Gauge Theory Is Dead!—Long Live Gauge Theory!

D. Kotschick

In late October and early November 1994 many mathematicians received e-mail from colleagues trumpeting the death of gauge theory. More than a decade earlier, S. K. Donaldson (Oxford) had found a deep but mysterious link between Yang-Mills theory from mathematical physics on the one hand and 4-dimensional differential topology on the other. Since then, many topologists had become fascinated by gauge theory. Last autumn they found themselves on the receiving end of wry comments, when rumour had it that a new set of equations proposed by E. Witten (IAS, Princeton) had made gauge theory obsolete in topology. This was, of course, an exaggeration.

Witten's equations, originating in his joint work [SW] with N. Seiberg (Rutgers) in quantum field theory and appearing in [W], do provide shortcuts to many of the consequences of gauge theory and quickly lead to proofs of very important new results. The equations are themselves part of a gauge theory and shed new light on the Yang-Mills equations that Donaldson used. Rather than making gauge theory obsolete, Witten's equations make gauge theory even more interesting and more powerful.

The Old Gauge Theory

Coming on the heels of the work of M. H. Freedman (University of California, San Diego) on topological 4-manifolds, Donaldson's use of gauge theory showed that the differentiable classification of smooth 4-manifolds is very different from their classification up to homeomorphism. Combined with Freedman's work, it produced exotic differentiable structures on Euclidean 4-space, an anomaly that does not arise in other dimensions.

> *Given the enormous impact of the $U(1)$ monopole equation on 4-dimensional topology, there are high expectations for the other equations.*

In gauge theory one considers connections or covariant derivatives A on a principal G-bundle over a smooth 4-manifold X endowed with an orientation and a Riemannian metric; here G is a compact Lie group. The connections of interest are the so-called instantons, the solutions of the anti-self-dual Yang-Mills equation, defined as follows: Let $*$ be the Hodge star operator defined by the orientation and the Riemannian metric on X. For A a G-connection, let F^A be its curvature . The self-dual part of the curvature is

$$F_+^A = \frac{1}{2}(F^A + *F^A)$$

and the anti-self-dual Yang-Mills equation is

$$*F^A = -F^A \iff F_+^A = 0.$$

D. Kotschick is professor of mathematics at the University of Basel in Switzerland.

Instantons are the minima of the Yang-Mills functional $YM(A) = \int_X |F^A|^2$. When G is $U(1)$, the anti-self-duality equation is linear and the instantons are completely described by Hodge theory. When G is $SU(2)$, say, the equation is a nonlinear PDE which is elliptic on the space of connections modulo bundle automorphisms or gauge transformations. Its space of solutions, the moduli space of instantons up to gauge transformations, is generically a finite-dimensional smooth manifold M. This manifold is usually noncompact, partly due to the conformal invariance of the equation, and this noncompactness leads to many technical difficulties in the topological applications of gauge theory.

In 1981–82, Donaldson had the insight that the algebraic topology of the moduli space M contains subtle information about the differentiable structure of X. He first proved that certain topological 4-manifolds do not support any differentiable structure at all. Later he defined differentiable invariants of large classes of manifolds which, although difficult to calculate completely, were very successful in distinguishing nondiffeomorphic differentiable structures on X.

This initially came as a complete surprise to topologists. Even after gauge theory had been firmly established as a tool in topology, there was no conceptual understanding of how and why instantons were related to the structure of 4-manifolds.

Rather than making gauge theory obsolete, Witten's equations make gauge theory even more interesting and more powerful.

Early on Donaldson proved that in the case of complex algebraic surfaces connections with anti-self-dual curvature are the same as stable holomorphic bundles in the sense of geometric invariant theory. Furthermore, the instanton invariants are nontrivial for algebraic surfaces. This established a strong link between gauge theory and algebraic geometry. Four-dimensional differential topology was seen as being very close to complex geometry.

