k.

The first advance during Singer's visit was that Richard Ward, one of Penrose's students, had shown that a complex solution to the self-dual Yang-Mills equations was given by a holomorphic vector bundle on twistor space – a nonlinear version of the correspondence which worked so well for linear equations. By chance Michael had attended the seminar and immediately saw how his Euclidean version of twistor theory would be relevant to the instanton problem – one needed a holomorphic rank 2 bundle on $\mathbf{C}\mathrm{P}^3$ which was holomorphically trivial on each real line. Since these lines are the fibres of the projection $\mathbf{C}\mathrm{P}^3 \to S^4$ this clearly defines a bundle on S^4. The connection is the unique extension of the holomorphic trivialization to the first order neighbourhood of a line.

The final answer to the original question was achieved in the ADHM (Atiyah-Hitchin-Drinfeld-Manin) construction [6] but Simon Donaldson has explained that in his talk in the series. Instead, let me point out the collaboration between Atiyah and Ward which preceded this [5].

How do you construct a holomorphic vector bundle on $\mathbf{C}\mathrm{P}^3$? By way of analogy, consider the problem of line bundles on a Riemann surface C. Suppose the line bundle L of degree $d \geq 0$ has a holomorphic section. It will vanish at points $x_1, \ldots, x_d$ and in the complement we have a non-vanishing section, namely a trivialization of L. On the other hand, by the definition of a line bundle there are local trivializations on open discs D_i containing x_i. Taking a covering $\{U_\alpha\}$ of C by the D_i and $C \backslash \{x_1, \ldots, x_d\}$, on twofold intersections $U_\alpha \cap U_\beta$ the trivializations differ by a holomorphic function $g_{\alpha\beta}$ with values in $\mathbf{C}^*$. Conversely, if z_i is a local coordinate with $z_i(x_i) = 0$ then z_i on the punctured disc D_i defines an equivalent family of transition functions. So the points $x_1, \ldots, x_d$ define L.

Suppose then that a rank 2 holomorphic vector bundle E on $\mathbf{C}\mathrm{P}^3$ has a section. Now the zero set is defined locally by two functions and is therefore a curve $C \subset \mathbf{C}\mathrm{P}^3$. In the complement we no longer have a trivialization but we do have a non-zero section giving a trivial subbundle and this means that there are transition functions of the form

$$g_{\alpha\beta} = \begin{pmatrix} 1 & a_{\alpha\beta} \\ 0 & b_{\alpha\beta} \end{pmatrix}.$$

Here $b_{\alpha\beta}$ is a transition function for the quotient line bundle L of E by the trivial subbundle and the multiplicative property of the transition function on threefold intersections means that $a_{\alpha\beta}$ is a Čech representative for a class in $H^1(\mathbf{C}\mathrm{P}^3 \backslash C, L^*)$. If $L \cong \mathcal{O}(2)$ then, as we have seen, this is a solution to Laplace's equation. The class does not extend to $\mathbf{C}\mathrm{P}^3$ but it does as a class in $\mathrm{Ext}^1_{\mathbf{C}\mathrm{P}^3}(\mathcal{J}_C, \mathcal{O}(L^*))$ where $\mathcal{J}_C$ is the ideal sheaf of C. Under certain conditions the curve C and some scalar information on C determines E up to equivalence. This was the so-called Serre construction of vector bundles.

Taking C to be a collection of disjoint lines in $\mathbf{C}\mathrm{P}^3$ yields the pseudoparticle solutions of (1). The class in $H^1(\mathbf{C}\mathrm{P}^3 \backslash \{C\}, \mathcal{O}(-2))$ corresponds via the Penrose transform to the solution ρ of the Laplace equation in the Ansatz. By considering L isomorphic to $\mathcal{O}(n)$ more solutions were obtained from a modified Ansatz using more general zero rest mass field equations, but the condition to be trivial on every real line turned out to be more difficult.

In the meantime a second advance was a use of the index theorem. Any automorphism of the principal bundle (a *gauge transformation*) takes one solution of the equations to another so a parameter count must take this into account by considering the moduli space – the set of solutions modulo gauge equivalence. To first order, a gauge transformation is given by a section ψ of the bundle $\mathfrak{g}$ of Lie algebras and its effect on the connection is the variation $\nabla\psi \in \Omega^1(S^4, \mathfrak{g})$, a 1-form with values in $\mathfrak{g}$. An arbitrary first order variation $\dot{A}$ of the connection satisfies the self-duality equation if the anti-self-dual component $(d_A)_-$ of the variation of the curvature $d_A\dot{A}$ vanishes. Then the kernel of the differential operator $(d_A)_-$ transforming $\Omega^1(S^4, \mathfrak{g})$ to $\Omega^2_-(S^4, \mathfrak{g})$ contains the first order deformations and this, modulo the image of ∇_A acting on sections of $\mathfrak{g}$ should give the parameter count.

This is the first cohomology of an elliptic complex on S^4:

$$\Omega^0(S^4, \mathfrak{g}) \stackrel{\nabla_A}{\to} \Omega^1(S^4, \mathfrak{g}) \stackrel{(d_A)_-}{\to} \Omega^2_-(S^4, \mathfrak{g}) \tag{2}$$

and the index theorem (for the elliptic operator $(d_A)_- \oplus \nabla_A^*$) gives $8k-3$ for the dimension if the cokernel is zero. The positivity of the (scalar) curvature of S^4 gives this via a differential-geometric vanishing theorem. The two topological invariants for the right hand side of the index theorem are the Euler characteristic 2 of the sphere and the second Chern class $-k$ of the rank 2 bundle. These $8k-3$ dimensions were significantly more than was being produced by constructions during this period and so pointed towards a better approach.

The algebraic geometry of bundles on projective space was a subject of considerable interest at the time, independent of the instanton problem, in particular work by Wolf Barth in Germany, Geoffrey Horrocks in England and Robin Hartshorne in the USA. More streamlined methods were being used for this problem, the key algebraic approach was to consider the direct sum of the sheaf cohomology groups $H^1(\mathbf{C}\mathrm{P}^3, E(n))$ for all n as a module over the direct sum of $H^0(\mathbf{C}\mathrm{P}^3, \mathcal{O}(n))$ which is the algebra of polynomials in four variables z_1, z_2, z_3, z_4. Each of these spaces has a Penrose transform interpretation but the algebraic geometry was more flexible and showed that ultimately all depended on the first part $H^1(\mathbf{C}\mathrm{P}^3, E(-1))$ so long as $H^1(\mathbf{C}\mathrm{P}^3, E(-2)) = 0$. Here the Penrose transform came to the rescue with positive curvature showing that the only solutions to the Laplace equation coupled to the connection were zero. This last result set the algebraic geometry machinery in motion and reduced the construction to one of matrices.

For Michael Atiyah this whole experience would be, I believe, a pivotal moment in his attitude to the relationship between geometry and physics.

A little later, Michael returned to twistor theory and revisited the material that went into the Serre construction. Physicists [28] had already in 1978 used the ADHM construction to explicitly calculate the Green's function for the Laplacian coupled to a self-dual connection and the issue of how to define a Green's function in terms of twistor theory arose.

It is more natural to use conformal invariance and consider the conformally invariant Laplacian which is $\Delta + R/6$ with R the scalar curvature of a metric in the conformal class, then because of the positivity of curvature of the 4-sphere, the Green's function is unambiguously defined. The answer, for the 4-sphere, as we remarked, is to take the ideal sheaf $\mathcal{J}_x$ of the twistor line L_x corresponding to $x \in S^4$ and the Serre class $\lambda(x) \in \mathrm{Ext}^1_{\mathbf{C}\mathrm{P}^3}(\mathcal{J}_x, \mathcal{O}(-2))$. The full two-variable Green's function $G(x, y)$ is the Serre class of the diagonal in $\mathbf{C}\mathrm{P}^3 \times \mathbf{C}\mathrm{P}^3$.

The Green's function paper [7] appeared in 1981 by which time there was a more general setting. By analogy with the gauge-theoretical instantons, Gary Gibbons and Stephen Hawking had produced gravitational analogues. These were Euclidean solutions of Einstein's vacuum equations with the additional property that the Levi-Civita connection was self-dual. In 1976, Penrose had also introduced his nonlinear graviton construction which replaced the twistor space $\mathbf{C}\mathrm{P}^3$ by a more general 3-manifold Z with the key property that it contained twistor lines – holomorphic $\mathbf{C}\mathrm{P}^1$s with normal bundle isomorphic to $\mathcal{O}(1) \oplus \mathcal{O}(1)$ and the Gibbons-Hawking examples fitted nicely into this picture. A competition then began with the physicist Don Page to calculate the Green's function for these examples. Michael and I used to walk from the Institute across the University Parks to St Catherine's College for lunch and I remember him complaining that he would never do a long intricate calculation like that again. I think a bottle of champagne was at stake, but I never found out who won it.

5 Gauge theory in mathematics

5.1 Two dimensions

The instanton episode clearly fitted the market force model – mathematicians responding to the needs of physicists. But it provided geometers with a new way of looking at familiar objects, namely connections on principal bundles. Differential geometers by and large had considered these one-by-one, but rarely as an infinite dimensional space of them. Automorphisms of bundles were also studied, though often moving the base too, so the consideration of gauge transformations as an infinite-dimensional group was unusual. The instanton question introduced a new viewpoint – considering the quotient of an infinite-dimensional space of connections by an infinite dimensional group of automorphisms. This was not a strange territory for physicists – the space of connections was the basis for the functional integral approach to gauge theories, and gauge fixing represented difficulties in carrying it out. The closest mathematical analogue was the analytical study of the moduli of complex structures on an underlying compact manifold as developed by Kodaira, Kuranishi and others. Following this analogy and using the properties of Sobolev spaces and Banach space inverse function theorems, the first order deformation theory arising from the elliptic complex (2) was used to define a moduli space, a smooth manifold of dimension $8k-3$.

When $k=1$ this moduli space is 5-dimensional and is acted on by the group of conformal transformations of the 4-sphere $SO(5,1)$. It did not need the ADHM construction to show that this was hyperbolic 5-space. The key point that its boundary was the original 4-manifold went unnoticed until Simon Donaldson started his momentous work.

