Yang-Mills on the lattice: New results and open problems

Sourav Chatterjee

Part 1: Overview

Quantum Yang–Mills theories

- Quantum gauge theories, also known as quantum Yang–Mills theories, are components of the Standard Model of quantum mechanics.
- In spite of many decades of research, physically relevant quantum gauge theories have not yet been constructed in a rigorous mathematical sense.
- The most popular approach to solving this problem is via the program of constructive field theory.
- In this approach, one starts with a statistical mechanical model on the lattice; the next step is to pass to a continuum limit of this model; the third step is to show that the continuum limit satisfies certain 'axioms'; if these axioms are satisfied, then there is a standard machinery which allows the construction of a quantum field theory.
- Taking this program to its completion is one of the Clay millennium problems.

Lattice gauge theories

- The statistical mechanical models considered in the first step of the above program are known as lattice gauge theories.
- A lattice gauge theory may be coupled with a Higgs field, or it may be a pure lattice gauge theory.
- We will only deal with pure lattice gauge theories in this talk.
- A pure lattice gauge theory is characterized by its gauge group (usually a compact matrix Lie group), the dimension of spacetime, and a parameter known as the coupling strength.
- These theories on their own, even without passing to the continuum limit or constructing the quantum theory, can yield substantial physically relevant information.
- Enormous amount of computational effort is spent in laboratories around the world to numerically predict masses of elementary particles and other quantities using lattice gauge theories.

Goals of this two-part talk

I will discuss the following topics about lattice gauge theories, explain the open problems related to these topics, and survey the known rigorous results:

- Mass gap.
- Quark confinement.
- Gauge-string duality.

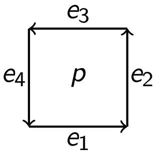
I will not discuss the problem of constructing continuum limits of lattice gauge theories, since a proper discussion of that would take several more lectures.

Notation

- Let $n \geq 1$ and $d \geq 2$ be two integers.
- Let G be a closed connected subgroup of the unitary group $U(n)$.
- Let $B_N := [-N, N]^d \cap \mathbb{Z}^d$.
- Let E_N be the set of positively oriented nearest-neighbor edges of B_N.
- Let Ω_N be the set of all functions from E_N into G.
- If $\omega \in \Omega_N$ and e is a negatively oriented edge, we define $\omega_e := \omega_{e^{-1}}^{-1}$, where e^{-1} is the positively oriented version of e.

Plaquettes

- A plaquette in $\mathbb{Z}^d$ is a set of four directed edges that form the boundary of a square.
- Let P_N be the set of all plaquettes in B_N.
- Given some $p \in P_N$ and $\omega \in \Omega_N$, we define ω_p as follows.
- Write p as a sequence of directed edges e_1, e_2, e_3, e_4, each one followed by the next.



- Let $\omega_p := \omega_{e_1}\omega_{e_2}\omega_{e_3}\omega_{e_4}$.
- Although there are ambiguities in this definition about the choice of e_1 and the direction of traversal, that is not problematic because we will only use the quantity $\Re(\mathrm{Tr}(\omega_p))$, which is not affected by these ambiguities.

Hamiltonian

- Let ∂E_N denote the set of positively oriented boundary edges of B_N.
- Let $\partial \Omega_N$ denote the set of all functions from ∂E_N into G.
- An element of $\partial \Omega_N$ will be called a boundary condition.
- Let $E_N^\circ := E_N \setminus \partial E_N$ be the set of positively oriented interior edges of B_N.
- Let Ω_N° be the set of all functions from E_N° into G. An element of Ω_N° will be called a configuration.
- Take any boundary condition δ. For each $\omega \in \Omega_N^\circ$, extend ω to an element $\widetilde{\omega} \in \Omega_N$ by defining

$$\widetilde{\omega}_e := \begin{cases} \omega_e & \text{if } e \in E_N^\circ, \\ \delta_e & \text{if } e \in \partial E_N. \end{cases}$$