Over the last four years, P. B. Kronheimer (Oxford) and T. S. Mrowka (California Institute of Technology) and others developed a structure theory for the instanton invariants. They found that under suitable hypotheses on X all the invariants derived from different $SU(2)$-bundles over X could be packaged into a single analytic function $q : H_2(X, \mathbf{R}) \to \mathbf{R}$ of the form

$$q = e^{Q/2} \sum_{i=1}^{s} a_i e^{K_i}, \tag{1}$$

where Q is the intersection form of X and the $a_i \in \mathbf{Q}$ and $K_i \in H^2(X, \mathbf{Z})$ are certain characteristic elements for the intersection form. The "basic classes" K_i are constrained by the inequality $2g(S) - 2 \geq Q(S,S) + K_i(S)$ for the genus g of any smoothly embedded surface $S \subset X$ with normal bundle of positive degree.

The New Gauge Theory

To write down Witten's equations on a smooth oriented Riemannian 4-manifold X, one has to choose a $Spin^c$-structure, that is, a lift of the frame bundle from $SO(4)$ to $Spin^c(4) = Spin(4) \times_{\pm 1} U(1)$. Associated with this structure are bundles $V_\pm$ of positive and negative spinors and a complex determinant line bundle $L = det(V_\pm)$. Further, there is a canonical map $\sigma : V_+ \times V_+ \to \Lambda^2_+$ defined by taking the trace-free part of an element in $V_+ \otimes V_+$ considered as an endomorphism of V_+.

A $U(1)$-connection A on L, together with the Levi-Civita connection of the Riemannian metric induce a covariant derivative $\Gamma(V_+) \to \Gamma(V_+ \otimes T^*X)$. Composing this with Clifford multiplication $\Gamma(V_+ \otimes T^*X) \to \Gamma(V_-)$ defines a Dirac operator $D_A : \Gamma(V_+) \to \Gamma(V_-)$. Witten's equations for a connection A and a positive spinor $\phi \in \Gamma(V_+)$ are

$$D_A\phi = 0$$
$$F_+^A = i\sigma(\phi, \phi).$$

These equations are invariant under bundle automorphisms of L, but they are not conformally invariant. The solutions, called monopoles, are the minima for the functional

$$\int_X (|F_+^A - i\sigma(\phi,\phi)|^2 + |D_A\phi|^2).$$

The space of monopoles modulo bundle automorphisms is generically a smooth manifold, and it is always compact. Compactness follows easily from the Weitzenböck formula for the Dirac operator combined with standard elliptic theory. Further, if the scalar curvature of the Riemannian metric is nonnegative, all the solutions of the monopole equation have $\phi = 0$ and so are $U(1)$ instantons. Even if the scalar curvature is negative somewhere, there are only finitely many cohomology classes which are the first Chern classes of complex line bundles L admitting nontrivial solutions to the monopole equations.

The compactness of the monopole moduli spaces makes them much easier to handle than the instanton moduli spaces. This is the main reason why Witten's approach is much simpler than Donaldson's. Although instanton and monopole moduli spaces seem to contain very similar information, the monopoles are tied much more closely to the Riemannian geometry of X. This inspires hope for the development of a combinatorial approach to gauge theory.

One can use the monopole moduli spaces to reprove and generalize most of Donaldson's theorems about the nonexistence of differentiable structures on certain topological 4-manifolds. One can also define invariants of differentiable structures, for example by counting the number of points in zero-dimensional moduli spaces. These invariants are trivial for manifolds admitting either a Riemannian metric of positive scalar curvature or a smooth connected sum decomposition in which both summands have intersection forms which are not negative definite.

On the other hand, the invariants are nontrivial for complex algebraic surfaces and, more generally, for symplectic manifolds [T]. This puts strong restrictions on the differential topology and geometry of symplectic 4-manifolds. It implies, for example, that the connected sum of three copies of the complex projective plane does not admit a symplectic structure, although all other known constraints for symplectic manifolds are satisfied in this case.

Witten [W] predicted that the "basic classes" K_i appearing in (1) should be precisely the first Chern classes of complex line bundles L for which the new monopole invariants are nonzero and that the coefficients a_i in (1) should be determined by the values of the monopole invariants. For large classes of manifolds, this prediction was quickly proved to be true, because one can calculate both sets of invariants, for instantons and for monopoles, and compare the answers. There are promising attempts to prove Witten's conjecture in a more direct way.