There was one example in the literature of such a moduli space of connections and this was the theorem of M.S.Narasimhan and C.S.Seshadri concerning stable holomorphic vector bundles on a Riemann surface Σ. In [46] they had shown that a holomorphic vector bundle on a compact Riemann surface which satisfied an algebro-geometric stability condition admitted a natural flat unitary connection. In their proof, the connection barely appears since a flat connection can be viewed geometrically rather than analytically as either a vector bundle with transition matrices which are locally constant, or simply as the quotient of $\tilde{\Sigma}\times\mathbf{C}^n$ (where $\tilde{\Sigma}$ is the universal covering) by the action of the fundamental group $\pi_1(\Sigma)$, acting on $\mathbf{C}^n$ via a homomorphism $\rho:\pi_1(M)\rightarrow U(n)$. This was an important result, linking topology and algebraic geometry, for it provided two very different structures on the moduli space. One was as an algebraic variety, the other as the space of unitary $n\times n$ matrices A_i,B_i satisfying the constraint $A_1^{-1}B_1^{-1}A_1B_1\cdots A_g^{-1}B_g^{-1}A_gB_g=1$ up to conjugation. In [37] Harder and Narasimhan had calculated the Betti numbers of these moduli spaces by number-theoretic means, using the proof of the Weil conjectures.

Atiyah and his long-time collaborator Raoul Bott, who had recently spent time at the Tata Institute with Narasimhan, set out in 1980 to use the Yang-Mills functional, not in four dimensions, but in two, to analyse this situation. Connections on a fixed smooth complex vector bundle E with a Hermitian inner product form an infinite-dimensional affine space $\mathcal{A}$ – the difference of any two covariant derivatives $\nabla_1-\nabla_2$ lies in $\Omega^1(\Sigma,\mathfrak{g})$ where $\mathfrak{g}$ is the bundle of skew-Hermitian endomorphisms of E. The Yang-Mills functional

$$\int_\Sigma |F_A|^2\nu$$

(where ν is the area form of a metric in the conformal class) is a natural gauge-invariant function on this space and the absolute minimum is the space of flat connections. The quotient by the group of unitary gauge transformations is the space studied by Narasimhan and Seshadri. The idea of Atiyah and Bott was to use Morse theory in infinite dimensions to give an alternative approach to finding the cohomology of the moduli space.

Recall the classical picture of Morse theory. We are given a compact manifold M and a real-valued function f with nondegenerate critical points. At a critical point the Hessian, the second derivative of f, is a well-defined quadratic form on the tangent space and its *index* is the number of negative eigenvalues. At a minimum value, say 0, the index is zero and one builds up the homotopy type of the manifold from that of $f^{-1}[0,t]$ by adding a cell of dimension n as t passes through a critical value of index n.

The aim of Atiyah and Bott was to use this the other way round, to determine by subtraction the topology of the absolute minimum from the topology of M and the other critical points, or critical submanifolds in this case. The infinite-dimensional affine space $\mathcal{A}$ is of course contractible, but the quotient $\mathcal{A}/\mathcal{G}$ by the group of gauge transformations is not and its homotopy type is determined by that of the maps from Σ to the classifying space $BU(n)$. The other critical points are where the first variation of the functional is zero for a variation $\dot{A} \in \Omega^1(\Sigma, \mathfrak{g})$ of the connection. This is

$$\int_\Sigma (d_A\dot{A}, F_A)\nu = \int_\Sigma (\dot{A}, d_A^* F_A)\nu$$

and hence vanishes if $d_A^* F_A = 0$ or equivalently $\nabla_A(*F_A) = 0$. This means that $*F_A$ is a covariant constant section of the Lie algebra bundle and so its eigenspaces decompose the vector bundle E as a direct sum of bundles of lower rank where the curvature of the connection is a scalar constant. The idea is then to generalize slightly to scalar curvature rather than zero curvature, use induction on the rank n of the vector bundle and set up an inductive procedure to determine the cohomology of the absolute minimum.

Although Bott had proved the eponymous periodicity theorems for the homotopy groups of the classical groups by using Morse theory on an infinite-dimensional loop space this programme represented considerable technical challenges and the interested reader is referred to [42]. The end result was perhaps less important than the features encountered on the journey which became influential in the further development of gauge theory in mathematics and in the importation of physical concepts. This spin-off took various forms.

5.2 Symplectic aspects

The affine space $\mathcal{A}$ is acted on transitively by $\Omega^1(\Sigma, \mathfrak{g})$ so the tangent space at any point A is naturally isomorphic to this space. Given two tangent vectors $\dot{A}, \dot{B} \in \Omega^1(\Sigma, \mathfrak{g})$ there is a canonical skew pairing

$$\int_\Sigma \mathrm{tr}(\dot{A} \wedge \dot{B})$$

and this is formally non-degenerate. It is also translation invariant and so formally it defines a closed 2-form, a symplectic form.

The relationship in Hamiltonian mechanics between linear momentum and the translation group, and angular momentum and the rotation group, leads to the more general notion of moment map for a manifold M with a symplectic form ω and the action of a Lie group G preserving ω. A momentum map is a map $\mu : M \to \mathfrak{g}^*$ to the dual of the Lie algebra which commutes with the actions of G and is such that for $a \in \mathfrak{g}$, the real valued function $f_a = \langle \mu, a \rangle$ is a Hamiltonian function for the vector field X_a generated by a. In other words $\omega(X_a, Y) = df_a(Y)$ for any tangent vector Y.

Here we have, formally, $M = \mathcal{A}$ the space of connections, $G = \mathcal{G}$ the group of gauge transformations with Lie algebra the space of sections a of the Lie algebra bundle. Then, as Atiyah and Bott pointed out, the curvature F_A is a moment map. We have

$$f_a = \langle \mu, a \rangle = \int_\Sigma \mathrm{tr}(F_A a)$$

and

$$df_a(\dot{A}) = \int_\Sigma \mathrm{tr}(d_A \dot{A} a) = \int_\Sigma \mathrm{tr}(\dot{A} \wedge d_A a)$$

and since $d_A a$ is the variation of the connection (just as in the 4-dimensional case of the complex (2)) F_A is the moment map.

From this viewpoint the space of flat connections is $\mu^{-1}(0)$ and the moduli space $\mu^{-1}(0)/\mathcal{G}$ is a symplectic quotient or reduced phase space and acquires a symplectic structure.

5.3 Holomorphic structures

The traditional view of a holomorphic structure on a vector bundle E over a Riemann surface is via transition functions $g_{\alpha\beta} : U_\alpha \cap U_\beta \to GL(n, \mathbf{C})$, but these give rise to a global $\bar{\partial}$-operator

$$\bar{\partial} : C^\infty(E) \to C^\infty(E \otimes \bar{K})$$

as we noted above. When Σ has a complex structure, by differentiating in the anti-holomorphic direction, any covariant derivative ∇ on E defines a differential operator $\nabla^{0,1}$ with the property that $\nabla^{0,1}(fs) = \bar{\partial} f s + f \nabla^{0,1}(s)$ where s is a section of E, f a function and $\bar{\partial}$ the Cauchy-Riemann operator. Conversely, given a $\bar{\partial}$-operator the Hermitian form defines a complex conjugate which is a ∂-operator and $\bar{\partial} + \partial$ is the covariant derivative ∇ of a connection. Atiyah and Bott showed, although in fact it had previously been proved by Koszul and Malgrange, that the local solutions to $\nabla^{0,1}(s) = 0$ are the local holomorphic sections of a holomorphic vector bundle.

Thus the space of connections is also the space of holomorphic structures and the complexified gauge group $\mathcal{G}^c$ is a well-defined object taking one $\bar{\partial}$-operator to another by conjugation. In particular $\mathcal{A}$ is an infinite-dimensional complex manifold with an infinite-dimensional group $\mathcal{G}^c$ of holomorphic transformations. Together with the symplectic structure $\mathcal{A}$ is an infinite-dimensional Kähler manifold and the moduli space inherits a natural Kähler metric.

5.4 Stability

These two differential-geometric observations placed the Narasimhan-Seshadri theorem in a new light. A holomorphic vector bundle on a Riemann surface Σ is said to be *stable* if for each subbundle $U \subset E$ we have

$$\frac{\deg U}{\operatorname{rk} U} < \frac{\deg E}{\operatorname{rk} E}.$$

The theorem states that a stable bundle of degree zero has a unique flat unitary connection. From the gauge-theoretic viewpoint this means that the stable holomorphic structures $A \in \mathcal{A}$ are those which are transformed by an action of $\mathcal{G}^c$ to the zero set of the moment map $A \mapsto F_A$. It presented the opportunity for a new proof of the theorem – to start with a $\bar{\partial}$-operator on a fixed C^∞ vector bundle with a Hermitian metric and try and find the minimum of the Yang-Mills functional on a $\mathcal{G}^c$-orbit. The stability condition should imply the existence of a minimum, which is zero. This was Donaldson's first paper [30] and he used the analysis developed by Karen Uhlenbeck mainly for the 4-dimensional Yang-Mills problem to get convergence results.

Simon Donaldson had been my research student but it was when he observed a similar phenomenon that I transferred him to Michael Atiyah, who was enthusiastically pursuing the moment map idea. The problem I suggested to Donaldson was the issue of stability for holomorphic vector bundles in higher dimensions, and specifically in two complex dimensions. A two-dimensional Kähler manifold has a canonical orientation and with respect to this the anti-self-dual condition on a connection implies that $\nabla^{0,1}$ defines a holomorphic structure but there is one more condition, that contraction of the curvature F_A with the Kähler form ω (the Λ-operator in Hodge theory) should be zero. There was a clear conjecture here that the algebraic geometry notion of stability was related to this curvature condition. There was evidence:

1. in one complex dimension this is just the vanishing of the curvature as in Narasimhan-Seshadri

2. Yau's proof of the Calabi conjecture gave examples for the tangent bundle of a K3 surface

3. an instanton connection on S^4 pulls back to $\mathbf{C}\mathrm{P}^3$ under the projection $\mathbf{C}\mathrm{P}^3 \to S^4$ to a connection which satisfies this condition in three complex dimensions. It was also known that these were stable bundles.

