Quantum Yang-Mills Theory

Arthur Jaffe and Edward Witten

1 The Physics of Gauge Theory

Since the early part of the twentieth century, it has been understood that quantum mechanics is needed to describe nature at the subatomic scale. In quantum mechanics, the position and velocity of a particle are noncommuting operators acting on a Hilbert space, and classical notions such as "the trajectory of a particle" do not apply.

But quantum mechanics of particles is not the whole story. In nineteenth and early twentieth century physics, many aspects of nature were described in terms of fields – the electric and magnetic fields that enter in Maxwell's equations, and the gravitational field governed by Einstein's equations. It was clear by the late 1920's that, since fields interact with particles, to give an internally coherent account of nature, the quantum concepts must be applied to fields as well as particles.

When this is done, quantities such as the components of the electric field at different points in spacetime become noncommuting operators. When one constructs a Hilbert space in which these operators act, one finds many surprises. The distinction between fields and particles breaks down, since the Hilbert space of a quantum field is constructed in terms of particle-like excitations. Conventional particles, such as electrons, are reinterpreted as arising from the quantization of a field. In the process, one finds the prediction of "antimatter"; for every particle, there must be a corresponding antiparticle, with the same mass and opposite electric charge. Soon after being predicted by P. A. M. Dirac on the basis of quantum field theory, the "positron" or oppositely charged antiparticle of the electron was discovered in cosmic rays.

The Quantum Field Theories (QFT's) that have proved to be most important in describing elementary particle physics are gauge theories. The classical example of a gauge theory is the theory of electromagnetism. The gauge group is the abelian group $U(1)$. If A denotes the $U(1)$ gauge connection, which locally can be regarded as a one-form on spacetime, then the curvature or electromagnetic field tensor is the two-form $F = dA$, and Maxwell's equations read $0 = dF = d^*F$. Here $*$ is the Hodge duality operator; Hodge indeed introduced his celebrated theory of harmonic forms as a generalization to p-forms of Maxwell's equations. Maxwell's equations describe large scale electric and magnetic fields and also – as Maxwell discovered – the propagation of light waves, at a characteristic velocity, the speed of light.

Yang-Mills theory or non-abelian gauge theory can, at the classical level, be described similarly, with $U(1)$ replaced by a more general compact gauge group G. The definition of the curvature must be modified to $F = dA + A \wedge A$, and Maxwell's equations are modified to the Yang-Mills equations,

$0 = d_A F = d_A * F$, where d_A is the gauge-covariant extension of the exterior derivative. These equations can be derived from the Yang-Mills Lagrangian

$$L = \frac{1}{4g^2} \int \mathrm{Tr}\, F \wedge *F, \tag{1}$$

where Tr denotes an invariant quadratic form on the Lie algebra of G. The Yang-Mills equations are nonlinear, so, in contrast to the Maxwell equations, but like the Einstein equations for the gravitational field, they are not explicitly soluble in general. But they have certain properties in common with the Maxwell equations, and in particular, they describe at the classical level massless waves that travel at the speed of light.

By the 1950's, when Yang-Mills theory was introduced, it was known that the quantum version of Maxwell theory – known as Quantum Electrodynamics or QED – gave a very accurate account of quantum behavior of electromagnetic fields and forces. The question arose of whether the non-abelian analog was important for describing other forces in nature, notably the weak force (responsible among other things for certain forms of radioactivity) and the strong or nuclear force (responsible among other things for the binding of protons and neutrons into nuclei). The masslessness of classical Yang-Mills waves was a serious obstacle to applying Yang-Mills theory to the other forces, for the weak and nuclear forces are not associated with long range fields or massless particles.

In the 1960's and 1970's, these obstacles to physical applications of nonabelian gauge theory were overcome. In the case of the weak force, this was accomplished by the Weinberg-Salam-Glashow electroweak theory [38, 31] with gauge group $H = SU(2) \times U(1)$. The masslessness of classical Yang-Mills waves was avoided by elaborating the theory with an additional "Higgs field." This is a scalar field, transforming in a two-dimensional representation of H, whose non-zero and approximately constant value in the vacuum state reduces the structure group from H to a $U(1)$ subgroup (diagonally embedded in $SU(2) \times U(1)$). This theory describes both the electromagnetic and weak forces, in a more or less unified way; because of the reduction of the structure group to $U(1)$, the long range fields are those of electromagnetism only, in accord with what we see in nature.

The solution of the problem of massless Yang-Mills fields for the strong interactions was of a completely different nature. The solution was not obtained by adding additional fields to Yang-Mills theory but by discovering a remarkable property of the quantum Yang-Mills theory itself, that is, of the quantum theory whose classical Lagrangian has been given in (1). This property [18, 30] is called "asymptotic freedom." When a quantum theory is asymptotically free, this means, roughly, that the quantum behavior at short distances is very similar to the classical behavior, but that the classical theory is not a good guide to the quantum behavior at long distances.

