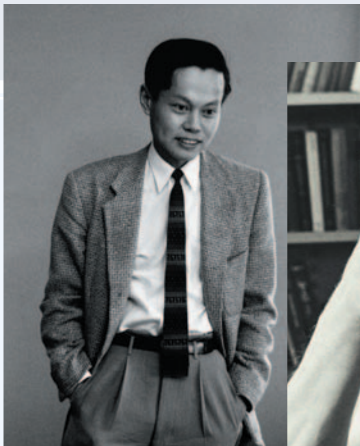
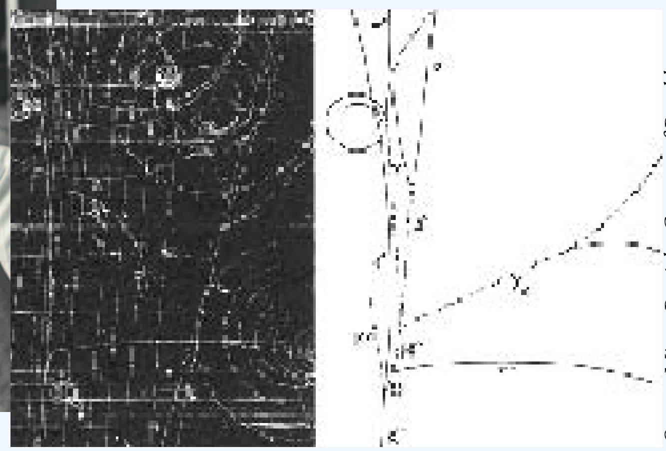
Report on the status of the Yang-Mills problem

by Michael R. Douglas, Professor of Physics at Rutgers University



Chen Ning Yang (b. 1922)
© Estate of François Bello, Science Photo Library, Photo Researchers, Inc.

Robert Leroy Mills (b. 1927)
Ohio State University Photo Services



Omega Minus Particle, Courtesy of Brookhaven National Laboratory

Yang-Mills Existence and Mass Gap: *Prove that for any compact simple gauge group G, quantum Yang-Mills theory of $\mathbb{R}^4$ exists and has a mass gap $\Delta > 0$.*

As explained in the official CMI problem description set out by Arthur Jaffe and Edward Witten, Yang-Mills theory is a generalization of Maxwell's theory of electromagnetism, in which the basic dynamical variable is a connection on a G-bundle over four-dimensional space-time. Its quantum version is the key ingredient in the Standard Model of the elementary particles and their interactions, and a solution to this problem would both put this theory on a firm mathematical footing and demonstrate a key feature of the physics of strong interactions.

To illustrate the difficulty of the problem, we might begin by comparing it to the study of classical Yang-Mills theory. Mathematically, this is a system of non-linear partial differential equations, obtained by extremizing the Yang-Mills action,

$$S = \int \mathrm{Tr}\ F \wedge *F,$$

where $F = dA + A \wedge A$ is the curvature of the G-connection A. Perhaps the most basic mathematical question here is to specify a class of initial conditions for which we can guarantee existence and uniqueness of solutions. Among the qualitative properties of these solutions, one with some analogies to the "mass gap" problem would be to establish or falsify the existence of solitonic solutions, whose energy density remains localized for all times (in fact, these are not believed to exist). Such questions have seen a great deal of mathematical progress in recent years, using techniques that are founded on the well-developed theory of linear PDE's.

By contrast, there is at present no satisfactory mathematical definition of the quantum Yang-Mills theory, because of the famous difficulties of renormalization. Conceptually, the simplest starting point for

Continued on page 16

Report on the status of the Yang-Mills problem

by Michael R. Douglas

Continued from page 5

discussing this is probably to consider Yang-Mills theory on a "lattice," in other words a graph Γ with vertices, edges and a set of faces or "plaquettes," each of which is a closed loop in Γ. For example, we could take the vertices to be integral points $\mathbb{Z}^4 \subset \mathbb{R}^4$; the edges to be the straight lines connecting pairs of points at unit distance, and the plaquettes to be the loops of total length four. A G-connection on Γ is then a map from edges into G, and the curvature of the connection on a specified plaquette is (minus) the trace of the holonomy around that loop minus the identity matrix. The Yang-Mills action can again be taken to be the sum of squared curvatures.

Given this explicit description of the space of configurations, we can define "quantum lattice Yang-Mills" on a finite subgraph γ of Γ in terms of a functional integral,

$$Z[\gamma, g^2, G] = \int \prod dU_i \; e^{-\frac{1}{g^2}S},$$

where the integral is over all holonomies in γ, in other words all maps from edges into G; the measure is the product of Haar measure for the holonomy on each edge, and S is the Yang-Mills action. Finally, the real parameter g^2 is the "bare coupling constant," which much like Planck's constant controls the strength of quantum fluctuations.