- Define the Hamiltonian

$$H_{N,\delta}(\omega) := \sum_{p \in P_N} \Re(\mathrm{Tr}(\widetilde{\omega}_p)).$$

Definition of lattice gauge theory

- The pure lattice gauge theory on B_N with gauge group G, coupling parameter β, and boundary condition δ, is the probability measure $\mu_{N,\delta,\beta}$ on Ω_N° defined as

$$d\mu_{N,\delta,\beta}(\omega) = Z_{N,\delta,\beta}^{-1} e^{\beta H(\omega)} d\lambda_N(\omega),$$

  where λ_N is the product Haar measure on Ω_N° and $Z_{N,\delta,\beta}$ is the normalizing constant.
- Here $\beta = 1/g_0^2$, where g_0 is the coupling strength.
- Given a measurable function $f : \Omega_N^\circ \to \mathbb{C}$, the expected value of the function under the above lattice gauge theory is the quantity

$$\langle f \rangle := \int_{\Omega_N^\circ} f(\omega) d\mu_{N,\delta,\beta}(\omega),$$

  provided that the integral on the right is well-defined.

Mass gap

- In lattice gauge theories, mass gap means exponential decay of correlations.
- The importance of mass gap arises from the fact that the rate of decay (i.e. the constant in the exponent) can be used to predict the masses of elementary particles.
- There is huge activity in this area, broadly known as lattice QCD. Supercomputers are employed to predict the masses and other properties of various elementary particles using lattice theories.

Mass gap: Known results

- Mass gap is relatively easy to establish in 2D theories, because 2D theories can be reduced to 1D problems using a procedure called gauge fixing.
- Mass gap is also easy to prove at small β using standard methods, such as series expansions or the methods of Dobrushin and Shlosman (1985).
- Abelian theories do not have mass gap at large β in 4D. This was proved for 4D $U(1)$ theory by Fröhlich and Spencer (1982), and for 4D theories with finite Abelian gauge groups by Kotecký and Shlosman (1982) and Borgs (1984).

Mass gap: Open problems

- In 3D and 4D, mass gap has not been established at large β for any theory.
- It is believed that 4D $SU(3)$ theory has mass gap at all large β. This completely open problem would have huge significance, if proved. (For example, it explains why protons and neutrons have mass.)
- Even showing any kind of correlation decay at large β for non-Abelian theories in 3D and 4D is open.

Quark confinement

- Quarks are the constituents of various elementary particles, such as protons and neutrons.
- It is an enduring mystery why quarks are never observed freely in nature.
- The problem of quark confinement has received a lot of attention in the physics literature, and yet the current consensus seems to be that a satisfactory theoretical explanation does not exist (let alone a rigorous proof!).
- Wilson (1974) argued that quark confinement is equivalent to showing that the relevant lattice gauge theory satisfies what's now known as Wilson's area law.

Area law

- Let π be a finite-dimensional unitary representation of the group G, and let χ_π be the character of π.
- Let ℓ be a closed loop in B_N, with directed edges $e_1, \ldots, e_k$.
- Given a configuration ω, the Wilson loop variable $W_\ell(\omega)$ is defined as

$$W_\ell(\omega) := \chi_\pi(\omega_{e_1}\omega_{e_2}\cdots\omega_{e_k}).$$

- The lattice gauge theory is said to satisfy Wilson's area law for the representation π if

$$|\langle W_\ell \rangle| \le C_1 e^{-C_2 \operatorname{area}(\ell)}$$

  for any rectangular loop ℓ, where C_1 and C_2 are positive constants that depend only on G, β, d and π, and $\operatorname{area}(\ell)$ is the area enclosed by ℓ.
- The analogous bound with area replaced by perimeter is known as perimeter law.

Why does area law imply quark confinement?