Asymptotic freedom, together with other experimental and theoretical discoveries made in the 1960's and 70's, made it possible to describe the nuclear force by a non-abelian gauge theory in which the gauge group is $G = SU(3)$. The additional fields describe, at the classical level, "quarks," which are spin 1/2 objects somewhat analogous to the electron, but transforming in the fundamental representation of $SU(3)$. The non-abelian gauge theory of the strong force is called Quantum Chromodynamics (QCD).

The use of QCD to describe the strong force was motivated by a whole series of experimental and theoretical discoveries made in the 1960's and 70's, involving the symmetries and high energy behavior of the strong interactions. But classical nonabelian gauge theory is very different from the observed world of strong interactions; for QCD to successfully describe the strong force, it must have at the quantum level the following three properties, each of which is dramatically different from the behavior of the classical theory:

(1) It must have a "mass gap," that is, there must be some strictly positive constant Δ such that every excitation of the vacuum has energy at least Δ.

(2) It must have "quark confinement," that is, even though the theory is described in terms of elementary fields, such as the quarks, that transform non-trivially under $SU(3)$, the physical particle states – such as the proton, neutron, and pion – are $SU(3)$-invariant.

(3) It must have "chiral symmetry breaking," which means that the vacuum is potentially invariant (in the limit that the quark bare masses vanish) only under a certain subgroup of the full symmetry group that acts on the quark fields.

The first point is necessary to explain why the nuclear force is strong but short-ranged; the second is needed to explain why we never see individual quarks; and the third is needed to account for the "current algebra" theory of soft pions that was developed in the 1960's.

Both experiment – since QCD has numerous successes in confrontation with experiment – and computer simulations [5] carried out since the late 1970's have given strong encouragement that QCD does have the properties cited above. These properties can be seen, to some extent, in theoretical calculations carried out in a variety of highly oversimplified models (like strongly coupled lattice gauge theory [39]). But they are not fully understood theoretically; there does not exist a convincing, even if not mathematically complete, theoretical computation demonstrating any of the three properties in QCD, as opposed to a severely simplified truncation of it.

2 Quest For Mathematical Understanding

In surveying the physics of gauge theories in the last section, we have considered both classical properties – such as the Higgs mechanism for the electroweak theory – and quantum properties that do not have classical analogs – like the mass gap and confinement for QCD. Classical properties of gauge theory are within the reach of established mathematical methods, and indeed classical non-abelian gauge theory has played a very important role in mathematics in the last twenty years, especially in the study of three- and four-dimensional manifolds. On the other hand, one does not yet have a mathematical understanding of the quantum behavior of four-dimensional gauge theory, or even a precise definition of quantum gauge theory in four dimensions. Will this change in the twenty-first century? We hope so!

At times, mathematical structures of importance have first appeared in physics before their mathematical importance was fully recognized. This happened with the discovery of calculus, which was needed to develop Newtonian mechanics, with functional analysis and group representation theory, topics whose importance became clearer with quantum mechanics, and even with the study of Riemannian geometry, whose development was greatly intensified once it became clear, through Einstein's invention of General Relativity to describe gravity, that this subject plays a role in the description of nature.

Quantum Field Theory (QFT) became increasingly central in physics throughout the twentieth century. There are reasons to believe that it may be important in twenty-first century mathematics. Indeed, many mathematical subjects that have been actively studied in the last few decades can be naturally formulated – at least at a heuristic level – in terms of QFT. New structures spanning analysis, algebra, and geometry have emerged.

On the analytic side, a byproduct of the mathematical construction of certain quantum field theories was the construction of a new class of measures: non-gaussian, quasi-invariant Borel measures on generalized functions. These measures are quasi-invariant under the action of the Euclidean group. Renormalization theory arises from the physics of quantum field theory and statistical physics, and provides the basis for the mathematical investigation of the local (ultra-violet) regularity and of the global decay (infra-red regularity) of quantum field theories. Surprisingly the ideas from renormalization theory also apply in other areas of mathematics, including in classic work on the convergence of Fourier series, and in recent progress on classical dynamical systems.

On the algebraic side, investigations of soluble models of quantum field theory and statistical mechanics have led to many new discoveries involving topics such as Yang-Baxter equations, quantum groups, bose-fermi equivalence in two dimensions, and rational conformal field theory.

Geometry abounds with new mathematical structures rooted in quantum field theory, many of them very actively studied in the last twenty years. Examples include Donaldson theory of four-manifolds, the Jones polynomial of knots and its generalizations, mirror symmetry of complex manifolds, elliptic cohomology, and $SL(2, \mathbf{Z})$ symmetry in the theory of affine Kac-Moody algebras.

QFT has in certain cases suggested new perspectives on mathematical problems, while in other cases its mathematical value up to the present time is motivational. In the case of the geometric examples cited above, a mathematical definition of the relevant QFT's (or one in which the relevant physical techniques can be justified) is not yet at hand. Existence theorems that put QFT's on a solid mathematical footing are needed to make the geometrical applications of QFT into a full-fledged part of mathematics. Regardless of the future role of QFT in pure mathematics, it is a great challenge for mathematicians to understand the physical principles that have been so important and productive throughout the twentieth century.

