The Yang–Mills free energy

Sourav Chatterjee

Quantum Yang–Mills theories

- Quantum Yang–Mills theories are the basic components of the Standard Model of quantum mechanics.
- In spite of huge strides made in the 80's, there are still no rigorous mathematical definitions of these objects in dimensions three and four.
- Dimension four is the most important case.
- The problem of constructing quantum Yang–Mills theories was posed as the Yang–Mills existence and mass gap problem in the list of seven millennium prize problems by the Clay Institute in the year 2000.

Connection form

- A quantum YM theory starts with a compact Lie group, called the gauge group.
- In this talk, the gauge group is $U(N)$, the group of unitary matrices of order N.
- Recall that the Lie algebra $\mathfrak{u}(N)$ of the Lie group $U(N)$ is the set of all $N \times N$ skew-Hermitian matrices.
- A $U(N)$ connection form on $\mathbb{R}^d$ is a smooth map from $\mathbb{R}^d$ into $\mathfrak{u}(N)^d$.
- If A is a $U(N)$ connection form, its value $A(x)$ at a point x is a d-tuple $(A_1(x), \ldots, A_d(x))$ of skew-Hermitian matrices.
- In the language of differential forms, $A = \sum_{j=1}^d A_j dx_j$.

Curvature form

- The curvature form F of a connection form A is the 2-form $F = dA + A \wedge A$.
- Explicitly, $F(x)$ is a $d \times d$ array of skew-Hermitian matrices of order N, whose $(j,k)^{\mathrm{th}}$ entry is the matrix

$$F_{jk}(x) = \frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k} + [A_j(x), A_k(x)]\,.$$

Yang–Mills action

- Let $\mathcal{A}$ be the space of all $U(N)$ connection forms on $\mathbb{R}^d$.
- The Yang–Mills action on $\mathcal{A}$ is the function

$$S_{\mathrm{YM}}(A) := -\int_{\mathbb{R}^d} \mathrm{Tr}(F \wedge *F)\,,$$

  where F is the curvature form of A and $*$ denotes the Hodge $*$-operator.
- Explicitly,

$$S_{\mathrm{YM}}(A) = -\int_{\mathbb{R}^d} \sum_{j,k=1}^{d} \mathrm{Tr}(F_{jk}(x)^2)\,dx\,.$$

Physics definition of Euclidean Yang–Mills theory

$U(N)$ quantum Yang–Mills theory in d-dimensional Euclidean spacetime is informally described as the probability measure

$$d\mu(A) = \frac{1}{Z}\exp\left(-\frac{1}{4g^2}S_{\mathrm{YM}}(A)\right)\mathcal{D}A\,,$$

where A belongs to the space $\mathcal{A}$ of all $U(N)$ connection forms, S_{YM} is the Yang–Mills functional,

$$\mathcal{D}A = \prod_{j=1}^{d}\prod_{x\in\mathbb{R}^d} d(A_j(x))$$

is "infinite dimensional Lebesgue measure" on $\mathcal{A}$, g a positive coupling constant, and Z is the normalizing constant (partition function) that makes this a probability measure.

The Yang–Mills existence problem

- The physics definition of Euclidean Yang–Mills theory is not mathematically valid, due to the non-existence of infinite dimensional Lebesgue measure on $\mathcal{A}$.
- The Yang–Mills existence problem has two parts: First, give a rigorous mathematical definition of Euclidean YM theory. Second, extend the theory to Minkowski spacetime by Wick rotation.
- The second part can be deduced from the first by standard tools from constructive quantum field theory if the Euclidean theory can be shown to satisfy certain properties (Wightman axioms or Osterwalder-Schrader axioms).

Lattice gauge theories

- In 1974, Kenneth Wilson introduced a discrete approximation of Euclidean YM theory, that is now known as lattice gauge theory.
- Lattice gauge theories are well-defined probability measures on subsets of $\mathbb{Z}^d$.