- Let $V(R)$ be the potential energy of a static quark-antiquark pair separated by distance R.
- QFT calculations imply that for a rectangular loop ℓ of side-lengths R and T in the continuum limit of 4D $SU(3)$ theory and a suitable representation π, $\langle W_\ell \rangle$ should behave like $e^{-V(R)T}$.
- So if the area law holds, then $V(R)$ grows linearly in the distance R between the quark and the antiquark. By the conservation of energy, this implies that the pair will not be able to separate beyond a certain distance.
- Renormalization group arguments predict that β has to be sent to infinity as the lattice spacing goes to zero to obtain the continuum limit of 4D non-Abelian theories. This indicates that we need the area law to hold at arbitrarily large values of β in 4D $SU(3)$ theory for it to imply confinement of quarks.

Area law: Basic facts

- It is easy to show that the area law holds at all β in any 2D theory, since gauge fixing can be used to reduce any 2D theory to a 1D model.
- Seiler (1978) proved an area law lower bound for any theory at any β.
- Simon and Yaffe (1982) proved a perimeter law upper bound for any theory at any β.

Area law: Deeper results

- Osterwalder and Seiler (1978) showed that the area law holds at small enough β (strong coupling) for any theory in any dimension.
- Guth (1980) and Fröhlich and Spencer (1982) showed that for 4D $U(1)$ theory, area law breaks down at large enough β; instead, perimeter law holds. This is known as the deconfinement transition for this theory.
- The deconfinement transition was physically expected, because 4D $U(1)$ theory is related to photons, which are not confined.
- Göpfert and Mack (1982) showed that the area law holds at all β for 3D $U(1)$ theory. This is still the only nontrivial case where the area law has been established at large β.

Area law: Various other theorems

- Fröhlich (1979) showed that confinement holds in $SU(n)$ theory if it holds in the corresponding $\mathbb{Z}_n$ theory.
- Durhuus and Fröhlich (1980) showed that confinement in a d-dimensional pure lattice gauge theory holds if there is exponential decay of correlations in a $(d-1)$-dimensional nonlinear σ model.
- Borgs and Seiler (1983) investigated confinement in lattice gauge theories at nonzero temperature, building on technology developed by Brydges and Federbush (1980).
- A toy model exhibiting a sharp transition from the confining to the deconfining regime was studied by Aizenman, Chayes, Chayes, Fröhlich and Russo (1983).

Area law: Recent progress

- Area law at small β for arbitrary loops (where area(ℓ) is the minimal surface area enclosed by ℓ) was established in large N limit of $SO(N)$ and $SU(N)$ theories by Chatterjee (2019), Jafarov (2016), and Chatterjee and Jafarov (2016).
- Chatterjee (2020a) computed the exact leading order behavior of Wilson loop expectations in 4D $\mathbb{Z}_2$ theory at large β.
- This result was extended to all 4D theories with finite Abelian gauge groups by Forsström, Lenells and Viklund (2020), and to all 4D theories with finite gauge groups by Cao (2020).
- Chatterjee (2020b) proved that exponential decay of correlations under arbitrary boundary conditions implies Wilson's area law if the gauge group has a nontrivial center, in any dimension.
- It is believed that 4D $SU(3)$ theory has exponential decay of correlations at all β. If indeed true, the above result would give a proof of quark confinement.
- I will say more about the above results in part 2 of the talk.

Area law: Open questions

- The main open question about the area law is to prove that it holds in 4D $SU(3)$ theory at arbitrarily large values of β.
- In fact, proving it for any other 4D non-Abelian theory would be tremendously interesting too.
- It is not expected to hold for 4D Abelian theories at large β, nor even for 4D non-Abelian theories with finite gauge groups.

Gauge-string duality

- Gauge theories (i.e. Yang–Mills theories) are theories of the quantum world. String theories are theories of gravity.
- It is a major goal of theoretical physics to make a connection between the above two.
- Although gravitational effects are negligible in the quantum world, there is one situation where both gravitational and quantum effects are important — in black holes. This is one of the practical motivations behind the quest for unifying gauge theories and string theories.
- Physicists have been aware of a duality between gauge theories and string theories since the 1970s. A concrete duality relation found by Maldacena (1997) kicked off a vast field of research, now known as gauge-string duality or gauge-gravity duality or AdS-CFT duality.