One of the spectacular new theorems proved using Witten's equations concerns the problem of finding the minimal genus of a smoothly embedded surface in a 4-manifold representing a given homology class. While formula (1) solved this problem in many cases, it seemed very difficult to generalize the argument using instantons to cover, for example, the case when the manifold is the complex projective plane. Now the nontriviality of the monopole invariants associated with the first Chern classes of algebraic surfaces allows one to show that algebraic curves have minimal genus among all smoothly embedded surfaces representing the same homology class (if it has nonnegative self-intersection). This statement for the case of algebraic curves in the complex projective plane [KM] was previously known as the Thom conjecture.


Laura Pedrick/NYT Pictures

E. Witten

The Physical Origins

Yang-Mills theory is a conformally invariant classical field theory whose groundstates are the instantons. Experience with quantum field theory suggests a recipe for turning this into a so-called topological field theory whose correlation functions are Donaldson's instanton invariants. As they are invariants of the differentiable structure, one can vary the Riemannian metric used without affecting the invariants. Witten [W] outlines how studying a family of metrics $g_t = tg_1$, where $t > 0$ is a real parameter, leads naturally to the appearance of the monopole equations.

For small t, the classical approximation to quantum field theory coincides with Donaldson's definition of instanton invariants. For large t, however, the quantum vacuum states of the theory, parametrized by a complex variable u, become relevant. It turns out [SW] that the quantum theory naturally leads one to consider the family of elliptic curves

$$y^2 = (x^2 - 1)(x - u).$$

For generic u, the elliptic curve is smooth, but it degenerates to a rational curve at $u = \pm 1$.

For large classes of manifolds, all the topological information of the instanton theory can be extracted by looking at appropriate elliptic functions in infinitesimal neighbourhoods of the points ± 1 in the u-plane. The reason is that for these special values of the parameter, certain particles in the quantum theory—the monopoles—become massless, although they are not massless in the classical theory. (The monopole equations, like the whole quantum theory, are not conformally invariant, although the classical theory is.) The information about instantons can then be derived from the monopole equations alone, leading to formulae for the Donaldson invariants. In fact, (1) is the simplest such formula, and understanding this was one of the motivations for Witten's approach.

S.K. Donaldson

The monopole equations seem to detect the simplest part of Donaldson theory, which for simple manifolds is all there is. However, there are manifolds like the complex projective plane, on which there are no monopoles at all, for which Donaldson theory has another, more complicated part to it. In the quantum field theory description, this is said to be detected by integration over the u-plane, because for these manifolds there is more to the u-plane than one sees near the special values $u = \pm 1$.

The physical theory developed by Seiberg and Witten [SW] suggests a whole family of coupled equations, of which the monopole equations are just the simplest example. If one considers $SU(N)$ instantons (instead of $SU(2)$), then there is a related set of equations for G-connections A, where G is the dual of the maximal torus of $SU(N)$, and certain spinor fields ϕ. The equations are the Dirac equation for ϕ and an equation expressing F_+^A by a certain hyper-Kähler moment map, generalizing $\sigma(\phi, \phi)$ in the case when $G = U(1)$. These generalized monopole equations are related to higher-dimensional Abelian varieties rather than elliptic curves. Given the enormous impact of the $U(1)$ monopole equation on 4-dimensional topology, there are high expectations for the other equations. Far from being dead, gauge theory is more active and exciting than ever.

References

[KM] P. B. Kronheimer and T. S. Mrowka, *The genus of embedded surfaces in the projective plane*, Math. Res. Lett. **1** (1994), 797-808.

[SW] N. Seiberg and E. Witten, *Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD*, Nuclear Phys. B (to appear).

[T] C. H. Taubes, *The Seiberg-Witten invariants and symplectic forms*, Math. Res. Lett. **1** (1994), 809-822.

[W] E. Witten, *Monopoles and four-manifolds*, Math. Res. Lett. **1** (1994), 769-796.