Donaldson formulated the conjecture in moment map terms on the space of connections $\mathcal{A}$. Here one needs the Kähler form to define a formal symplectic structure, which was unnecessary in the one-dimensional case. The curvature now has different components corresponding to the decomposition of 2-forms on a complex manifold into types $(1,1)$, generated by $dz_i \wedge d\bar{z}_j$, $(2,0)$ generated by $dz_i \wedge dz_j$ and $(0,2)$ by $d\bar{z}_i \wedge d\bar{z}_j$. The vanishing of the $(0,2)$ part of F_A is the condition for $\nabla_A^{0,1} s = 0$ to have local solutions and define a holomorphic structure. This subspace of $\mathcal{A}$ is a nonlinear complex submanifold and then $\Lambda F_A = 0$ is the vanishing of the restricted moment map, so stability should be equivalent to transforming by a $\mathcal{G}^c$ gauge transformation to a zero of the moment map. Donaldson proved this for surfaces [32] and Uhlenbeck and Yau in higher dimensions [54], but this came after Donaldson, as Atiyah's student, turned his attention to instantons on general four-dimensional manifolds with spectacular results.

Michael's other student at the time, Frances Kirwan, was given the *finite dimensional* problem of using moment maps and equivariant Morse theory to calculate the cohomology of quotients in algebraic geometry using the same ideas and the norm squared of the moment map instead of the Yang-Mills functional. Lessons learned from a gauge-theoretic approach were valuable in pure algebraic geometry.

5.5 Convexity

There was another aspect of symplectic geometry which made its presence felt during this piece of research. The brief sketch above of how Morse theory operates should be amplified by using a Riemannian metric and considering the gradient flow of f. Each point flows down to a critical point x and the upward flow from x is isomorphic to $\mathbf{R}^m$ where m is the number of positive eigenvalues of the Hessian. Thus the upward flow from the absolute minimum is open. These stable submanifolds give a cell decomposition of M and the behaviour of their closures encompasses the complexity of the cohomology. When the critical points form submanifolds $N \subset M$ then these are total spaces of vector bundles over N.

There is a partial ordering on the critical submanifolds N_i. Considering the downward gradient flow of f then $N_1 < N_2$ if there is a trajectory starting on N_1 and whose limit lies in N_2. In Atiyah and Bott's case the critical submanifolds are direct sums of bundles $E_1, \ldots, E_k$ of rank $n_1, \ldots, n_k$ and degree $d_1, \ldots, d_k$. A partial ordering was needed among these sets of integers.

In fact, the gradient flow of the Yang-Mills functional was difficult to apply but instead there existed a ready-made stratification by considering $\mathcal{A}$ as the space of holomorphic structures. A holomorphic vector bundle may not be stable (or semi-stable which is where equality occurs in the definition) but instead there is the Harder-Narasimhan filtration [37]

$$0 = E_0 \subset E_1 \subset E_2 \subset \cdots \subset E_n = E$$

where $D_i = E_i/E_{i-1}$ is semistable and $\mu(D_1) > \mu(D_2) \cdots > \mu(D_n)$ with slope $\mu(V) = \deg V/\operatorname{rk} V$. Stability is an open condition so the stratum for $E = E_n$ and $\mu = 0$ is the analogue of the upward flow from the absolute minimum of the functional.

Using a result of Shatz [52] about how the $\mu(D_i)$ specialize under limits the partial ordering is organized by taking the vector of slopes $(\mu_1, \ldots, \mu_n)$ including repetition and defining $\lambda \geq \mu$ if and only if

$$\sum_{j \leq i} \lambda_j \geq \sum_{j \leq i} \mu_j \tag{3}$$

for $1 \leq i \leq n-1$. Here $\sum \mu_i = \sum \lambda_i = \deg E$ which is a topological invariant.

In subsequent papers [9],[8] Atiyah presents this in the context of the action of the torus T of diagonal matrices on the symplectic manifold $U(n)/T$, a coadjoint orbit. The moment map takes for each generator of the Lie algebra $\mathfrak{t}$ of T a Hamiltonian function and gives an equivariant map $\mu : U(n)/T \to \mathfrak{t}^* \cong \mathbf{R}^n$. The theorem is that the image is convex. To see the link, let $a \in \mathfrak{u}(n)$ be a skew Hermitian matrix with distinct eigenvalues $(i\lambda_1, \ldots, i\lambda_n)$ and consider the orbit in the Lie algebra which we can identify with its dual using the inner product $\operatorname{tr}(ab)$. The λ_i are of course constant on the orbit. The moment map for the action of $U(n)$ is just the inclusion but restricted to the torus of diagonal matrices in $U(n)$ it is the projection onto the Lie algebra. So the moment map is $(\mu_1, \ldots, \mu_n)$ where these are the diagonal entries. The convexity of the image gives the inequality (3).

In fact, as Atiyah writes:

> When I visited Harvard I lectured on this material in Bott's seminar and was mystified by the looks of amusement on several faces in the audience. It transpired that Guillemin and Sternberg had almost simultaneously found the same result!.

These insights fed into other fields, clearly into generalizing to Lie groups, where Kostant had already trodden, but also into stationary phase approximations and equivariant cohomology.

This journey into two dimensions was motivated by the four-dimensional problem of instantons and its moduli space was not neglected, but we shall return to this later.

6 From mathematics to physics

6.1 Edward Witten

Ideas flowed in both directions in the 1980s when Michael assumed the roles of enthusiast, educator and facilitator. Many of the crucial interactions arose through discussions with Edward Witten. Atiyah and Witten first met in MIT in 1977. Witten had written a paper relating $SO(3)$-invariant instantons on S^4 to the 2-dimensional gauge theoretic problem of abelian vortices on the hyperbolic plane [55]. Atiyah invited him to visit Oxford the following year and a new phase in the interaction of mathematics and physics began. As Witten comments [53].:

> Physicists at the time were very interested in solving the instanton equations because of speculation by Alexander Polyakov about the dynamics of gauge theories. However, the ingredients in the twistor transform of the instanton equation – complex manifolds, sheaf cohomology, fiber bundles – were quite unfamiliar to me and most other physicists

Without necessarily going into the details, Atiyah developed a nose for the parts of mathematics which physicists would find useful and absorbed through his conversations with physicists, Witten in particular, and also his collaborators Bott and Singer, the radically different viewpoints in the two disciplines.

Here is Witten's recollection [53]: '

> At the 1979 Cargèse summer school, Atiyah and Raoul Bott undertook to educate physicists about Morse theory. I and most (or all?) of the physicists there had certainly never been exposed to Morse theory before. Another highlight was a conference in Texas where Atiyah and Is Singer began to elucidate the topological meaning of what physicists know as perturbative anomalies in gauge theory. This helped introduce physicists to a deeper understanding of fermion path integrals. Two papers by Atiyah and Bott in these years were ultimately influential for physicists. Their 1983 paper "The Yang - Mills equations over a Riemann surface" introduced ideas that were important later in understanding quantum gauge theories in two dimensions. Their 1984 paper "The moment map

and equivariant cohomology" helped lead to the important technique of " localization" in supersymmetric quantum field theory.

Atiyah's contribution to the interaction was one of issuing challenges to the physics community and educating them in the parts of mathematics he thought would be of most interest and use. The process accelerated in the mid 1980s when string theory provided a much wider interface between the two disciplines. The challenges included making sense of the relationship between Langlands duality of Lie groups and electric and magnetic charges, and of finding a quantum field theory to provide a context for the ongoing work of Donaldson on four-manifold invariants and, perhaps most successfully, understanding the Jones polynomial as a topological quantum field theory (see Edward Witten's article based on a talk in the series). Witten's description above of part of the educational process represents just a few of the interactions. In the next section we shall see what happened to Morse theory in Witten's hands. To paraphrase Goethe, replacing mathematicians by physicists: "whatever you say to them they translate into their own language and forthwith it is something entirely different."

6.2 Morse theory à la Witten

Atiyah and Bott's use of Morse theory led to an evaluation of the cohomology because they had shown that their Morse function was *perfect*, or that the stratification behaved that way. This was really a consequence of the complex structure on the strata – if the index of a critical point is even then much more can be read off. In general one obtains inequalities.

If m_p is the number of critical points of index p then the Morse inequalities are $\dim H^p(M,\mathbf{R}) \leq m_p$ and

$$\sum_0^k (-1)^p m_p \geq \sum_0^k (-1)^p \dim H^p(M,\mathbf{R}).$$

Witten uses the de Rham description of the cohomology groups

$$H^p(M,\mathbf{R}) = \frac{\ker d : \Omega^p \to \Omega^{p+1}}{\operatorname{im} d : \Omega^{p-1} \to \Omega^p}$$

and the Hodge theorem: if M is compact, then each cohomology class has a unique harmonic representative: $d\alpha = 0, d^*\alpha = 0$.

For Witten the direct sum of forms of odd degree Ω^{od} is to be regarded as a space of fermionic states and the even forms Ω^{ev} as bosonic states. There are supersymmetry operators exchanging bosons and fermions and these are

$$Q_1 = d + d^*, \quad Q_2 = i(d - d^*)$$

while the Hamiltonian $H = dd^* + d^*d$ is the Laplacian on forms. These satisfy the supersymmetry relations

$$Q_1^2 = Q_2^2 = H, \quad Q_1Q_2 + Q_2Q_1 = 0.$$

Now introduce the Morse function f and conjugate d to give $d_t = e^{-ft}de^{ft}$ and $d_t^* = e^{-ft}d^*e^{ft}$. Clearly $d_t^2 = 0$ and the cohomology groups defined by d_t are isomorphic to the de Rham cohomology. However the Hamiltonian is now

$$H_t\psi = (d_td_t^* + d_t^*d_t)\psi = H\psi + t^2(df)^2\psi + t(\nabla^2 f)_{ij}[e_i\wedge, i_{e_j}]\psi$$

(where i_{e_j} denotes contraction, or inner product, with an element of a local orthonormal basis of tangent vectors). The expression is a Laplacian plus a potential term V, familiar territory for a physicist.