The quantity Z is known as the partition function. The other quantities of physical interest are expectation values under this measure; for example, the joint expectation value of the curvatures on a pair or set of plaquettes.

Taking these finite-dimensional integrals as our starting point, the question of the existence of quantum Yang-Mills is essentially whether there is any sensible way to define the limit of the $Z[\gamma, g^2, G]$ over successively larger subgraphs $\gamma \subset \Gamma$, and thus to define a functional integral on Γ. Taking this limit will clearly involve renormalization, and this example was Wilson's primary motivation for his study of the renormalization group.

A great deal is known about the expected properties of the limit from a variety of physical arguments. Most importantly, in the limit of $g^2 \to 0$ and large γ, it is believed that the specific choice of Γ approximating $\mathbb{R}^4$ will not matter; the resulting expectation values will converge to "correlation functions in a continuum quantum field theory," satisfying formal properties that include invariance under the isometry group of the flat metric on $\mathbb{R}^4$, and others formalized in the Osterwalder-Schrader axioms. The additional axioms provide sufficient conditions for the construction of a Hilbert space and operator interpretation of the theory, analogous to the standard operator interpretation of quantum mechanics.

Establishing these axioms is the "existence" part of the problem, while the "mass gap" claim involves the fall-off of correlation functions with distance, as explained in more detail in Jaffe and Witten's problem description. We should say that starting with lattice Yang-Mills is not essential, and a variety of other starting points, such as other approximations to the functional integral, have been considered. The point is to define the limiting continuum field theory. In principle, this might even be done without taking limits.

So far as I know, no breakthroughs have been made on this problem in the last few years. In particular, while progress has been made in lower-dimensional field theories, I know of no significant progress toward a mathematically rigorous construction of the quantum Yang-Mills theory. The state of the art remains the works of Balaban and of Magnen, Rivasseau and Sénéor cited in the problem description.

There has, however, been interesting progress on various related problems. First, there is much to be learned from the mathematical study of more general lower-dimensional field theories. A recent overview of constructive quantum field theory is Rivasseau, math-ph/0006017.

There are two classes of quantum field theories which are generally believed to bear a close similarity to four-dimensional Yang-Mills theory. The first is the two-dimensional nonlinear sigma model with target space a group manifold G, or, more generally, a symmetric space M of positive Riemannian curvature.

This is a theory whose fields are maps from two-dimensional space-time into M. Apparently, even these theories have not yet been constructed to the standards required in the problem description. On the other hand, the mass gap has been exhibited in a regulated version of the $O(N)$ sigma model [3]. Physically, these theories are known to be integrable, a point we will return to below.

The other broad class of models with great similarity to Yang-Mills, but significantly simpler, are the four-dimensional supersymmetric Yang-Mills theories. These are modifications of Yang-Mills for which the fields include, in addition to the connection of our previous discussion, various "fermionic" and other fields chosen to realize the following property: the Hamiltonian, the operator on the quantum Hilbert space which generates time translations, has a square root, called the "supercharge."

Many wonderful properties and simplifications follow from supersymmetry, as discussed in [1, 2]. Physically, the most important is that the renormalization problem is mitigated, which one would hope could make defining the theory easier. Mathematically, the most important is the relation to "topological field theory," first proposed by Witten, and discussed in some detail in the references we just gave.

Furthermore, although one has changed the problem, one still has a fairly close relation to the original problem using the ideology of the renormalization group. Namely, one can start with a supersymmetric theory, and add supersymmetry breaking terms to the action which only become important at long distances (many lattice spacings in the earlier discussion). This will produce a theory with the better renormalization properties of the supersymmetric theory at short distances, but which reduces to conventional Yang-Mills theory at longer distances. Thus, a solution to the problem in a sufficiently general class of supersymmetric theories would in fact imply the solution of the original problem.

While at present we appear no closer to this goal than the original one, there has been significant progress in the past few years in understanding supersymmetric gauge theory. Perhaps the most interesting recent results are Nekrasov's rederivation of the Seiberg-Witten solution using instanton methods [4, 5], and the recent evidence for integrability of the large N supersymmetric $N = 4$ gauge theory [6] (for a recent review, see [7]).

The first of these starts from the celebrated 1994 Seiberg-Witten solution of $N = 2$ supersymmetric gauge theory. While mathematicians are perhaps more familiar with the consequences of this work for four-dimensional topology, the original physics motivation and significance of this work was in fact that it addressed the "mass gap" and related problems in a supersymmetric analog of Yang-Mills theory. Indeed, the basic reason for the simplicity of the Seiberg-Witten invariants, compared to the Donaldson invariants, is the mass gap property, which allows reducing the computation of the invariants to a lower-dimensional problem.