Wilson's discretization

- Suppose that $A = \sum_{j=1}^d A_j dx_j$ is a $U(N)$ connection form on $\mathbb{R}^d$.
- Discretize $\mathbb{R}^d$ as $\epsilon\mathbb{Z}^d$.
- Let $e_1, \ldots, e_d$ be the standard basis.
- For any $x \in \epsilon\mathbb{Z}^d$ and $1 \le j \le d$, let $U(x, x+\epsilon e_j) := e^{\epsilon A_j(x)}$.
- Let $U(y,x) := U(x,y)^{-1}$ for any edge (x,y) of $\epsilon\mathbb{Z}^d$.
- Take any x and $1 \le j < k \le d$. Let x_1, x_2, x_3, x_4 be the four vertices x, $x+\epsilon e_j$, $x+\epsilon e_j+\epsilon e_k$ and $x+\epsilon e_k$.
- Define $U(x,j,k)$

$$
\begin{aligned}
&:= U(x_1,x_2)U(x_2,x_3)U(x_3,x_4)U(x_4,x_1)\\
&= e^{\epsilon A_j(x_1)} e^{\epsilon A_k(x_2)} e^{-\epsilon A_j(x_4)} e^{-\epsilon A_k(x_1)} .
\end{aligned}
$$

- Observe that

$$
A_k(x_2) = A_k(x+\epsilon e_j) \approx A_k(x) + \epsilon \frac{\partial A_k}{\partial x_j},
$$

$$
A_j(x_4) = A_j(x+\epsilon e_k) \approx A_j(x) + \epsilon \frac{\partial A_j}{\partial x_k}.
$$

- Recall Baker–Campbell–Hausdorff formula

$$
e^B e^C = e^{B+C+\frac{1}{2}[B,C]+\cdots}.
$$

- Using the last two displays, one can show that

$$
\begin{aligned}
&\mathrm{Re}(\mathrm{Tr}(I - U(x,j,k)))\\
&\approx -\frac{\epsilon^4}{2}\mathrm{Tr}\biggl(\frac{\partial A_k}{\partial x_j} - \frac{\partial A_j}{\partial x_k} + [A_j(x), A_k(x)]\biggr)^2\\
&= -\frac{\epsilon^4}{2}\mathrm{Tr}(F_{jk}(x)^2).
\end{aligned}
$$

Wilson's definition of $U(N)$ lattice gauge theory

- Let $B_n := \{0, 1, \ldots, n-1\}^d$.
- On each edge (x, y) of ϵB_n, attach a unitary matrix $U(x, y)$ with the constraint that $U(y, x) = U(x, y)^{-1}$.
- Any such assignment of unitary matrices to edges will be called a configuration.
- Define $U(x, j, k)$ as in the previous slide, and let $\mu_{n,\epsilon,g}$ be the probability measure on the set of all configurations that has density

$$\frac{1}{Z(n, \epsilon, g)} \exp\biggl(-\frac{1}{g^2 \epsilon^{4-d}} \sum_{x,j,k} \mathrm{Re}(\mathrm{Tr}(I - U(x, j, k)))\biggr)$$

  with respect to the product Haar measure. Here $Z(n, \epsilon, g)$ is the partition function (normalizing constant).
- The probability measure $\mu_{n,\epsilon,g}$ is called $U(N)$ lattice gauge theory on ϵB_n with coupling strength g.

From lattice gauge to YM

- The lattice gauge theory defined in the previous slide is a well-defined probability measure on the set of configurations on ϵB_n.
- Euclidean $U(N)$ Yang–Mills theory is supposed to be the scaling limit of the above lattice gauge theory as $\epsilon \to 0$ and $n \to \infty$.
- If $n\epsilon \to R$, then we should get YM theory in a continuum box of side-length R. If $n\epsilon \to \infty$, then we should get YM theory in the whole of $\mathbb{R}^d$.
- No one knows what the limit object is, except in 2D.

2D Yang–Mills

- Mathematicians have been successful in constructing and understanding 2D Yang–Mills theories, sometimes as scaling limits of lattice gauge theories, sometimes in other ways.
- Some of the main contributors: Brydges, Fröhlich, Seiler, Bralić, Klimek, Kondracki, Gross, King, Sengupta, Driver, Fine, Witten, Lévy. Long body of work, from the 80's to recent years.
- Has been extended to 2D surfaces of higher genus, with connections to topological quantum field theory.
- A variety of theoretical computations, such as calculation of expected values, can also be carried out with the available tools.