Gauge-string duality: Rigorous results

- Almost none of the results in gauge-string duality are rigorous, since all the physical results are for continuum theories, which are not rigorously defined.
- An old result of Brydges, Giffen, Durhuus and Fröhlich (1986) expresses Wilson loop expectations in $\mathbb{Z}_2$, $U(1)$ and $SU(2)$ theories (in any dimension) as weighted sums over surfaces, similar to the ones arising in string theories.
- Chatterjee (2019) showed that in the large N limit at small β, Wilson loop expectations of $SO(N)$ theory can be expressed as weight sums of trajectories in a certain string theory on the lattice.
- This result was extended to other theories by Jafarov (2016) and Chatterjee and Jafarov (2016).
- Basu and Ganguly (2018) showed that the lattice string theory from Chatterjee (2019) permits exact calculations in 2D.
- As far as I know, these are the only known rigorous results. Will say more in part 2.

Gauge-string duality: Open problems

- The main challenge is to prove duality in the continuum, because there are many techniques for dealing with string theories in the continuum. Lattice string theories, just like lattice gauge theories, are hard to handle.
- One approach would be to extend the lattice dualities to arbitrarily large β. Currently, it is not known how to do that.

Part 2: Some recent results

Computing Wilson loop expectations to leading order

Wilson loop expectations at large β

- For constructing continuum limits of non-Abelian theories, we need a very precise understanding of Wilson loop expectations as $\beta \to \infty$.
- Such a result was obtained for 4D $\mathbb{Z}_2$ theory in Chatterjee (2020). This will be presented in the next slide.
- Although this is an Abelian theory, it is a first step towards a more general understanding.
- Indeed, the result was soon extended to all 4D theories with finite Abelian gauge groups by Forsström, Lenells and Viklund (2020), and to all 4D theories with arbitrary finite gauge groups by Cao (2020).
- Cao's extension is highly nontrivial, involving substantial amounts of algebraic topology and other ingredients.
- The natural next step would be to extend it to compact Lie groups, but this seems to be very challenging.

Wilson loop expectation in 4D $\mathbb{Z}_2$ lattice gauge theory

- Call an edge $e \in \ell$ a corner edge if there is some other edge $e' \in \ell$ such that e and e' share a common plaquette.
- For example, a rectangular loop with length and width greater than one has exactly eight corner edges.

Theorem (Chatterjee, 2020a)

Consider 4D $\mathbb{Z}_2$ lattice gauge theory. There exists $\beta_0 > 0$ such that the following holds when $\beta \geq \beta_0$. Let ℓ be a non-self-intersecting loop. Let r be the number of edges in ℓ and let r_0 be the number of corner edges of ℓ. Then

$$|\langle W_\ell \rangle - e^{-2re^{-12\beta}}| \leq C_1 \left(e^{-2\beta} + \sqrt{\frac{r_0}{r}} \right)^{C_2},$$

where C_1 and C_2 are two positive universal constants.

(To put it more simply, if $\beta \gg 1$ and $r_0 \ll r$, and $r = \alpha e^{12\beta}$, then $\langle W_\ell \rangle \approx e^{-2\alpha}$.)

Cao's result

- Let the gauge group G be a finite group of $m \times m$ unitary matrices.
- Define $\Delta_G := \min_{g \neq I} \Re(\mathrm{Tr}(I) - \mathrm{Tr}(g))$, and let G_0 be the set of minimizers.
- Define
$$A := \frac{1}{|G_0|} \sum_{g \in G_0} g.$$

Theorem (Cao, 2020)

Let $\beta \geq \Delta_G^{-1}(1000 + 14 \log |G|)$. Let ℓ be a loop of length r. Let $X \sim \mathrm{Poisson}(r|G_0|e^{-6\beta\Delta_G})$. Then

$$|\langle W_\ell \rangle - \mathrm{Tr}(\mathbb{E}A^X)| \leq 10me^{-c(G)\beta}.$$

Here $c(G) > 0$ is a constant which depends only on G.