Finally, QFT is the jumping off point for a quest that may prove central in twenty-first century physics – the effort to unify gravity and quantum mechanics, perhaps in string theory. For mathematicians to participate in this quest, or even to understand the possible results, QFT must be developed as a mathematical subject.

3 The Problem

For all of these reasons, the Clay Mathematics Institute has chosen to formulate a Millenium Problem involving the quest for mathematical insight into four-dimensional QFT. This problem requires understanding one of the deep unsolved physics mysteries about the behavior of QFT and producing a mathematical framework for its solution.

The problem that has been chosen involves quantum gauge theory in four dimensions. This has been selected because of the great physical importance of four-dimensional quantum gauge theory and because of the importance of four-dimensional gauge theory in geometry.

A quantum field, or local quantum field operator, is an operator-valued generalized function on spacetime obeying certain axioms. The quantum fields act in a Hilbert space $\mathcal{H}$ that furnishes a positive energy representation of the Poincaré group (or inhomogeneous Lorentz group). Among other things, we require that $\mathcal{H}$ contains a vacuum vector Ω, namely a unit vector that is invariant under the representation of the Poincaré group and unique up to multiplication by a phase. The properties required of the quantum fields are described at a physical level of precision in many textbooks, see for example [22]. Mathematically precise systems of axiom have been given for quantum field theories on $\mathbf{R}^4$ with a Minkowski signature by Wightman, see [36], and on $\mathbf{R}^4$ with a Euclidean signature by Osterwalder and Schrader [29]. One of the achievements of twentieth century axiomatic quantum field theory was the proof of theorems showing how to convert a Euclidean quantum field theory to a Lorentzian one, and vice-versa. Under certain technical assumptions on the growth of the n-point Euclidean Green's functions as a function of n, Osterwalder and Schrader prove the equivalence of these two axiom schemes. The continued mathematical exploration of quantum field theory may possibly lead to refinements of the axiom sets that have been in use up to now, perhaps to incorporate properties considered important by physicists such as the existence of an operator product expansion or of a local stress-energy tensor. At any rate, for purposes of the CMI Millenium Problem, an existence proof for a quantum field theory must establish axioms at least as strong as those cited in [36, 29].

To establish existence of four-dimensional quantum gauge theory with gauge group G, one should define (in the sense of the last paragraph) a quantum field theory with local quantum field operators in correspondence with the gauge-invariant local polynomials in the curvature F and its covariant derivatives, such as $\operatorname{Tr} F_{ij}F_{kl}(x)$. (A natural $1-1$ correspondence between such classical 'differential polynomials' and quantized operators does not exist, since the correspondence has some standard subtleties involving renormalization [22]. The space of classical differential polynomials of dimension $\leq d$ does correspond to the space of local quantum operators of dimension $\leq d$.) Correlation functions of the quantum field operators should agree at short distances with the predictions of asymptotic freedom and perturbative renormalization theory, as described in textbooks. Those predictions include among other things the existence of a stress tensor and an operator product expansion, having perscribed local singularities predicted by asymptotic freedom.

The Hamiltonian H of a quantum field theory is the element of the Lie algebra of the Poincaré group that generates time translations. Since the vacuum vector Ω is Poincaré invariant, it is an eigenstate with zero energy, namely $H\Omega = 0$. The positive energy axiom asserts that in any quantum field theory, the spectrum of H is supported in the region $[0, \infty)$. A quantum field theory has a *mass gap* if H has no spectrum in the interval $(0, \Delta)$ for some $\Delta > 0$. The supremum of such Δ is called the mass m.

Now we can state the Millenium Problem:

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

An important consequence of the existence of a mass gap is that for any positive constant $C < \Delta$ and for any local quantum field operator $\mathcal{O}$ such that $\langle \Omega, \mathcal{O}\Omega \rangle = 0$, one has

$$|\langle \Omega, \mathcal{O}(x)\mathcal{O}(y)\Omega \rangle| \leq \exp(-C|x-y|) \ , \tag{2}$$

as long as $|x-y|$ is sufficiently large (depending on C and $\mathcal{O}$). This gives a locality property that, roughly speaking, may make it possible to apply results on $\mathbf{R}^4$ to any four-manifold, as argued at a heuristic level (for a supersymmetric extension of four-dimensional gauge theory) in [40]. Thus, it may be important in mathematical applications of four-dimensional quantum gauge theories. The importance of the mass gap of four-dimensional gauge theory for physics has already been explained in the introduction.

There are many natural extensions of the Millenium Problem. Among other things, one would like to prove to prove confinement as well as the mass gap, to prove existence of various generalizations of four-dimensional quantum gauge theory in which one incorporates additional fields (preserving asymptotic freedom), and to understand dynamical questions (such as the possible mass gap, confinement, and chiral symmetry breaking) in these more general theories, and to extend the existence theorems from $\mathbf{R}^4$ to an arbitrary four-manifold.

But a solution of the existence and mass gap problem as stated above would be a turning point in the mathematical understanding of quantum field theory, with a chance of opening new horizons for its applications.