Now let $t \to \infty$ and V becomes large except at the critical points where $df = 0$ and this implies that the eigenfunctions of H_t are concentrated near the critical points. Moreover there is an asymptotic expansion of the eigenvalues

$$\lambda \sim t(a + bt^{-1} + ct^{-2} + \ldots)$$

where the coefficients are local expressions around the critical points. Fix p and consider the eigenvalues on p-forms. For large t the number of eigenvalues which vanish is no larger than the number of coefficients a which vanish and this is equal to the number of negative eigenvalues of the Hessian term $\nabla^2 f$. Thus the number of critical points of index $p = m_p \geq \dim H^p(M)$, the first Morse inequality.

To obtain the stronger form of the inequality requires studying the orbits of the gradient flow, passing from one critical point to another, the "paths of steepest descent". These now acquire the terminology of "instantons" "tunneling" between two states. Witten introduced thereby a complex which was in some sense a revival of early work by Morse, Smale, Thom, and Milnor but because of its setting it was highly influential, in particular the development of Andreas Floer [35].

The meeting of Hodge theory with Morse theory also meant that the method could be adapted by replacing the exterior derivative d with the $\bar{\partial}$ operator on a complex manifold [29], [57]. This leads to asymptotic inequalities as $p \to \infty$ of the form

$$\sum_0^k (-1)^{k-j} \dim H^j(M, L^p) \leq \frac{p^n}{n} \int_N p(R)$$

where the integral is a curvature term over a subspace N determined by k. These have direct consequences in algebraic and complex geometry.

7 From physics to mathematics

One of the enduring ideas coming from physics is that of a *topological quantum field theory*. Until the mid 1980s, mathematicians were appreciating the doors which were opening by using gauge theories, but regarded as issues surrounding classical solutions to nonlinear partial differential equations. Quantum theory may well have been in the motivation but did not appear in the mathematical problem. That began to change, especially when Vaughan Jones saw the link between von Neumann algebras and knot theory and Atiyah began to issue challenges to the physics community to find a structure which would explain this.

At the same time, photocopies of a draft paper by Graeme Segal called "The definition of conformal field theory" began to circulate. Years later it appeared in modified form [51]. This introduced the idea of a categorical formulation associating to a Riemann surface with k boundary components the tensor product of Hilbert spaces, one for each component. Orientations on the boundary distinguish inbound and outbound components and a gluing process enables the picture for a general surface to be built up from basic pieces.

This scenario, stripped of the complex structure, was the basis for Atiyah's axiomatic approach to the topological version [12], influenced by Witten. This associates to each Σ, a compact oriented d-dimensional manifold, a complex vector space $Z(\Sigma)$, and to each compact oriented $d+1$-manifold with boundary Σ a vector $Z(M) \in Z(\Sigma)$. Then the axioms are:

1. Z is functorial with respect to orientation-preserving diffeomorphims
2. if Σ^* is Σ with the opposite orientation then $Z(\Sigma^*) = Z(\Sigma)^*$
3. for a disjoint union $Z(\Sigma_1 \cup \Sigma_2) = Z(\Sigma_1) \otimes Z(\Sigma_2)$
4. if $M = \emptyset$ then $Z(\emptyset) = 1 \in Z(\partial\emptyset) = \mathbf{C}$
5. $Z(M^*) = \bar{Z}(M)$.

Axiom 3 for the tensor product of the disjoint union is the feature which to a topologist is the most unusual, and somehow represents the quantum content of the definition. If M has empty boundary then $Z(M) \in \mathbf{C}$ is an invariant.

This abstraction is a long way from the physical motivation but Σ is modelled on space and the extra dimension of M is time, $Z(M)$ is the vacuum state in the Hilbert space $Z(\Sigma)$. The cylinder $M = \Sigma \times [0, 1]$ has two boundary components with opposite orientation and then Axiom 2 says that the "propagation" from one Hilbert space to the other is the identity, or that the Hamiltonian $H = 0$, which is what the "topological" adjective is meant to imply.

The relationship in differential topology between manifolds and their boundaries is called cobordism and the axioms represent a functorial relationship between two corresponding categories. They offer an organizing goal for associated results which in many cases is difficult to fully achieve, but this is one of the challenges which physics presents to mathematics.

In low dimensional cases where manifolds are all known up to diffeomorphism, we can recognize the structure the axioms yield. In particular when $d = 1$, the circle is the only compact 1-dimensional manifold and all surfaces with boundary are obtained by gluing discs and pairs of pants. So $Z(S^1) = V$ is a vector space and if we take a pair of pants M with one incoming boundary and two outgoing ones we get a vector

$$v \in V^* \otimes V^* \otimes V$$

which is a commutative bilinear product on V. More generally if there are p incoming boundaries and q outgoing then M defines a homomorphism from $\otimes^p V$ to $\otimes^q V$. Taking M to be a disc with a single ingoing boundary defines a vector in V^* which is a "trace" $\theta(a)$, and the disc with an outgoing boundary a unit 1. This defines a *Frobenius algebra* structure on V. It is a theorem that any Frobenius algebra can be viewed this way [1].

The even degree cohomology $H^{ev}(M, \mathbf{C})$ of a compact oriented manifold M is an example, the trace of a de Rham representative is the integral of the top degree component. For a complex projective variety, the quantum cohomology is a deformation of this which involves the enumerative geometry of rational curves in M.

One could also take the group ring of a finite abelian group Γ – any $u \in V$ is a linear combination

$$u = \sum_{i=0}^{n} \alpha_i g_i$$

with $g_i \in \Gamma, g_0 = e$ the identity and $\alpha_i \in \mathbf{C}$. The trace is defined by $\theta(u) = \alpha_0$. Dually, the space V^* consists of functions on Γ and group multiplication corresponds to convolution

$$f * g(x) = \sum_{h \in \Gamma} f(xh^{-1})g(h)$$

and $\theta(f) = f(e)$, evaluating at the identity. For a non-abelian group, the space of functions on conjugacy classes is the character ring, spanned by the characters of irreducible representations, and this with the convolution product as functions on Γ is a commutative Frobenius algebra. The characters χ_1, χ_2 of two inequivalent irreducible representations satisfy $\chi_1 * \chi_2 = 0$ and $\chi_1 * \chi_1 = |\Gamma| \chi_1 / \chi_1(e)$ so the structure of the algebra is quite simple.

This example is related to the question of flat connections on surfaces: a case with $d = 1$. A flat connection on a space X with holonomy in the finite group Γ is just a principal Γ-bundle, a finite Galois covering. Up to equivalence these are parametrized by $\mathrm{Hom}(\pi_1(X), \Gamma)$ modulo the conjugation action of Γ. So if X is the circle these are the conjugacy classes.

If $X = M$ is a surface with incoming boundary Σ^- and outgoing boundary Σ^+ then a flat connection on M restricts to one on the boundary, so a function on $\mathrm{Hom}(\pi_1(\Sigma^-), \Gamma)$ can also be considered as a function on the finite set of flat connections on M. Summing over the inverse images of the map to $\mathrm{Hom}(\pi_1(\Sigma^+), \Gamma)$ and weighting by the inverse of the number of automorphisms of the connection defines $Z(M) \in Z(\Sigma)$. Together with the Frobenius algebra above, this satisfies the axioms.

In physical language the conjugacy classes of Γ form a classical phase space and the functions on it are the quantization. Then clearly we have the tensor product property for the disjoint union. The invariant for a closed surface M is the number of equivalence classes of flat Γ-connections.

Replacing Γ by a compact Lie group G gives an infinite-dimensional Hilbert space V and Witten computed the Frobenius algebra structure giving a similar result to the finite group case. Now instead of counting the number of flat connections one obtains the volume of the moduli space $\mathcal{M}$. Recall that Atiyah and Bott had observed the natural symplectic form ω on $\mathcal{M}$. Witten's formula is

$$\int_{\mathcal{M}} \frac{1}{n} \omega^n = |Z(G)| \, \mathrm{vol}(G)^{2g-2} \sum_R (\dim R)^{2-2g}$$

where R runs through the irreducible representations of G and $Z(G)$ is the centre and $\dim \mathcal{M} = 2n$. Since ω is closed the volume is the cohomology class $[\omega]^n/n$ evaluated on a fundamental class and this brought into play the *multiplicative* properties of the generators of the cohomology of $\mathcal{M}$ which Atiyah and Bott had determined. This was being pursued by geometers from various points of view but one of them related to moving TQFTs one dimension higher.

This was the context for the Jones polynomials but it is characteristic of the difference in approach of physicists and mathematicians, or at least those with a geometric background, that instead of considering a simple space like the 2-sphere with distinguished points on it (which led to knots and braid groups), geometers prefer higher genus surfaces with no embellishments. For a closed oriented surface Σ the analogue of the set of conjugacy classes to discuss connections on the circle is the moduli space

$\mathcal{M}$ of flat G-connections. Its quantization is no longer the Hilbert space of functions on $\mathcal{M}$, but instead one has to apply geometric quantization as developed by Kostant and Souriau in the 1970s. Firstly it is a Hilbert space of sections of a *line bundle* rather than just functions – a line bundle with connection whose curvature is a multiple of the symplectic form ω, which introduces an integer k, the level, the Chern class of this bundle. Secondly the section has to be constant relative to a polarization of the symplectic structure and given the holomorphic interpetation of $\mathcal{M}$ this means being holomorphic with respect to a complex structure such that ω is the Kähler form. Determining its dimension is the first task and this, for a geometer, involves the full multiplicative structure of the cohomology of $\mathcal{M}$ and not just the volume, which gives just the leading term in a polynomial in the level k. The formula was known to physicists (the "Verlinde formula") in terms of conformal blocks.

Creating a genuine topological field theory out of this data was only achieved much later by roundabout methods. Key issues were to prove, in a suitable setting, that the space of sections is independent of the choice of complex structure on the surface Σ and also proving the unitarity of Axiom 5. For this case at least, the structure was a framework for proving what had to be true rather than a means of attaining it.

Atiyah pursued these ideas with enthusiasm, explaining them to a general audience later in the little book [15]. A turning point occurred at a dinner in a restaurant in 1988 at the International Conference on Mathematical Physics. As Witten recalls [53],

> Eventually, at a meeting in Swansea where I had the benefit of further discussions with Atiyah and with Graeme Segal, I had the good fortune to put some of the pieces together and interpret the Jones polynomial in terms of a three-dimensional gauge theory with the Chern-Simons function as its action.