Explicit form of Cao's result

- There is a more explicit formula, as follows.
- Not hard to show that A is Hermitian and $\|A\|_{op} \leq 1$.
- Let $-1 \leq \lambda_1, \ldots, \lambda_m \leq 1$ be the eigenvalues of A. Then

$$\mathrm{Tr}(\mathbb{E}A^X) = \sum_{i=1}^{m} e^{-(1-\lambda_i)r|G_0|e^{-6\beta\Delta_G}}.$$

- In other words, if β is large and $r = \alpha e^{6\beta\Delta_G}$, then

$$\langle W_\ell \rangle \approx \sum_{i=1}^{m} e^{-(1-\lambda_i)|G_0|\alpha}.$$

- If $G = \mathbb{Z}_2$, then $m = 1$, $|G_0| = 1$, $\lambda_1 = -1$, $\Delta_G = 2$, and so the above reduces to $e^{-2\alpha}$.

Relation between mass gap and quark confinement

Strong mass gap implies quark confinement

- Main result in Chatterjee (2020b): *If the gauge group is compact, connected, and has a nontrivial center, then the presence of exponential decay of correlations under arbitrary boundary conditions implies that Wilson's area law holds.* (Precise statement in next few slides.)
- The nontriviality of the center is known as center symmetry.
- The exponential decay assumption stated above is stronger than usual mass gap, which means exponential decay under a specific boundary condition dictated by the transfer matrix of the theory. This is rather subtle and often not spelled out in the literature.
- To distinguish it from usual mass gap, I call the above condition strong mass gap.

Physical background

- There are theories which have both mass gap and center symmetry, and yet are believed to be deconfining.
- Therefore, the strengthening of the correlation decay assumption is necessary to deduce confinement.
- There is a longstanding belief in physics, originating in the work of 't Hooft (1978), that mass gap plus unbroken center symmetry implies confinement. (Note: 'Unbroken center symmetry' is not a rigorously defined concept.)
- The result from Chatterjee (2020b) shows that *strong mass gap plus center symmetry implies confinement*.
- The physical explanation (according to Ed Witten, personal communication) is that strong mass gap prevents breaking of center symmetry.
- Does $SU(3)$ theory have strong mass gap at all β? It is possible, but we cannot be sure without a proof.

Recall notation

- The gauge group G is a closed connected subgroup of the unitary group $U(n)$.
- $B_N = [-N, N]^d \cap \mathbb{Z}^d$.
- E_N is the set of positively oriented nearest-neighbor edges of B_N.
- Ω_N is the set of all functions from E_N into G.
- ∂E_N is the set of positively oriented boundary edges of B_N.
- $\partial \Omega_N$ is the set of all functions from ∂E_N into G.
- An element of $\partial \Omega_N$ is called a boundary condition.
- $E_N^\circ = E_N \setminus \partial E_N$ is the set of positively oriented interior edges of B_N.
- Ω_N° is the set of all functions from E_N° into G. An element of Ω_N° is called a configuration.

Assumptions

- Center symmetry: Assume that the center of G is nontrivial, and there is an element g_0 in the center such that $\pi(g_0) = cI$ for some $c \neq 1$, where I is the $m \times m$ identity matrix, m being the dimension of π.
- Strong mass gap: Say that two edges are neighbors if they both belong to some common plaquette. Say that a measurable map $f : \Omega_N^\circ \to \mathbb{R}$ is a *local function* supported on an edge $e \in E_N$ if $f(\omega)$ depends only on the values of ω_u for u that are neighbors of e. Given two local functions f and g, let $\mathrm{dist}(f, g)$ denote the Euclidean distance between the midpoints of their supporting edges. Assume that there are positive constants K_1 and K_2 depending only on G, β and d, and not on N or the boundary condition δ, such that for any local functions $f, g : \Omega_N^\circ \to [-1, 1]$,

$$|\langle fg\rangle - \langle f\rangle\langle g\rangle| \leq K_1 e^{-K_2 \,\mathrm{dist}(f,g)}.$$