3D and 4D Yang–Mills

- No rigorous construction in 3D and 4D, except for 3D $U(1)$ Yang–Mills theory, which was constructed by Gross (1983) and more completely by King (1986).
- For non-Abelian YM theories in 3D and 4D, the best known results are qualitative in nature, coming in the form of inequalities and compactness theorems, mainly for the partition function.
- The most famous results are due to Bałaban, who proved ultraviolet stability of 3D and 4D non-Abelian lattice gauge theories in a sequence of 14 long papers over a period of 6 years from 1983 to 1989.
- Ultraviolet stability is a form of compactness that implies the existence of certain subsequential limits. I do not have the time to define or discuss this in this talk.
- Other important results are due to Federbush (1986 - 1989), and Magnen, Rivasseau and Sénéor (1993).

Limit of the partition function

- In the absence of a limit object, one can at least get an understanding of the limiting behavior of the partition function $Z(n, \epsilon, g)$.
- It would be a significant breakthrough to be able to understand the exact asymptotics of $Z(n, \epsilon, g)$ as $\epsilon \to 0$ and $n\epsilon \to \infty$.

Exact asymptotics versus leading term

- For example, the exact asymptotics of $n!$ are given by Stirling's formula
$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$
whereas the leading term in the exponent is given by
$$n! = \exp\left(n\log\frac{n}{e} + o(n)\right).$$
- Large deviations theory, for instance, is concerned mainly with understanding the leading term. Same in statistical mechanics. In number theory, on the other hand, exact asymptotics are often very important.
- Why do we need the exact asymptotics of the partition function, instead of the leading term? Because one would eventually want to analytically continue and understand the behavior of the limit in Minkowski spacetime.

The main result

- Let $d = 3$. Recall the box $B_n = \{0, 1, \ldots, n-1\}^3$ and $U(N)$ lattice gauge theory on ϵB_n with coupling strength g.
- Let $Z(n, \epsilon, g)$ be the partition function of this theory.
- $\log Z(n, \epsilon, g)$ is called the free energy.
- The main result of this talk is an explicit formula for the leading term of the free energy as $\epsilon \to 0$ and $n\epsilon \to \infty$. Displayed in the next slide.
- This may be the first rigorous mathematical computation of any kind for Yang–Mills theories in dimensions higher than two.
- A similar formula for the 4D theory is also obtained, but only if g is also sent to zero.

The leading term of the free energy

Theorem (C., 2016)

Let $Z(n, \epsilon, g)$ be the partition function of 3D $U(N)$ lattice gauge theory in the box ϵB_n. Then as $\epsilon \to 0$ and $n\epsilon \to \infty$,

$$\log Z(n, \epsilon, g) = n^3 \left(N^2 \log(g^2 \epsilon) + 2 \log \left(\frac{\prod_{j=1}^{N-1} j!}{(2\pi)^{N/2}} \right) + N^2 K_3 \right) + o(n^3),$$

where K_3 is a numerical constant that will defined in the next slide.

Definition of K_3

- Let $d = 3$. Let E_n be the set of all positively oriented edges of B_n.
- For $t \in \mathbb{R}^{E_n}$ and $(x, y) \in E_n$, let $t(y, x) := -t(x, y)$.
- Take any $x \in B_n$ and $1 \le j < k \le 3$. Let x_1, x_2, x_3, x_4 be the four vertices x, $x + e_j$, $x + e_j + e_k$ and $x + e_k$.
- Define

$$t(x, j, k) := t(x_1, x_2) + t(x_2, x_3) + t(x_3, x_4) + t(x_4, x_1)$$

  provided that $x_1, x_2, x_3, x_4 \in B_n$.
- Let

$$M_n(t) := \sum_{x,j,k} t(x, j, k)^2 .$$

- Then M_n is a quadratic form on $\mathbb{R}^{E_n}$.
- Let E_n^0 be the subset of E_n consisting of all edges (x, y) that have one of the following forms:
  (1) $x = (a, 0, 0)$, $y = (a + 1, 0, 0)$.
  (2) $x = (a, b, 0)$, $y = (a, b + 1, 0)$.
  (3) $x = (a, b, c)$, $y = (a, b, c + 1)$.
- Let M_n^0 denote the restriction of M_n to the subspace of $\mathbb{R}^{E_n}$ consisting of all t such that $t(x, y) = 0$ for each $(x, y) \in E_n^0$.
- Define

$$K_3 := \lim_{n\to\infty} \frac{-\log\det M_n^0}{2n^3} .$$

  (Existence and finiteness of the limit is proved in the preprint.)