The original work of Seiberg and Witten obtained the mass gap property by an ingenious argument that involved a clever ansatz for the "supersymmetric effective action" describing the long-distance properties of the theory. Subsequent work showed that their ansatz was the only consistent possibility out of a large class, but this still left the justification of the overall framework more or less as pure physical intuition.

On the other hand, their ansatz had precisely the form expected from a direct computation of the sum of an instanton expansion, defined as a succession of integrals over moduli spaces of Yang-Mills instantons. Thus one could hope to validate it by direct computation. Furthermore, these moduli spaces have an explicit description, due to Atiyah, Drinfeld, Hitchin and Manin, so the necessary computation was rather explicit. It was, however, too difficult to perform directly.

This problem was solved in a *tour de force* work of Nekrasov, which started from the problem of instanton sums in the noncommutative deformation of the gauge theory originally introduced by Alain Connes and Marc Rieffel. In the noncommutative theory, instanton moduli spaces are smooth and significantly simpler, but many further clever tricks were required to come to a result. In a sense, this result establishes Seiberg and Witten's result as a sort of four-dimensional analog of mirror symmetry, in which the explicit instanton sum is the analog of numbers of holomorphic curves on a Calabi-Yau, and the mirror is the Riemann surface encoding the Seiberg-Witten solution.

One can argue that this is the simplest result in four-dimensional gauge theory that carries any of the physics of the mass gap, and as such might be the best way in to the complexities of the problem.

The second result, integrability of $N = 4$ super Yang-Mills, is harder to explain in mathematical terms, but we mention it because of its novelty and the possibility that it will drastically change how we think about these problems. Its starting point is the "AdS/CFT correspondence" of Maldacena, according to which this version of Yang-Mills theory can be reformulated as a string theory in anti de Sitter (AdS) space. This is a difficult string theory to work with, even by physicists' standards, and few explicit results are known. But the striking claim of Minahan, Zarembo and others is that, at least for the $SU(N)$ gauge group in the limit of large N, this string theory becomes exactly solvable (or integrable) in precisely the sense that we mentioned above for the two-dimensional sigma model. This is essentially because of the high degree of symmetry of AdS and the other spaces involved. The full picture is still rather mysterious at present, but might lead to a detailed connection between the much better understood two-dimensional field theories, and the original Yang-Mills problem in four dimensions.

To summarize, while I would not bet on this problem being solved in the next few years, it remains a fertile ground for mathematical discovery.

References

[1] P. Deligne, P. Etingof, D. Freed, L. Jeffrey, D. Kazhdan, J. Morgan, D. Morrison, E. Witten, eds., *Quantum Fields and Strings: A Course for Mathematicians*, American Mathematical Society, Providence, RI , 1999, two volumes.

[2] D. Freed, *Five Lectures on Supersymmetry*, American Mathematical Society, Providence, 1999.

[3] C. Kopper, Comm. Math. Phys. 202, 89 (1999).

[4] N. A. Nekrasov, "Seiberg-Witten prepotential from instanton counting," Proceedings of the ICM, Beijing 2002, vol. 3, 477–496, arXiv:hep-th/0306211.

[5] N. A. Nekrasov and A. Okounkov, "Seiberg-Witten Theory and Random Partitions," arXiv:hep-th/0306238.

In the late 1960s, I began a new formulation of gauge fields through the approach of nonintegrable phase factors. It happened one semester I was teaching general relativity, and I noticed that the formula in gauge theory,

$$F_{\mu\nu} = \frac{\partial B_\mu}{\partial x_\nu} - \frac{\partial B_\nu}{\partial x_\mu} + i\epsilon\left(B_\mu B_\nu - B_\nu B_\mu\right)$$

and the formula in Riemannian geometry

$$R^{\ell}_{ijk} = \frac{\partial}{\partial x^j}\left\{\frac{\ell}{i\,k}\right\} - \frac{\partial}{\partial x^k}\left\{\frac{\ell}{i\,j}\right\} + \left\{\frac{m}{i\,k}\right\}\left\{\frac{\ell}{m\,j}\right\} - \left\{\frac{m}{i\,j}\right\}\left\{\frac{\ell}{m\,k}\right\}$$

are not just similar --- they are, in fact, the same if one makes the right identification of symbols! It is hard to describe the thrill I felt at understanding this point.

Excerpts from an interview with C. N. Yang in the Mathematical Intelligencer, Vol. 15, No. 4 (Fall 1993) 13–21.