4 Remarks, Known Results, and Other Directions

We review here some results on the existence and physical properties of QFT's obtained by methods of constructive quantum field theory (CQFT). This work concerns quantum fields in two and in three space-time dimensions. We aim to give the reader an orientation and a set of references for background material. The final section sketches some existing ideas that may be useful in the study of the Yang-Mills problem.

Since the inception of quantum field theory, two central methods have emerged to show the existence of quantum fields on non-compact configuration space (such as Minkowski space). These known methods are:

(i) Find an exact solution in closed form.

(ii) Establish convergence of a sequence of solutions to a sequence of approximate problems.

The first method, which is only useful in very special cases, arises in the study of problems without interaction (linear quantum fields); it has been used to investigate examples that are integrable; and it also arises in examples that can be solved by a clever change of variables. The second method lies at the basis of *constructive quantum field theory* or CQFT, and it relies on establishing uniform *a priori* estimates. These estimates give quantitative control over the deviation of the limiting solution to a sequence of approximations, as well as eventually leading to convergence of these approximate solutions. Often these estimates are sufficient to establish the existence of a uniform mass gap in the sequence of approximating problems, and also to show that this gap persists in the limiting solution.

Two schemes for establishing estimates are

(ii.a) Correlation inequalities.

(ii.b) Convergent expansions.

The correlation inequality methods have an advantage of being general, and when applicable hold for all coupling constants. But their disadvantage is that they rely on special properties of scalar bosons or abelian gauge theories, and are not applicable to all interactions.

4.1 Small Coupling Constants

The rest of this discussion focuses on expansion methods. A powerful set of expansion methods have been developed to establish the existence of certain examples, to exhibit uniqueness or non-uniqueness of these solutions, as well as to establish properties of the small-energy spectrum of the Hamiltonians for these theories. These methods are limited to cases where one can establish the existence of a convergent expansion about an approximate solution; this means they apply to weak coupling — in an appropriate parameterization of the coupling constants.

These expansion methods, known as *cluster expansions*, are based on iterative analysis of degrees of freedom on different length scales. These methods occur in two generic settings: in Hilbert space and in functional integrals. Quantum theory is based on the existence of a Hilbert space of physical states, and Hilbert space methods are central to many mathematical constructions. Operator methods are complemented by estimates established using functional integration. The intimate connection between quantum theory (operators) and probability theory (functional integration) is beautifully summarized in the formulations of Nelson and of Osterwalder and Schrader.

There is another convergence method that one can characterize as *weak*. These methods rely on compactness arguments, and they end up proving the convergence of subsequences of approximations, without actually establishing convergence. But using a weak method may lead to lost control over properties of the limiting solution: for example, properties of the spectrum that hold in each approximating theory (such as a uniform mass gap), do not necessarily carry over to the limit. For this reason, we exclude weak existence results from consideration as a part of the solution to this prize problem.

In many cases, proofs of a uniform gap in a solution obtained through method (ii) involve a semi-classical approximation. Often this semi-classical approximation can be made directly, while in other cases, the semi-classical approximation is only possible after a change of variables, replacing the original field by a function of this field. We illustrate these possibilities by examples that have been investigated in detail in space-time dimension two or three. The easiest cases concern a small perturbation of the massive free field φ. Let us first consider the simple example with a normal-ordered energy density

$$\lambda\varphi^4 + \varphi^2 \,, \qquad \text{with } \; 0 < \lambda \ll 1 \,. \tag{3}$$

It has been shown in [16, 15] that this field theory exists, is unique, and has a mass gap equal to

$$m = \sqrt{2} + O(\lambda^2) \,. \tag{4}$$

Assume that the operators $\mathfrak{A}$ have the property that the vectors $\mathfrak{A}\Omega$ are dense in $\mathcal{H}$. The proof of the existence of the mass gap relies on proving an estimate: given any constant $C < \sqrt{2}$, there exists a constant $\lambda_0 > 0$, and constants B depending on A, such that for all $A \in \mathfrak{A}$, and all $0 \leq \lambda \leq \lambda_0$, the expectation of e^{-tH} decays exponentially,

$$\left\langle A\Omega, e^{-tH} A\Omega \right\rangle \leq B\, e^{-Ct} \,, \qquad \text{for} \qquad 1 \leq t \,. \tag{5}$$

The states of a single particle are eigenstates of a "mass" operator acting on the Hilbert space, with the corresponding eigenvalue equal to the particle's mass. According to special relativity, the mass M, energy H and momentum P are commuting operators related (in units with the velocity of light equal to one) by $M = \sqrt{H^2 - P^2}$. For this example, much is known, and elaborations of the cluster expansion allow more detailed study of the spectrum of M. In fact, m is an isolated eigenvalue in the spectrum of M, and the corresponding eigenvalues are the states of one observable particle [16]. These states transform under an irreducible representation of the Poincaré group.