Some remarks about the proof

- The proof involves an interesting interplay of random matrix theory, Selberg-type integral formulas, properties of Gaussian measures and bare-hands probability theory.
- The term

$$\log\left(\frac{\prod_{j=1}^{N-1} j!}{(2\pi)^{N/2}}\right)$$

  arises from a passage to the Lie algebra from the Lie group. This is the part that involves random matrix theory and Selberg integrals.
- This term is closely related to the volume of $U(N)$, recently also appearing in a preprint of Diaconis and Forrester, where the volume is computed using an old formula of Hurwitz.
- The term K_3 arises from a Gaussian integral, therefore it is not surprising that a determinant is involved.

Main ideas in the proof

- When ϵ is small, the coupling parameter $g^2\epsilon$ of 3D lattice gauge theory is also small.
- The smallness of $g^2\epsilon$ forces the theory to behave like a Gaussian process at very small scales. This is the first step of the proof.
- The second step is to show that the leading term in the free energy is determined wholly by small scale behavior.
- The main challenge lies in proving that the cumulative effect of the larger scales is of smaller order than the leading term.
- An important point is that the proof avoids the highly complex renormalization methods of earlier works.

Some details

- Consider $U(N)$ lattice gauge theory on a box B_n of side-length n in $\epsilon\mathbb{Z}^3$, with coupling strength g.
- Let Z_n be the partition function. The free energy is $F_n = n^{-3} \log Z_n$.
- We get an upper bound for F_n and then a matching lower bound.
- Let $\phi(p)$ be the contribution to the Hamiltonian from a plaquette p. Note that $\phi(p) \geq 0$.
- Thus, if some plaquettes are dropped from the Hamiltonian, Z_n increases.
- Choose some m that divides n. Divide B_n as a disjoint union of translates of B_m, and remove the plaquettes that touch the boundaries of multiple boxes.
- The resulting integral breaks up as a product of integrals.
- From this it follows that $F_n \leq F_m$.

Details of the upper bound

- The inequality $F_n \leq F_m$ reduces the upper bound problem to understanding the behavior of F_m as m remains fixed (or grows very slowly) and $\epsilon \to 0$.
- If ϵ is small, one can argue that with high probability, any three of the four unitary matrices attached to the edges of the plaquette approximately determine the fourth.
- If m is not too large, then the above property can be used to deduce inductively that the matrices attached to a subset of the edges in B_m approximately determine the matrices attached to all the edges in B_m (approximate axial gauge fixing).
- This allows us to carry out the integration for computing F_m by first fixing the values of the matrices attached to a subset of the edges, then using a kind of Laplace approximation.

Details of the lower bound

- Note that the argument for the upper bound hinged on the fact that we could replace n by $m \ll n$ and let m grow very slowly as $\epsilon \to 0$.
- For the lower bound, however, we cannot directly replace a large box by a small box, because the inequality goes in the opposite direction.
- For the lower bound, the main trick is to define a small ball A of certain special configurations.
- The integral over A can be computed by a Laplace approximation. Call this integral $Z_n(A)$.
- Note that $Z_n = p^{-1} Z_n(A)$, where $p =$ the probability of A.
- Thus, $Z_n(A)$ can be used as a surrogate for Z_n if $p = \exp(-o(n^3))$, since $\log Z_n$ is of order n^3.

Details of the lower bound, continued

- Now, p is the probability that roughly n^3 matrices are close to some prescribed values.
- It is therefore natural to expect that p behaves like $\exp(-Cn^3)$ for some constant C rather than like $\exp(-o(n^3))$.
- The reason why we can show that p behaves like $\exp(-o(n^3))$ is that the set of all matrices can be controlled by controlling a subset of size $o(n^3)$ when ϵ is small.

Summary

- Quantum Yang–Mills theories are the basic components of the Standard Model of quantum mechanics.
- As yet undefined rigorously in dimensions higher than two. 4D is the most important case.
- Lattice gauge theories are well-defined discrete approximations of these hypothetical objects.
- The asymptotic behavior of the partition function of a lattice gauge theory, as the lattice spacing goes to zero, is an important object of interest.
- The logarithm of the partition function is called the free energy.
- The main result presented in this talk gives an explicit formula for the leading term of the free energy of 3D $U(N)$ lattice gauge theory as the lattice spacing goes to zero.
- This may be the first rigorous mathematical computation for YM theories in any dimension higher than two.
- The proof is based on a new technique that avoids renormalization.
- There is considerable interest in computing the full asymptotics of the free energy in 4D (instead of only the leading term, as is done here).