Precise statement

Theorem (Chatterjee, 2020b)

Let G be a compact connected subgroup of $U(n)$ for some n, and let π be a finite-dimensional unitary representation of G. Take any $d \geq 2$ and $\beta \in \mathbb{R}$, and consider the lattice gauge theory on the cube B_N. Suppose that the center symmetry and strong mass gap assumptions are satisfied. Then there are positive constants C_1 and C_2 depending only on G, β, π and d, such that the following holds. Take any $N \geq 2$, any boundary condition δ on B_N, and any rectangular loop ℓ contained in $B_{N'}$ for some $N' \leq N/2$. Then

$$|\langle W_\ell \rangle| \leq C_1 e^{-C_2 \operatorname{area}(\ell)},$$

where area(ℓ) *is the area enclosed by ℓ. Moreover, there is a unique infinite volume Gibbs state, and the above bound holds for any rectangular loop ℓ if the expectation on the left is taken with respect to this Gibbs state.*

Remark

- The strong mass gap assumption is satisfied by any lattice gauge theory if β is small enough. This recovers the old result of Osterwalder and Seiler (1978) which says all theories are confining at small β; but it says much more, since we have no requirement that β be sufficiently small.
- In some sense, this result says that the area law holds in the entire strong coupling regime of lattice gauge theories, in the presence of center symmetry. It is possible that the strong coupling regime of 4D non-Abelian theories consists of the whole of $[0, \infty)$.
- This is the first rigorous result that throws light on the role of center symmetry in confinement, which is well-established in the theoretical physics discourse.

Gauge-string duality

Large N gauge theories: 't Hooft's approach

- Gauge groups such as $SU(5)$, $SU(3)$ and $SU(2) \times U(1)$ are the ones that are relevant for physical theories.
- However, theoretical understanding is difficult to achieve.
- 't Hooft (1974) suggested a simplification of the problem by considering groups such as $SU(N)$ where N is large.
- The $N \to \infty$ limit, after replacing β by $N\beta$, simplifies many theoretical problems. This is known as the 't Hooft limit.
- Calculation of normalizing constants by 't Hooft's planar diagram theory.
- Some aspects of the planar diagram theory for systems with a fixed number of vertices has been made rigorous by various authors (Guionnet, Ercolani, McLaughlin, Eynard, Jones, ...).

Gauge-string duality

- 't Hooft's approach pointed to a connection between gauge theories and string theories.
- In string theories, particles are replaced by one-dimensional strings.
- A string moving through time traces out a surface. From that perspective, string theories may be viewed as theories of random surfaces.
- The planar diagrams in 't Hooft's approach give a relation between computations in gauge theories and counting certain kinds of surfaces.

Gauge-string duality, contd.

- This connection was not very well understood in physics until the work of Maldacena (1997), which gave birth to the flourishing field of AdS-CFT duality, sometimes called gauge-gravity duality.
- This theory gives, among other things, an explicit way to write the expectations of Wilson loop variables in a certain supersymmetric gauge theory (the 'CFT' part) in the large N limit in terms of weighted sums over trajectories of strings in a high dimensional string theory (the 'AdS' part).
- None of this is rigorous math, however. Mainly because the underlying objects are not mathematically defined.
- It was shown in Chatterjee (2019) that in the large N limit of $SO(N)$ theory at small enough β, in any dimension, Wilson loop expectations can be written as sums over trajectories in a certain string theory on the lattice.

A string theory on the lattice

- Basic objects: Collections of finitely many loops in $\mathbb{Z}^d$, called 'strings'. Analogous to strings in the continuum.
- Strings can evolve in time according to certain rules.
- At each time step, only one loop in a string is allowed to be modified.
- Four possible modifications:
  - Positive deformation. (Addition of a plaquette without erasing edges.)
  - Negative deformation. (Addition or deletion of a plaquette that involves erasing at least one edge.)
  - Positive splitting. (Splitting a loop into two loops without erasing edges.)
  - Negative splitting. (Splitting a loop into two loops in a way that erases at least one edge.)