The estimate (5) shows that for any $\epsilon > 0$, the gap satisfies $\Delta > \sqrt{2} - \epsilon$ for sufficiently small λ. However, it neither establishes the existence of a particle (of mass m), nor the asymptotic expansion (4) of the mass in λ. In [16] one develops an expansion yielding a linear operator E_2 that is approximately a projection onto vectors in $\mathcal{H}$ that have energy less than the mass of 2 particles. One shows that there exists E_2 whose range is spanned by vectors of the form Ω and $e^{-sH}\varphi\Omega$, and for which the exponential rate of decay in (5) can be improved to twice its value,

$$\left|\left\langle A\Omega, e^{-tH}(I - E_2) A\Omega \right\rangle\right| \leq B\, e^{-2Ct} \,, \qquad \text{for} \qquad 1 \leq t \,. \tag{6}$$

This "two-particle" decay, along with an analysis of the states of the form $e^{-sH}\varphi\Omega$ leads to the proof of the existence of the isolated eigenvalue m of M satisfying (4), as well as to the existence of an upper gap in the spectrum of M. In fact, using correlation inequalities valid only for this example, one can show that on the subspace of $\mathcal{H}$ of vectors even under $\varphi \to -\varphi$ symmetry, the Hamiltonian has no spectrum in the interval $(0, 2m)$ [14]. In the language of physics, this establishes the absence of *bound states* and the repulsiveness of forces in this $\lambda\varphi^4 + \varphi^2$ example with small λ.

The existence of an isolated eigenvalue m for M, in addition to the validity of the Wightman axioms, justifies the applicability of two standard forms of scattering theory: the Haag and Ruelle theory (a wave operator approach), and the Lehmann, Symanzik, and Zimmermann theory (based on matrix elements). The consequence of applying these scattering theories is the existence of asymptotic states with an arbitrary number of particles, see the discussion in [15]. Leaving aside the important open question of *asymptotic completeness, i.e.* whether these asymptotic states span the Hilbert space $\mathcal{H}$, this gives a full particle interpretation for these examples.

Existence and an isolated one-particle spectrum is also known for more general polynomial interactions of the form $\lambda\mathcal{P}(\varphi)+\varphi^2$ on $\mathbf{R}^2$, with $\mathcal{P}$ bounded from below and with $0<\lambda<\lambda_0\ll 1$. (These examples are known in the literature as $\mathcal{P}(\varphi)_2$ models with small coupling, with the subscript denoting the space-time dimension.) In these examples, the presence or the absence of a bound states depends on the details of $\mathcal{P}$. A field theory with a bound state would have a mass operator M whose spectrum has the mass gap $(0,m)$ and the upper gap (m,m_b), with $m<m_b<2m$, while the bound state would occur as an isolated eigenvalue m_b. The *binding energy* is $2m-m_b>0$, and this binding energy vanishes as $\lambda\to 0$. For references, see [35, 6, 15]. In the three-dimensional case, there is a solution for interaction $\lambda\varphi^4+\varphi^2$. A major step in this example was the proof of stability for this Hamiltonian, after appropriate renormalization [12]. Feldman and Osterwalder completed the construction of the φ_3^4 field theory, establishing existence, uniqueness, and a mass gap for sufficiently small λ [7].

The existence of boson-fermion interactions is also known. It was carried out in two dimensions for Yukawa-type interactions of the form $\overline{\psi}\psi\varphi$ and their generalization of the sort $\mathcal{P}(\varphi)+\overline{\psi}\psi\mathcal{Q}''(\varphi)$, with appropriate renormalization, see [15]. The supersymmetric case arises from $\mathcal{P}=\left|\mathcal{Q}'\right|^2$, and requires extra care in dealing with cancellations of divergences, see [23] for references. Partial results on the three-dimensional Yukawa interaction (the case $\mathcal{P}=0$, $\mathcal{Q}''=\varphi$) have been established, see [25], as well in other more singular interactions [10].

The proof of these results depends on an analysis of ultra-violet regularization (smoothing) and infra-red regularization (compactification). One removes these regularizations after establishing *a priori* estimates. Both operator inequalities as well as methods based on functional integration play a role, and these two methods are related through Feynman-Kac representations for the regularized examples. It is important to understand the contribution to estimates arising from degrees of freedom (Fourier components of the field) corresponding to approximate localizations in both configuration space and Fourier space. This philosophy of *phase-cell localization* generally applied iteratively, forms the basis of proving all but the most straightforward estimates, see [16, 15, 12]. In the two-dimensional and three-dimensional examples, one obtains a local field theory, a representation of the canonical commutation relations for the initial data, and also a representation of the inhomogeneous Lorentz group. In the two-dimensional $\mathcal{P}(\varphi)_2$ examples the representation of the canonical commutation relations in each bounded region of space-time is unitarily equivalent to the representation on Fock space [11], but it is globally inequivalent to the Fock representation. In contrast, in the three-dimensional φ_3^4 example, the representation of the commutation relations is locally unitarily inequivalent to the Fock representation [12].

Some gauge theory examples have also been considered. In two dimensional space-time, the interaction of an abelian gauge field with a complex (charged) scalar field with a quartic self coupling is known as the Higgs model. Charge invariance of the interaction is associated with a $U(1)$

symmetry, whose breaking leading to a mass for the gauge boson, the Higgs effect, see [3].