In 1990 Fields Medals were awarded to Edward Witten, Vaughan Jones, Vladimir Drinfeld and Shigefumi Mori. The Selection Committee consisted of L.Faddeev (chair), M.Atiyah, J-M.Bismut, E.Bombieri, C.Fefferman, K. Iwasawa, P.Lax, and I.Shafarevich. It was said at the time that there were three quantum prizewinners (Drinfeld for the introduction of quantum groups) and one mathematician Mori.

Physics had arrived in mathematics.... but not without some controversy.

8 The Jaffe-Quinn article

In 1993 the Bulletin of the American Mathematical Society published an article by Arthur Jaffe and Frank Quinn [40] which began

> Is speculative mathematics dangerous? Recent interactions between physics and mathematics pose the question with some force: traditional mathematical norms discourage speculation, but it is the fabric of theoretical physics.

The paper provoked an avalanche of responses, published the following year. Michael Atiyah, writing from the Master's Lodge in Trinity College Cambridge (which, together with being President of the Royal Society took away most of his research time) wrote [17]:

> I find myself agreeing with much of the detail of the Jaffe-Quinn argument, especially the importance of distinguishing between results based on rigorous proofs and those which have a heuristic basis. Overall, however, I rebel against their general tone and attitude which appears too authoritarian.
>
> My fundamental objection is that Jaffe and Quinn present a sanitized view of mathematics which condemns the subject to an arthritic old age. ...The history of mathematics is full of instances of happy inspiration triumphing over a lack of rigour..... The marvelous formulae emerging at present from heuristic physical arguments are the modern counterparts of Euler and Ramanujan, and they should be accepted in the same spirit of gratitude tempered with caution.

Jaffe and Quinn isolated in their view four major problems,

- theoretical (or speculative) work, if taken too far, goes astray because it lacks the feedback and corrections provided by rigorous proof
- further work is discouraged and confused by uncertainty about which parts are reliable
- a dead area is often created when full credit is claimed by vigorous theorizers: there is little incentive for cleaning up the debris that blocks further progress
- students and young researchers are misled

and made certain specific proposals:

- theoretical work should be explicitly acknowledged as theoretical and incomplete; in particular, a major share of credit for the final result must be reserved for the rigorous work that validates it
- within a paper, standard nomenclature should prevail: in theoretical material, a word like "conjecture" should replace "theorem"; a word like "predict" should replace "show" or "construct"; and expressions such as "motivation" or "supporting argument" should replace "proof". Ideally the title and abstract should contain a word like "theoretical", "speculative", or "conjectural".

The article was probably deliberately provocative, but engendered various responses, such as

> A case in point, in my experience, was E. Cartan's work on exterior differential forms and connections .... Personally, I felt rather comfortable with it but later, after having been exposed to the present points of view, could hardly understand what I had thought to understand.. .... But I do believe in the self-correcting power of mathematics (*A.Borel*)
>
> My main objection to JQ is that, in their search for credit for some individuals at the expense of others they consider rogues, they propose to set up a police state within Charles (River) mathematics, and a world cop beyond its borders. (*B.Mandelbrot*)
>
> There is no doubt that a mathematician that gave a rigorous proof of a statement heuristically proved by another mathematician or physicist deserves essential credit. However I don't think that one can say a priori that "a major share of credit for the final result must be reserved for the rigorous work." (*A.Schwartz*)

The article raised valid points but Atiyah remained unconvinced:

> What is unusual about the current interaction is that it involves front-line ideas both in theoretical physics and in geometry. This greatly increases its interest to both parties, but Jaffe-Quinn want to emphasize the dangers. They point out that geometers are inexperienced in dealing with physicists and are perhaps being led astray. I think most geometers find this attitude a little patronizing: we feel we are perfectly capable of defending our virtue.

Galileo could well have joined in the discussions [33]:

> I hear my adversaries shouting in my ears that it is one thing to deal with matters physically, and quite another to do so mathematically, and that geometers should stick to their fantasies and not get entangled in philosophical matters...

9 Geometry and particles

While Michael Atiyah was happy urging physicists to exploit quantum field theories to formulate new questions in mathematics or to impart order into physically-influenced results, he sometimes expressed doubts about quantum theory as a final description of the universe – it was too linear. If we look at the non-expository papers he wrote, then many of them feature, in different forms, particle-like objects. I think he felt that, though they were embedded in the behaviour of solutions to nonlinear equations, nonlinear localized properties revealed more of the physical essence. In this last section I will try and deal with some of them. In his later years he adhered to an unorthodox view of physics when he changed, in the words of Bernd Schroers, [50] from an "inadvertent physicist" to an "intentional one". This focus on particles then became more pronounced but here I confine myself to a few of the earlier examples.

9.1 Instanton moduli spaces

The ADHM construction, despite its explicitness, seemingly gave no method for determining the topology of the moduli space. In 1978, together with J.D.S.Jones, Atiyah used instead the pre-ADHM examples of instantons due to 't Hooft to obtain homological information about the moduli space M_k of instantons of charge k framed at infinity. The framing increases the dimension of the space by the dimension of the group and so for $SU(2)$ this is an $8k$-dimensional noncompact manifold.

Recall that these initial solutions were defined by k distinct points in $\mathbf{R}^4$ together with a positive scale for each point. The configuration space $C_k(\mathbf{R}^4)$ of such points carries non-trivial topology. This was examined by Graeme Segal who in 1973 defined a map from $C_k(\mathbf{R}^n)$ to the space $\Omega^n_k(S^n)$ of based maps of degree k from S^n to itself and showed that this gave an isomorphism in degree q homology if k is sufficiently large. In our case the map can be considered as the field generated by unit charges at each point: $\mathrm{grad}\, 1/r^2$ maps ∞ to the origin and the origin to infinity which is a map from $\mathbf{R}^4 \cup \{\infty\} \cong S^4$ to itself.

The infinite-dimensional space of all framed $SU(2)$ connections on S^4 modulo gauge equivalence is shown to be homotopically equivalent to $\Omega^3(SU(2)) = \Omega^3(S^3)$ by comparing the radial gauge on $\mathbf{R}^4$ with the framing at infinity. This has components $\Omega^3_k(S^3)$ indexed by k the degree of the map, or equivalently the instanton charge. There is therefore a map from M_k to $\Omega^3_k(S^3)$. But pulling back the 1-instanton by a map $f : S^4 \to S^4$ gives a map from $\Omega^4(S^4)$ to the moduli space of connections. It is then shown that, up to homotopy, the map from the configuration space to $\Omega^4_k(S^4)$ composed with the map to $\Omega^3_k(S^3)$ factors through M_k and one produces non-trivial homology classes in M_k. The authors left the paper with a series of conjectures which were finally resolved in [27] in 1992. The end result is that the natural inclusion $M_k \subset \Omega^3_k(SU(2))$ is a homotopy equivalence through dimension $q = [k/2] - 2$.

By that time the interest of physicists in these spaces had faded somewhat. The general feature that, up to a certain level, the homology of these moduli spaces of absolute minima captures the homology of the infinite-dimensional spaces on which the functional is defined, holds in many situations, from the Riemann surface case (a consequence of Atiyah and Bott's calculations), through instantons to the theory of harmonic maps which had now acquired the physicist's notation of the nonlinear sigma model. In the Riemann surface case, the "correction terms" are specifically carried by the higher critical points and although the existence of such non-minimal solutions to the Yang-Mills equations on $\mathbf{R}^4$ is well-established, the picture is not so clear. Nevertheless from this approach it is clear that the "pseudoparticle" interpretation of instantons captures a large amount of the topology of the moduli space.

9.2 Magnetic monopoles

Following closely on the heels of the instanton problem came the question of magnetic monopoles. These are static configurations defined on $\mathbf{R}^3$ and consisting of a connection A and an associated *Higgs field*, a section ϕ of the bundle of Lie algebras. The equations they satisfy, the Bogomolny equations, are

$$F_A = *\nabla_A \phi.$$

The physical origin arises from the Yang-Mills-Higgs functional

$$\int_{\mathbf{R}^3} |F_A|^2 + |\nabla_A \phi|^2 + \frac{\lambda}{4}(1 - |\phi|^2)^2.$$

The boundary conditions for this integral to exist include $|\phi| \to 1$ as $|x| \to \infty$ and the Bogomolny equations describe the absolute minimum when $\lambda = 0$ but keeping the condition $|\phi| \to 1$. In the case of an $SU(2)$ connection, this implies that on a large sphere in $\mathbf{R}^3$ of radius R the eigenspaces of ϕ define line bundles of degree $\pm k$ and here k is the magnetic charge. The charge 1 monopole, centred at the origin was given in 1975 by Prasad and Sommerfield:

$$A_i = -i\epsilon_{ijk}\sigma_j x_k \frac{\sinh r - r}{r^2 \sinh r} \qquad \phi = i x_j \sigma_j \frac{r\cosh r - \sinh r}{r^2 \sinh r}$$

The Yang-Mills density $|F_A|^2$ for this solution is concentrated around the origin and suggests a nonlinear particle-like object. In 1980, Jaffe and Taubes [41] produced by analytic means an existence theorem for solutions representing "widely spaced monopoles" where the energy density is approximately localized around k points in $\mathbf{R}^3$, further supporting the particle-like interpretation. Since the Higgs field for the 1-monopole vanishes at the origin one might also track the "locations" by the zeros of the Higgs field.

I began working on monopoles in the early 1980s using twistor methods. Here there was a more direct interpretation of the twistor space as the space of oriented straight lines in $\mathbf{R}^3$. As a complex surface the twistor space is the total space of the line bundle $\mathcal{O}(2)$ over $\mathbf{C}\mathrm{P}^1$. Holomorphic sections are quadratic in ζ, $\eta = a + b\zeta + c\zeta^2$, so that $(a, b, c) \in \mathbf{C}^3$ parametrizes these twistor lines and is the complexification of $\mathbf{R}^3$. The null cone through a point is the space of sections tangential to the corresponding line: for the zero section this means $a + b\zeta + c\zeta^2$ has a double zero, or $b^2 - 4ac = 0$.