Trajectories

- The evolution of a string is called a trajectory.
- A trajectory is called vanishing if it ends in nothing in a finite number of steps.

Action of a trajectory

- String theories in the continuum prescribe actions for trajectories of strings, analogous to the Lagrangian action of classical mechanics.
- For example, in Bosonic string theory, the Nambu–Goto action is the area of the surface traced out by the string as it moves through time.
- We would like to define an action for a trajectory of strings in our lattice string theory.

Action of a trajectory, contd.

- The lattice string theory defined here has a parameter β.
- Depending on the value of β, each step of a trajectory is given a weight, as follows.
- Let m be the total number of edges in the string before the step is taken.
- The weight of the step is defined to be:

$$
\begin{aligned}
-\beta/m &\text{ if the step is a positive deformation;} \\
\beta/m &\text{ if the step is a negative deformation;} \\
-2/m &\text{ if the step is a positive splitting;} \\
2/m &\text{ if the step is a negative splitting.}
\end{aligned}
$$

- The weight or action of a vanishing trajectory X is defined to be the product of the weights of the steps in the trajectory. Denoted by $w_\beta(X)$.

Gauge-string duality in the 't Hooft limit

Theorem (Chatterjee, 2019)

There exists $\beta_0 > 0$ *such that the following is true. Let* ℓ *be a fixed loop in* $\mathbb{Z}^d$*. Let* $\langle W_\ell \rangle$ *denote the expectation of the Wilson loop variable* W_ℓ *in the* $SO(N)$ *lattice gauge theory in a region* $\Lambda_N \subseteq \mathbb{Z}^d$*, with coupling parameter* $N\beta$*. If* $|\beta| \le \beta_0$ *and* $\Lambda_N \uparrow \mathbb{Z}^d$ *as* $N \to \infty$*, then*

$$\lim_{N\to\infty} \frac{\langle W_\ell \rangle}{N} = \sum_{X \in \mathcal{X}(\ell)} w_\beta(X)\,,$$

where $\mathcal{X}(\ell)$ *is the set of all vanishing trajectories starting at* ℓ *and* $w_\beta(X)$ *is the action of* X *defined earlier. Moreover, the infinite sum on the right is absolutely convergent.*

Implication for Wilson's area law

- Recall that area(ℓ) was defined to be the minimum surface area enclosed by a loop ℓ.
- In the context of the lattice, one can take a 'surface' to mean a 2-chain in the standard cell complex of $\mathbb{Z}^d$.
- Then the notion of a surface having boundary ℓ is well-defined, if we look at ℓ as a 1-cycle.
- Since any 2-chain can be written as a linear combination of plaquettes with integer coefficients, the sum of the absolute values of these coefficients give a natural notion of the area of a lattice surface.

Implication for Wilson's area law, contd.

- Key observation, making use of the Poincaré lemma: The minimum number of steps required to gradually modify a loop (according the our deformations and splittings) into nothing, is equal to the minimum surface area enclosed by the loop.
- From this and the main theorem, it is not hard to deduce that

$$\lim_{N\to\infty} \frac{|\langle W_\ell \rangle|}{N} \le C_1 e^{-C_2 \text{area}(\ell)} .$$

- This is the first rigorous result that proves the area law upper bound for an arbitrary loop. All previous results were for rectangles.

Further developments

- The analogous result for $SU(N)$ theory was proved by Jafarov (2016).
- A full asymptotic series expansion for $\langle W_\ell \rangle$ in powers of $1/N$ was obtained by Chatterjee and Jafarov (2016). The terms in the expansion are similar sums of trajectories in string theories, except that for the k^{th} term, various components of the string are allowed to merge up to k times, thus tracing a surface of genus $\leq k$.
- Basu and Ganguly (2018) showed that the lattice string theory permits exact calculations in 2D. This result has not yet been extended to higher dimensions.
- A similar duality result has not been proved for large β. This would be of great interest, because one can then hope to pass to the continuum limit. String theories are mathematically tractable objects in the continuum.