In three dimensional space-time, the pure Yang-Mills theory has been studied on a toroidal lattice, especially by Balban [1]. Estimates on the effective action has been established, uniform estimates in the lattice spacing, however, the Wightman axioms have not been proved for the Yang-Mills theory in three-dimensional space-time. Over the past few years there has also been substantial progress toward the proof of existence of a Yang Mills theory on a compactified four-dimensional space-time, namely on studying Yang-Mills theory on $\mathbf{T}^4$, the toroidal compactification of $\mathbf{R}^4$. Balaban studied this question on a toroidal lattice [2]; other work centered on non-gauge invariant approximations in the continuum [24]. This represents a major advance in understanding the behavior of the renormalization of the Yang-Mills action, and in obtaining the structure of the effective action for large values of the gauge field. In the lattice theory, one starts by constructing a classical configuration obtained by minimizing the action in a particular local choice of gauge. Then one applies a renormalization transformation integrating out fluctuations about this classical solution. The main feature of this progress is obtaining analytic control over effective actions after the action of a renormalization transformation. These results make plausible that with sufficient new work a proof of the existence of Yang-Mills theory might be given on $\mathbf{T}^4$. They do not, however, shed light on the mass gap and for that reason they do not give any information about the limit as $\mathbf{T}^4 \to \mathbf{R}^4$.

4.2 Large Coupling Constant

The field theory with energy density $\lambda\varphi^4 + \varphi^2$ exists on $\mathbf{R}^2$ for all positive λ, a result that has been proved using correlation inequality methods and monotonicity in the volume. For sufficiently large λ, this interaction has (at least) two different solutions. Each solution corresponds to a solution to the Wightman axioms with a unique, Poincaré-invariant vacuum state. These bosonic field theories in a finite spatial volume have a unique ground state, represented by a positive wave function. The distinct ground states in infinite spatial volume appear as a bifurcation of the unique solutions, as one passes to the limit of infinite volume. This phenomenon is called a *phase transition*. In the case that $\langle \Omega, \varphi\Omega \rangle = 0$ in each approximating theory with spatial volume $R < \infty$, the limit $R \to \infty$ does not yield an expectation that is a pure state; rather it gives an even statistical *mixture* of the two pure states that characterize the individual phases. Some intermediate value λ_c of the coupling constant provides a boundary between the uniqueness of the solution (for $\lambda < \lambda_c$) and the existence of a phase transition $\lambda > \lambda_c$, and λ_c is known as the *critical point*. As λ increases to λ_c, the mass gap $m = m(\lambda)$ decreases monotonically and continuously to zero [19, 13, 27]; furthermore, the function $m(\lambda)$ is known to be a positive, increasing function of λ for $\lambda \gg \lambda_c$ [17]. The behavior of the field theory in the neighborhood of $\lambda = \lambda_c$ is the most difficult to analyze, since there are no known asymptotic expansions of the Green's functions or the mass about this point. It is not even known mathematically whether the mass is monotonic increasing for all $\lambda > \lambda_c$. However, physicists have a qualitative picture based on the assumed fractional power-law behavior $m(\lambda) \sim |\lambda_c - \lambda|^\nu$ above or below the critical point, where the exponent ν is a function of the dimension. The critical coupling λ_c corresponds to the largest effective forces between the particles in the theory. In this example it is known that there is a particle interpretation for $\lambda < \lambda_c$ and for $\lambda \gg \lambda_c$. While one expects this behavior for all $\lambda \neq \lambda_c$, and not for $\lambda = \lambda_c$ in two dimensions, this remains to be proved. See [15] for further discussion and for references.

Just as the interaction $\lambda\varphi^4 + \varphi^2$ with large coupling $\lambda \gg 1$ has two solutions, the interaction $\lambda'\varphi^4 - \varphi^2$, with small coupling $0 < \lambda' \ll 1$ also has two solutions. Each solution has a mass gap and a unique vacuum. The solutions for $\lambda'\varphi^4 - \varphi^2$ correspond to a field localized in one of two minima of the quartic energy density, with minima separated by $O(\lambda'^{-1/2})$. In this example, the different solutions for a given λ break the $\varphi \to -\varphi$ symmetry of the Hamiltonian. The proof of these facts was implemented by developing a cluster expansion about a quadratic interaction approximating one of the minima of the potential, and obtaining a sufficiently small estimate of the probability to tunnel into the other ground state. This expansion also leads to the mass gap $\Delta > 0$. A choice of boundary conditions selects a particular ground state Ω_+ or Ω_-. This choice both breaks $\varphi \to -\varphi$ symmetry of the energy density, and also leads to the expectation $\langle \Omega_\pm, \varphi\Omega_\pm \rangle = \pm\,(2\lambda')^{-1/2} + O(\lambda')$. Each solution has a mass gap $m = 2 + O(\lambda'^2)$, see [17].