The real points for Euclidean space are $(a, b, c) = (x_1 + ix_2, 2ix_3, x_1 - ix_2)$ giving the Euclidean metric quadratic form $(4ac - b^2) = 4(x_1^2 + x_2^2 + x_3^2)$. The sections passing through the point (η, ζ) in twistor space satisfy the equation

$$\eta = (x_1 + ix_2) + 2ix_3\zeta + (x_1 - ix_2)\zeta^2$$

which, taking the real and imaginary parts, gives linear equations for (x_1, x_2, x_3) – two planes in $\mathbf{R}^3$ which intersect in a line.

A solution of the Bogomolny equations now corresponds to a holomorphic vector bundle on this minitwistor space, but there is no compactification to an algebraic surface which works here – the transcendental nature of the 1-monopole solution suggests this. Nevertheless algebraic geometry comes to the aid of finding solutions. My own approach involved solving the ODE $\nabla s + i\phi s = 0$ along the straight lines in $\mathbf{R}^3$. The lines which admitted an L^2 solution formed an algebraic curve of genus $(k-1)^2$ – the spectral curve – which is subject to a transcendental constraint.

Then in 1981 a preprint from CERN by the physicist Werner Nahm [45] gave an alternative construction. This involved three $k \times k$ matrices $T_i(t)$ satisfying the equations

$$\frac{dT_1}{dt} = [T_2, T_3], \quad \frac{dT_2}{dt} = [T_3, T_1], \quad \frac{dT_3}{dt} = [T_1, T_2].$$

In the preprint Nahm shows how to solve these for $k = 2$ by using elliptic functions which suggests a link with the spectral curve and indeed writing $T = (T_1 + iT_2) + 2iT_3\zeta + (T_1 - iT_2)\zeta^2$ the equations become

$$\frac{dT}{dt} = [iT_3 + (T_1 - iT_2)\zeta, T]$$

so that $\det(\eta - T)$ is independent of t and is a polynomial $p(\eta, \zeta)$. Its zero set is the spectral curve in general.

The matrices are defined from the monopole by taking an orthonormal basis ψ_α of the k-dimensional space of L^2 eigenfunctions with eigenvalue t of the Dirac equation coupled to the monopole and defining the $k \times k$ matrix

$$T_{\alpha\beta} = \int_{\mathbf{R}^3} (x_i \psi_\alpha, \psi_\beta).$$

Though Atiyah was following these developments at the time he was more involved with the issues discussed in the previous sections, although in 1987 he wrote about the parallel version on hyperbolic 3-space. He did however begin to have discussions concerning monopoles with Nick Manton, a theoretical physicist in Cambridge. Manton had written [43] about conjectured forces between well-separated monopoles, but also showed how the classical dynamics of slowly moving monopoles should be well approximated by geodesic motion on the moduli space of static monopoles [44]. To put this into effect required a knowledge of the natural Riemannian metric on the moduli space.

The academic year 1983-84 I spent on sabbatical in Stony Brook and there with physicists Roček, Lindstrom and Karlhede we developed the hyperkähler quotient construction. A hyperkähler metric has three symplectic forms $\omega_1, \omega_2, \omega_3$ which are Kähler forms for complex structures I, J, K which satisfy the algebraic condition of quaternions. Until the work of Gibbons and Hawking in the late 1970s there were hardly any explicit examples but the quotient construction, adapted from the standard symplectic quotient, meant that any quaternionic representation yielded in principal an example.

When I returned to Oxford Michael immediately pointed out that the Bogomolny equations had an interpretation as the zero set of an infinite-dimensional hyperkähler moment map and as a consequence the moduli space had a hyperkähler metric – one complex structure for each direction in $\mathbf{R}^3$. Moreover he wanted the metric for two monopoles in order to test Manton's ideas about monopole dynamics. The formalism for implementing the hyperkähler moment map produces an extra circle factor – the hyperkähler manifold has dimension $4k$ instead of $4k-1$ for the genuine moduli space.

For $k = 1$ this space is $\mathbf{R}^3$ given by a centre for the monopole, but Atiyah argued that we should really think of the 4-manifold $S^1 \times \mathbf{R}^3$ as describing a location and an internal $U(1)$ phase.

The charge 2 moduli space is 8-dimensional but a centred version is 4-dimensional with the rotation group acting isometrically. We determined the metric, which depended on complete elliptic integrals, as one might have expected from the fact that the spectral curves were elliptic, and calculated some geodesics. By chance, a seminar Atiyah was giving was attended by members of IBM's UK research laboratory in Southampton and as a test of their parallel processors they proposed making a movie of the monopole scattering. It can be viewed on my home page [38].

In its preparation, I observed Atiyah as an experimental scientist and it revealed one other difference between mathematicians and physicists – an appreciation of scale. For the monopole dynamics it took a great deal of iteration to get any indication of the nonlinearity of the problem. The first "takes" just showed linear motion with elastic scattering in various directions and it was a while before, by adjusting the scale, one could see from the changing distribution of the Yang-Mills density the nonlinear effect of two monopoles colliding and scattering.

Atiyah suggested after this work that we should write a book based on his Porter Lectures [13] and it was here that the particle picture of monopoles entered in a rather novel manner. We needed to describe the metric on the moduli space of charge k monopoles but explicit formulas like the $k = 2$ situation were out of reach. Hyperkähler manifolds however have twistor spaces which are complex manifolds fibering over $\mathbf{C}\mathrm{P}^1$ with a (twisted) holomorphic symplectic form along the fibres. In the case of monopoles, given a direction in $\mathbf{R}^3$, i.e. a point $(u_1, u_2, u_3) \in S^2 \cong \mathbf{C}\mathrm{P}^1$ the fibre over u is the moduli space with complex structure $u_1 I + u_2 J + u_3 K$. The rotation group $SO(3)$ is an isometry and takes one complex structure to another so the fibres are all holomorphically equivalent. Donaldson in fact had identified each as a space of rational maps as follows.

In [31] Donaldson took Nahm's description of monopoles and, for the complex structure in the x_1 direction, broke the equations into a complex one and a real one, introducing a fourth matrix T_0 which can be removed eventually by a gauge transformation. With skew-adjoint matrices T_i one sets $\alpha = (T_0 + iT_1)/2, \beta = (T_2 + iT_3)/2$ and the equations are

$$\frac{d\beta}{dt} + 2[\alpha, \beta] = 0 \qquad \frac{d}{ds}(\alpha + \alpha^*) + 2([\alpha, \alpha^*] + [\beta, \beta^*]) = 0.$$

Acting on solutions to the complex equation by complex gauge transformations is the analogue of the moment map/stability situation for holomorphic vector bundles which began with Atiyah and Bott. Given the boundary conditions from Nahm, Donaldson describes the "stability condition" on (α, β) for there to exist a solution. The end result is a classification as pairs (B, v) where B is a complex symmetric $k \times k$ matrix ($B = \beta(1)$ at the midpoint $t = 1$ of the t-interval) and $v \in \mathbf{C}^k$ a cyclic vector, up to the action of a complex orthogonal matrix. The rational map is then

$$f(z) = w^T(z - B)^{-1}w$$

taking $z = \infty$ to 0, so is of the form

$$f(z) = \frac{p(z)}{q(z)} = \frac{a_0 + a_1 z + \ldots + a_{k-1}z^{k-1}}{z^k + b_1 z^{k-1} + \ldots + b_k}$$

The charge 1 monopole is

$$f(z) = \frac{a}{z + b}$$

where $(a, b) \in \mathbf{C}^* \times \mathbf{C}$. We can take a rational map

$$f(z) = \sum_{i=1}^{k} \frac{\alpha_i}{z - \beta_i} = \frac{p(z)}{q(z)}$$

and think of this as an approximate superposition of single monopoles located at $(-\log|\alpha_i|, \beta_i)$ and phase $\alpha_i/|\alpha_i|$, though this description is of course direction-dependent.

The β_i above are the zeros of the denominator $q(z)$. In general if $q(\beta_i) = 0$, then $p(\beta_i) \neq 0$ since the degree of the map f is k, so $f(z)$ defines an *unordered* sequence of points $((p(\beta_1), \beta_1), \ldots, (p(\beta_k), \beta_k)) \in (\mathbf{C}^* \times \mathbf{C})^k$, i.e. in the quotient by the symmetric group Σ_k. The symmetric product of k surface factors is singular but the space of rational maps is smooth (the complement of the resultant equation $R(p, q) = 0$ in $\mathbf{C}^k \times \mathbf{C}^k$) so the space of rational maps is a smooth resolution of the singularities of the symmetric product. Or put another way, the monopole moduli space smoothes out the singularities that particles with $U(1)$-charges acquire under collision.

In algebraic geometry the standard way to resolve the singularities of the symmetric product of a complex surface X is the Hilbert scheme $X^{[k]}$, the space of ideal sheaves $\mathcal{I} \subset \mathcal{O}_X$ such that $\dim \mathcal{O}_X/\mathcal{I} = k$. Moreover if X has a holomorphic symplectic form then the natural form on the product X^k extends to a symplectic form on the Hilbert scheme. The space of rational maps contains no projective spaces as the Hilbert scheme does, but Atiyah in Chapter 6 of [13] introduces the notion of *transverse Hilbert scheme*. This applies to a surface X with a projection $\pi : X \to \mathbf{C}$. The transverse Hilbert scheme is the subspace of points $D \in X^{[k]}$ such that $\pi(D)$ is an isomorphism onto its scheme-theoretic image. Surprisingly, according to Bielawski

[26] this is smooth for all dimensions of X, though the Hilbert scheme itself is not for $\dim X > 2$.

To describe the twistor space now, one takes the twistor space for the flat metric on $S^1 \times \mathbf{R}^3$ (the quotient of $\mathcal{O}(1) \oplus \mathcal{O}(1)$ by the action of $\mathbf{Z}$ generated by the additive action of $(\zeta, 1)$) and applies the fibrewise transverse Hilbert scheme, a more canonical description than appealing to rational maps.

The symplectic form coming from $dz/z \wedge dw$ on $\mathbf{C}^* \times \mathbf{C}$ is

$$\omega = \sum_{i=1}^{k} \frac{dp(\beta_i)}{p(\beta_i)} \wedge d\beta_i$$

and this we know extends to the Hilbert scheme. It can be seen more invariantly as follows [34].