While this example with a quartic potential exhibits both a phase transition and symmetry breaking, these concepts are not necessarily connected with one another. In theories with an energy density that is a sum of two polynomials $\mathcal{P}(\varphi) + \mathcal{Q}(\varphi)$, with $\deg \mathcal{Q} < \deg \mathcal{P}$, one can vary the coefficients of the polynomial $\mathcal{Q}(\varphi)$ in order to obtain a wide variety of phases [21].

4.3 Possible Approaches to the Yang-Mills Problem

One approach to the existence problem for four-dimensional gauge theory is to begin with a lattice regularization and to demonstrate the existence of a limit as the lattice spacing tends to zero. The gauge invariant action [39] can be defined on a toroidal Euclidean space-time, yielding a well-defined path integral. From this point of view, one must verify the existence of limits of appropriate expectations of gauge-invariant observables as the lattice spacing tends to zero and the volume tends to infinity.

Because four-dimensional gauge theory is a theory in which the mass gap is not visible classically, to demonstrate it may require a non-classical change of variables, known to physicists as a duality transformation. An example in which a duality has been used to establish a mass gap is the problem of a Coulomb gas, where the mass gap is known as Debye screening. A neutral Coulomb gas is composed of charges that at a microscopic level interact pairwise with a force determined by the long-range Coulomb potential (the Green's function for the Laplacian). But macroscopically, test charges interact by a force that decays exponentially with the distance between them. The mathematical proof of this phenomenon relies on a duality transformation relating the long-range Coulomb interaction of particles with the short-range $\cos(\lambda\varphi)$ (Sine-Gordon) interaction of fields. This duality, investigated by Stratonovich in 1957, forms the basis of the proof of Debye screening [4].

One view of the mass gap in Yang-Mills theory involves it as arising from the quartic potential $(A\wedge A)^2$ in the action, where $F = dA + gA\wedge A$, see [8]. This elementary picture can be interpreted as arising from curvature space of connections, see [34]. In fact, certain quantum mechanics problems with potentials having flat directions (directions for which the potential remains bounded as $|x| \to \infty$) do lead to bound states [33]. However, a mathematical analysis leading to such an effective potential in Yang-Mills theory remains currently out of reach.

A prominent speculation about a duality that might shed light on dynamical properties of four-dimensional gauge theory involves the $1/N$ expansion [20]. It is suspected that four-dimensional quantum gauge theory with gauge group $SU(N)$ (or $SO(N)$, or $Sp(N)$) may be equivalent to a string theory with $1/N$ as the string coupling constant. Such a description might give a clear-cut explanation of the mass gap and confinement, and perhaps a good starting point for a rigorous proof (for sufficiently large N). There has been surprising progress along these lines for certain strongly coupled four-dimensional gauge systems with matter [26] but as of yet, no effective approach to the pure four-dimensional gauge theory. Investigations of supersymmetric theories and string theories have given a variety of other approaches to understanding the mass gap in certain four-dimensional gauge theories with matter fields; for example, see [32].

References

[1] Tadeusz Balaban, Ultraviolet stability of three-dimensional lattice pure gauge field theories, *Commun. Math. Phys.* **102** (1985), 255–275.

[2] Tadeusz Balaban, Renormalization group approach to lattice gauge field theories. I: generation of effective actions in a small field approximation and a coupling constant renormalization in 4D, *Commun. Math. Phys.* **109** (1987), 249–301.

[3] Tadeusz Balaban, David Brydges, John Imbrie, and Arthur Jaffe, The mass gap for Higgs models on a unit lattice, *Ann. Physics* **158** (1984), 281–319.

[4] David Brydges and Paul Federbush, Debye screening, *Commun. Math. Phys.* **73** (1980), 197–246.

[5] M. Creutz, Monte Carlo Study of Quantized $SU(2)$ Gauge Theory, *Phys. Rev.* **D21** (1980), 2308–2315.

[6] J. Dimock and J.-P. Eckmann, On the bound state in weakly coupled $\lambda(\varphi^6 - \varphi^4)_2$, *Commun. Math. Phys.* **51** (1976), 41–54.

[7] Joel Feldman and Konrad Osterwalder, The Wightman axioms and the mass gap for weakly coupled φ_3^4 quantum field theories, *Ann. Physics* **97** (1976), 80–135.

[8] Richard Feynman, The quantitative behavior of Yang-Mills theory in $2 + 1$ dimensions, *Nucl. Phys.* **B188** (1981), 479–512.

[9] Jürg Fröhlich, Konrad Osterwalder, and Erhard Seiler, On virtual representations of symmetric spaces and their analytic continuation, *Ann. of Math.* **118** (1983), 461–489.

[10] Kryzstof Gawedzki, Renormalization of a non-renormalizable quantum field theory, *Nucl. Phys.* **B262** (1985), 33–48.

[11] James Glimm and Arthur Jaffe, The $\lambda\varphi_2^4$ quantum field theory without cut-offs, III. The physical vacuum, *Acta Math.* **125** (1970), 203–267.

[12] James Glimm and Arthur Jaffe, Positivity of the φ_3^4 Hamiltonian, *Fortschr. Phys.* **21** (1973), 327–376.