From the expression above one can see that functions of the coefficients of p Poisson commute as do functions of the coefficients of q. A point z defines the function $p(z)$ of the numerator p and $q(z)$ of the denominator in p/q. Then the following formula gives the Poisson bracket even when there are multiple zeros of q:

$$\{p(z), q(w)\} = \frac{p(z)q(w) - q(z)p(w)}{z - w}$$

This is classically the Bezoutian.

9.3 Skyrmions

Atiyah's contact with Manton led to another field of exploration of the particle-like world – the Skyrme model. Mathematically it involves maps $f : \mathbf{R}^3 \to SU(2)$ and minima of a certain functional.

We have the derivative $df : T\mathbf{R}^3 \mapsto TSU(2)$ which is a 1-form on $\mathbf{R}^3$ with values in the vector bundle $f^*TSU(2)$, and the usual harmonic map functional is the L^2 norm square of this. For the Skyrme functional one introduces $\Lambda^2 df : \Lambda^2 T\mathbf{R}^3 \to \Lambda^2 TSU(2)$ which is a 2-form with values in $f^*\Lambda^2 TSU(2) \cong f^*TSU(2)$ and the Skyrme energy is

$$E = \int_{\mathbf{R}^3} c_1|df|^2 + c_2|\Lambda^2 df|^2 \tag{4}$$

for constants c_1, c_2. In this case there is a topological invariant, the degree k of f and $E \geq const.\sqrt{c_1c_2}k$. The problem is to find the maps f which are critical for this functional.

In two dimensions, the standard harmonic map functional behaves well and in a parallel development with Yang-Mills theory the physicists's nonlinear sigma model gave rise to a large mathematical literature giving explicit constructions where the target manifold is complex projective space or a Lie group, but the three-dimensional theory is less amenable – for example put $c_2 = 0$ and there is no positive lower bound on E. The situation may be compared to pure Yang-Mills in two dimensions as in Atiyah and Bott's work and the monopole equations in three dimensions: the 3-dimensional Yang-Mills-Higgs functional

$$\int_{\mathbf{R}^3} |F_A|^2 + |\nabla_A \phi|^2$$

with its 1-form and 2-form terms is comparable to (4).

For the physicists the Skyrme theory is a model for baryons and k is called the baryon number. In the paper [2], where Witten is a coauthor, here is the motivation:

> Since the proper large-N effective theory is unknown, we will consider here a crude description in which the large-N theory is assumed to be a theory of pions only. In this context, it is necessary to add a non-minimal term to the non-linear sigma model to prevent the solitons from shrinking to zero-size..... .... Although the Skyrme model is only a rough description, since it omits the other mesons and interactions that are present in the large-N limit of QCD, we regard it as a good model for testing the reasonableness of a soliton description of nucleons.

In the same paper the authors reduce the determination of the spherically symmetric $k = 1$ solution to an ODE and solve it numerically. The obvious Ansatz for f is

$$f(x) = \exp(\phi(r)(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)/r)$$

where $\sigma_1, \sigma_2, \sigma_3$ is a basis for the Lie algebra of $SU(2)$. It is straightforward then to derive the variational equation for f, and the differential equation for ϕ, from the functional.

One entry point for Atiyah derived from the work on monopoles. He was interested in monopoles which were symmetric with respect to a finite group of rotations and in particular how the energy density was distributed, reflecting the symmetry. He was assisted in this by Paul Sutcliffe who developed several methods for numerically solving monopole equations from the Nahm point of view and graphically displaying the energy densities. When applied to skyrmions there was a consistency in the shapes which suggested that the behaviour of soliton-like objects in three dimensions was largely model-independent. This was perhaps the reason why, with Manton, he sought an alternative approach.

The idea in [14] is to use instantons in $\mathbf{R}^4$ to generate maps $f : \mathbf{R}^3 \to SU(2)$ by decomposing $\mathbf{R}^4 = \mathbf{R}^3 \times \mathbf{R}$ and taking $f(x)$ to be the holonomy of the connection along the line (x, t) from $t = -\infty$ to $t = +\infty$. This provides a finite-dimensional family of maps and one looks to minimize the functional on this family rather than the full function space. The advantage is that the baryon number corresponds to the instanton charge. In principle one could use the full ADHM family but Atiyah chose to use the familiar pseudoparticle solutions.

The charge one instanton gives $\phi(r) = \pi(1 - (1 + \lambda^2/r^2)^{-1/2})$ and the minimum is determined numerically with $\lambda^2 = 2.11$. The paper [16] follows the idea of the dynamics of two monopoles to apply to two skyrmions but now there is no moduli space of absolute minima and it is replaced by a manifold of curves of steepest descent. In this situation there is a potential as well as a kinetic energy term.

9.4 The Berry-Robbins problem

Emerging from years of administration as President of the Royal Society and Master of Trinity College Cambridge, at the turn of the century Michael latched on to a problem suggested to him by Michael Berry [25]. The question was simple – is there a natural map

$$C_n(\mathbf{R}^3) \to U(n)/T$$

which commutes with the action of the symmetric group Σ_n?

Here $C_n(\mathbf{R}^3)$ is, as before, the configuration space of ordered n-tuples of distinct points $x_1, \ldots, x_n \in \mathbf{R}^3$ and $U(n)/T$ is the flag manifold, the quotient space of the unitary group by its diagonal elements T. The symmetric group acts on the left hand side by changing the order and on the right as the Weyl group $N(T)/T$. Part of the appeal was the association of "particles" in $\mathbf{R}^3$ with complex vectors in $\mathbf{C}^n$, a shadow of quantization, yet it was a purely mathematical question which led Atiyah into various different fields.

In [18] he gave an elementary example of such a continuous map and used this later to discuss its action in homotopy, in particular $SO(3)$-equivariant cohomology. But he disliked the construction, presumably as not "a proof consistent with the elegance of the problem" as he would often comment. He offered a different one but which was never fully established:

For each pair of points one associates the direction of $x_j - x_i$ as a point $t_{ij} \in S^2$ identified as $\mathbf{CP}^1$ (this is the celestial sphere that Penrose regarded as the key component of twistor theory). Then define p_i to be the polynomial of degree $(n-1)$ with roots t_{ij} $j \neq i$. The space of polynomials of this degree is n-dimensional and if the p_i are linearly independent they define a flag

$$\langle p_1 \rangle \subset \langle p_1, p_2 \rangle \subset \langle p_1, p_2, p_3 \rangle \ldots .$$

This map not only commutes with the symmetric group, but it is also invariant under translations in $\mathbf{R}^3$ and equivariant under rotations where $SO(3)$ acts on the projective space $\mathbf{C}\mathrm{P}^{n-1}$ via the n-dimensional irreducible representation of its covering $SU(2)$. The problem was proving linear independence.

The ad hoc continuous map follows the same pattern but instead one fixes an origin and identifies any sphere centred at the origin with the unit sphere of directions. Then for each i, one takes a polynomial whose roots are defined by $x_j/|x_j|$ if x_j is outside the sphere $|x| = |x_i|$ and if x_j is inside the root is $y_j/|y_j|$ where $|y_j| = |x_i|$ and y_j lies on the line joining x_i to x_j. Then linear independence is proved by induction.

Atiyah was unhappy with his solution, in particular because it was not translation-invariant and this inhibited his interest in "clusters" of well-separated points analogous to his earlier consideration of monopoles. He pursued the more invariant conjectural solution together with Sutcliffe [20] focusing on a determinant whose nonvanishing would give linear independence of the polynomials. This becomes a function V on the configuration space which is conjecturally greater than or equal to one. Numerical evaluation of V-minimizing polyhedra gave graphics which were reminiscent of the numerical work on monopoles or skyrmions, and proofs for small n provided the opportunity for some elementary geometry which Atiyah would delight in introducing in his papers when appropriate.

The best solution (in the author's opinion) came from using Nahm's equations, with their origins in gauge theory [19]. There was also a more general setting replacing $U(n)$ by a simple Lie group G and its maximal torus. Instead of the configuration space one takes the Cartan subalgebra $\mathfrak{h}$. tensors with $\mathbf{R}^3$ and removes the kernels Δ of the root homomorphisms $\alpha \otimes 1 : \mathfrak{h} \otimes \mathbf{R}^3 \to \mathbf{R}^3$. The theorem then is that there is a map

$$\mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta \to G/T$$

which commutes with the action of the Weyl group. The solution is also equivariant with respect to $SU(2)$ acting as $SO(3)$ on the second factor in the left hand side and the principal three-dimensional subgroup of G on the right hand side.

Nahm's equations are

$$\frac{dT_1}{dt} = [T_2, T_3], \quad \frac{dT_2}{dt} = [T_3, T_1], \quad \frac{dT_3}{dt} = [T_1, T_2].$$

and clearly make sense if T_i take values in any Lie algebra. If they acquire a simple pole, at $t = 0$ say, then $T_i = R_i/t + \dots$ and the equations give $R_1 = -[R_2, R_3]$ etc. which is a homomorphism from the Lie algebra of $SU(2)$ to $\mathfrak{g}$.

For the original application to $SU(2)$ monopoles there is a pole at each end of a finite interval $(0, 2)$. In this case, following work of Kronheimer, one takes solutions on $(0, \infty)$ with a pole at $t = 0$ given by the principal three-dimensional subgroup and as $t \to \infty$ the T_i approach a regular commuting triple in $\mathfrak{g}$. Fixing the subgroup $\rho : SU(2) \to G$, an existence theorem shows that there is a manifold N' of solutions with these boundary conditions and hence a map $N' \to \mathfrak{g} \otimes \mathbf{R}^3$ giving the asymptotic value τ.

Three commuting regular elements lie in a Cartan subalgebra so the orbit $G\tau$ meets a fixed $\mathfrak{h} \otimes \mathbf{R}^3$ in an orbit of the Weyl group W. This gives a map $N' \to (\mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta)/W$ and a W-covering N maps $N \to \mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta$. Fixing $\tau \in \mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta$ identifies $G\tau$ with G/T and so there is a map $N \to G/T$.

All of this works for any three-dimensional subgroup but for the principal one the map $N \to \mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta$. is an isomorphism and thus gives the required map

$$\phi : \mathfrak{h} \otimes \mathbf{R}^3 \backslash \Delta \to G/T.$$

The paper [19] makes interesting speculations about Hecke algebras and Kazhdan-Lusztig theory which goes far beyond the "particle" aspects.

10 Conclusion

This article by no means exhausts Michael Atiyah's interactions with physics. In particular the students he acquired on his return to Oxford from Princeton followed his physics interests from the mid 1970s onwards: there was Simon Donaldson of course with the applications of gauge theory to 4-manifold topology, Michael Murray describing monopoles for a general simple group, Peter Kronheimer producing a construction and classification of ALE gravitational instantons and Lisa Jeffrey providing a mathematically rigorous proof of results on the asymptotics of the three-manifold invariants of Witten and Reshetikhin and Turaev. Also Ruth Lawrence's thesis on braid group representations was motivated by the work of Jones and Witten.

In later years he wrote a long paper with Witten [21] on G_2-manifolds and, in response to the physicist's notion of D-brane charges, wrote one on twisted K-theory with Graeme Segal [22]. Even a paper as mathematical as [19] on the Dedekind η-function contains the remark

... Hirzebruch notes the appearance of Dedekind sums in number theory.. and topology. He asks whether there is some deep explanation for this fact. Hopefully this paper, following Hirzebruch's work on cusps, provides an answer by showing that the real connection between number theory and topology, in this context, hinges on fundamental ideas from the physics of gauge theories!

Atiyah's mathematical garden was certainly reinvigorated by the exotic specimens physicists provided.

References

[1] L. Abrams, *Two-dimensional topological quantum field theories and Frobenius algebras*, J. Knot Theory Ramifications **5** (1996) 569–587.

[2] G.Adkins, C.Nappi & E.Witten, *Static properties of nucleons in the Skyrme model*, Nucl. Phys. **B228** (1983) 552–566.

[3] L.Alvarez-Gaumé, *Supersymmetry and the Atiyah-Singer index theorem*, Commun. Math. Phys. **90** (1983) 161–173.

[4] M.F.Atiyah & I.M.Singer, *The index of elliptic operators IV*, Ann. of Math. **93** (1971) 119–138.

[5] M.F.Atiyah & R.S.Ward, *Instantons and algebraic geometry*, Commun. Math. Phys. **55** (1977) 117–124.

[6] M.F. Atiyah, V.G. Drinfeld, N.J.Hitchin & Y.I. Manin, *Construction of instantons*, Phys. Lett. **A 65** (1978), 185–187.

[7] M.F.Atiyah, *Green's functions for self-dual four-manifolds*, Advances in Mathematics Supplementary Studies **7 A** (1981) 129–158.

[8] M.F.Atiyah, *Convexity and commuting Hamiltonians*, Bull. Lond. Math Soc. **14** (1982) 1–15.

[9] M.F.Atiyah, *Angular momentum, convex polyhedra and algebraic geometry*, Proc. Edin. Math Soc. **26** (1983) 121–138.

[10] M.F.Atiyah & I.M.Singer, *Dirac operators coupled to vector potentials*, Proc. Natl. Acad. Sci. USA, **81** (1984) 2597–2600.

[11] M.F.Atiyah, *The logarithm of the Dedekind η-function*, Math.Ann. **278** (1987) 335–380.

[12] M.F.Atiyah, *Topological quantum field theory,* Publ. math. IHES **68** (1988) 175–186.

[13] M.F.Atiyah & N.Hitchin, "The geometry and dynamics of magnetic monopoles", Princeton Univ. Press (1988).

[14] M.F.Atiyah & N.S.Manton, *Skyrmions from instantons*, Phys. Lett. **B222** (1989) 438–442.

[15] M.F.Atiyah, "The geometry and physics of knots", Cambridge Univ. Press (1990)

[16] M.F.Atiyah & N.S.Manton, *Geometry and kinematics of two skyrmions*, Commun. Math. Phys. **152** (1993) 391–422.

[17] M.F.Atiyah et al, *Responses to "Theoretical Mathematics": towards a cultural synthesis of mathematics and theoretical physics, by A.Jaffe and F.Quinn*, Bull AMS **30** (1994) 178–207.

[18] M.F.Atiyah, *The geometry of classical particles,* Surveys in Differential Geometry **7** (2001) 1–15.

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

[20] M.F.Atiyah, *The geometry of point particles,* Proc. R. Soc. London A **458** (2002) 1089–1115.

[21] M.F.Atiyah & E.Witten *M-theory dynamics on a manifold of G_2 holonomy*, Adv. Theor. Math. Phys. **6** (2003) 1 – 106.

[22] M.F.Atiyah & G.Segal, *Twisted K-theory*, Ukr. Math. Bull. **1** (2004) 291-334.

[23] M.F.Atiyah, Collected Works Vol 6, Oxford University Press (2004) 11.

[24] M.F.Atiyah, *Geometry and Physics of the 20th Century,* in "Géometrie au XXe siècle,1930-2000: Histoire et horizons" , J.Kouneiher et al (eds.), Hermann, Paris (2005) 4–9. Collected Works Vol 7 259–264.

[25] M.Berry,& J.Robbins, *Indistinguishability for quantum particles: spin, statistics and the geometric phase*, Proc. R. Soc. London A **453** (1997) 1771–1790.

[26] R.Bielawski, *Transverse Hilbert schemes, bi-Hamiltonian systems, and hyperkähler geometry*, arXiv:2001.05669.

[27] C.Boyer, J.Hurtubise, B.Mann & J.Milgram (1993), *The topology of instanton moduli spaces. I. The Atiyah –Jones conjecture*, Ann. of Math. **137** (1993) 561-609.

[28] E.F. Corrigan, D.B. Fairlie, S. Templeton & P. Goddard, *A Green function for the general self-dual gauge field*. Nucl. Phys. **B 140** (1978) 31–44.

[29] J.-P. Demailly *Champs magnétiques et inégalités de Morse pour la d" - cohomologie*, Ann. Inst. Fourier **35** (1985) 189-229.

[30] S.K.Donaldson, *A new proof of a theorem of Narasimhan and Seshadri*, J. Differential Geom. **18** (1983) 269 – 277.

[31] S.K.Donaldson, *Nahm's equations and the classification of monopoles*, Commun. Math. Phys. **96** (1984) 387–407.

[32] S.K.Donaldson, *Anti-self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles*, Proc. Lond. Math. Soc. **50** (1985) 1– 26.

[33] S.Drake, "Galileo at work", Dover (1978) 172–173.

[34] L.Faybusovich & M.Gekhtman, *Poisson brackets on rational functions and multi-Hamiltonian structure from integrable lattices*, Phys. Lett. A **272** (2000) 236–244.

[35] A. Floer, *Witten's complex and infinite-dimensional Morse theory*, J. Differential Geom. **30** (1989), 207-221.

[36] E.Getzler, *Supersymmetry and the Atiyah-Singer index theorem*, Commun. Math. Phys. **90** (1983) 161–173.

[37] G.Harder & M.S.Narasimhan, *On the cohomology groups of moduli spaces of vector bundles over curves*, Math. Annalen. **212** (1975) 215–248.

[38] Nigel Hitchin homepage: https://people.maths.ox.ac.uk/hitchin/

[39] R. Jackiw, C. Nohl, & C. Rebbi, *Conformal properties of pseudoparticle configurations*, Phys. Rev. **D 15** (1977) 1642 1646.

[40] A.Jaffe & F.Quinn, *"Theoretical Mathematics": towards a cultural synthesis of mathematics and theoretical physics*, Bull. AMS **29** (1993) 1–13.

[41] A. Jaffe & C. Taubes, "Vortices and monopoles", Boston, Birkhäuser, (1980).

[42] F.C.Kirwan, *Moment maps and convexity: memories of Michael Atiyah*, Notices of the AMS **66** (2019) 540–567.

[43] N.S.Manton, *The force between 't Hooft-Polyakov monopoles*, Nucl. Phys. **B126** (1977) 525–541.

[44] N.S.Manton, *A remark on the scattering of BPS monopoles*, Phys. Lett. **110B** (1982) 54–56.

[45] W. Nahm, *All self-dual monopoles for arbitrary gauge groups*, CERN preprint TH.3172-CERN (1981).

[46] M.S.Narasimhan & C.S.Seshadri, *Stable and unitary vector bundles on a compact Riemann surface*, Ann. of Math. **82** (1965) 540–567.

[47] R.Penrose, *Twistor algebra*, J. Math. Phys. **8** (1967) 345–366.

[48] R.Penrose, *Twistor functions and sheaf cohomology*, Twistor Newsletter **2** 10th June (1976).

[49] R.Penrose, *Solutions of the zero rest mass equations*, J.Math.Phys. **10** 38 (1969).

[50] B.Schroers, *Michael Atiyah and Physics: the later years*, Notices of the AMS **66** (2019) 1849–1851.

[51] G.B.Segal, *The definition of conformal field theory* in " Topology, geometry and quantum field theory," U.Tillmann (ed), London Math. Soc. Lecture Note Ser., **308**, Cambridge Univ. Press, Cambridge, (2004) 421-577

[52] S.Shatz, *The decomposition and specialization of algebraic families of vector bundles*, Compositio math. **35** (1977) 163 – 187.

[53] E.Witten, *Michael Atiyah and physics*, Notices of the AMS **66** (2019) 1837–1839.

[54] K.Uhlenbeck & S-T Yau, *On the existence of Hermitean Yang-Mills-connections on stable bundles over Kähler manifolds*, Comm. Pure Appl. Math. **39** (1986) 257–293

[55] E.Witten, *Some exact multipseudoparticle solutions of classical Yang-Mills theory*, Phys. Rev. Lett. **38** (1977) 121–124.

[56] E.Witten, *Supersymmetry and Morse theory,* J.Differential Geom. **17** (1982) 661–692.

[57] E. Witten, *Holomorphic Morse inequalities*, In “Algebraic and differential topology – global differential geometry”, volume 70, Teubner-Texte Math., 318- 333. Teubner, Leipzig, (1984).