[13] James Glimm and Arthur Jaffe, φ_2^4 quantum field model in the single phase region: differentiability of the mass and bouns on critical exponents, *Phys. Rev.* **10** (1974), 536–539.

[14] James Glimm and Arthur Jaffe, A remark on the existence of φ_4^4, *Phys. Rev. Lett.* **33** (1974), 440–442.

[15] James Glimm and Arthur Jaffe, *Quantum Physics*, Second Edition, Springer Verlag 1987, and *Selected Papers, Volumes I and II*, Birkhäuser Boston, 1985. (Volume II includes reprints of [11, 12, 16, 17].)

[16] James Glimm, Arthur Jaffe, and Thomas Spencer, The Wightman axioms and particle structure in the $P(\varphi)_2$ quantum field model, *Ann. of Math.* **100** (1974), 585–632.

[17] James Glimm, Arthur Jaffe, and Thomas Spencer, A convergent expansion about mean field theory, I. *Ann. Phys.* **101** (1976), 610–630, and II. *Ann. Phys.* **101** (1976), 631–669.

[18] D. J. Gross and F. Wilczek, Ultraviolet behavior of non-abelian gauge theories, *Phys. Rev. Lett.* **30** (1973), 1343–1346.

[19] F. Guerra, L. Rosen, and B. Simon, Correlation inequalities and the mass gap in $\P(\varphi)_2$, III. Mass gap for a class of strongly coupled theories with non-zero external field, *Commun. Math. Phys.* **41** (1975), 19–32.

[20] G. 't Hooft, A planar diagram theory for strong interactions, *Nucl. Phys.* **B72** (1974), 461–473.

[21] John Imbrie, Phase diagrams and cluster expansions for low temperature $\mathcal{P}(\varphi)_2$ models, *Commun. Math. Phys.* **82** (1981), 261–304 and 305–343.

[22] Claude Itzykson and Jean-Bernard Zuber, *Quantum Field Theory*, McGraw-Hill, New York, 1980.

[23] Arthur Jaffe, Constructive quantum field theory, in *Mathematical Physics*, edited by T. Kibble, World Scientific, Singapore 2000.

[24] Jacques Magnen, Vincent Rivasseau, and Roland Sénéor, Construction of YM_4 with an infrared cutoff, *Commun. Math. Phys.* **155** (1993), 325–383.

[25] Jacques Magnen and Roland Sénéor, Yukawa quantum field theory in three dimensions, in *Third International Conference on Collective Phenomena*, J. Lebowitz et. al. editors, New York Academy of Sciences 1980.

[26] J. Maldacena, The large N limit of superconformal field theories and supergravity, Adv. Theor. Math. Phys. **2** (1998), 231–252.

[27] O. McBryan and J. Rosen, Existence of the critical point in φ^4 field theory, *Commun. Math. Phys.* **51** (1976), 97–105.

[28] Edward Nelson, Quantum fields and Markoff fields, in *Proc. Sympos. Pure Math. XXIII 1971*, pp. 413–420, Amer. Math. Soc., Providence, R.I. 1973.

[29] Konrad Osterwalder and Robert Schrader, Axioms for Euclidean Green's functions, *Commun. Math. Phys.* **31** (1973), 83–112, and *Commun. Math. Phys.* **42** (1975), 281–305.

[30] H. D. Politzer, Reliable perturbative results for strong interactions? *Phys. Rev. Lett.* **30** (1973), 1346–1349.

[31] A. Salam, Weak and electromagnetic interactions, pp. 367–377 in *Svartholm: Elementary Particle Theory, Proceedings of The Nobel Symposium held in 1968 at Lerum, Sweden,* Stockholm 1968.

[32] N. Seiberg and E. Witten, Electric-magnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang-Mills theory, *Nucl. Phys.* **B426** (1994), 19–52.

[33] Barry Simon, Some quantum operators with discrete spectrum but classically continuous spectrum, *Ann. Phys.* **146** (1983), 209–220.

[34] I. M. Singer, The geometry of the orbit space for nonabelian gauge theories, *Phys. Scripta* **24** (1981), 817–820.

[35] T. Spencer and F. Zirilli, Scattering states and bound states in $\lambda\mathcal{P}(\varphi)_2$, *Commun. Math. Phys.* **49** (1976), 1–16.

[36] Raymond Streater and Arthur Wightman, *PCT, Spin and Statistics and all That*, W. A. Benjamin, New York 1964.

[37] Kurt Symanzik, Euclidean quantum field theory, in *Local Quantum Theory*, pp. 152–226, R. Jost, Editor, Academic Press, New York 1969.

[38] S. Weinberg, A model of leptons, *Phys. Rev. Lett.* **19** (1967), 1264–1266.

[39] K. G. Wilson, Quarks and strings on a lattice, in *New Phenomena In Subnuclear Physics*, Proceedings of the 1975 International School of Subnuclear Physics, Erice, Sicily, ed. A. Zichichi, Plenum Press, New York 1977.

[40] Edward Witten, Supersymmetric Yang-Mills theory on a four-manifold, J. Math. Phys. **35** (1994), 5101–5135.

Harvard University
California Institute of Technology