Integrable lattice models from gauge theory

Edward Witten

These notes provide an introduction to recent work by Kevin Costello in which integrable lattice models of classical statistical mechanics in two dimensions are understood in terms of quantum gauge theory in four dimensions. This construction will be compared to the more familiar relationship between quantum knot invariants in three dimensions and Chern-Simons gauge theory. (Based on a Whittaker Colloquium at the University of Edinburgh and a lecture at Strings 2016 in Beijing.)

1. Preliminary remarks

Since the discovery of the Bethe ansatz in the early days of quantum mechanics, integrable quantum spin systems in $1 + 1$ dimensions and their close cousins of various sorts have been a topic of much fascination. Important advances have been made by many physicists and mathematicians, among them Onsager, Yang, Baxter, Lieb, Kruskal, Fadde'ev, Drinfeld, Miwa, Jimbo, and A. and Al. Zamolodchikov. The subject is so multi-faceted that no short summary can do justice to it.

In today's lecture, I will be describing a new approach to integrable lattice models of two-dimensional classical statistical mechanics, developed recently by Kevin Costello [1]. Costello's work has offered an essentially new perspective in which these models are understood in terms of four-dimensional gauge theory.[1] Arguably this perspective is in line with a vision relating theories in different dimensions that was offered many years ago by Michael Atiyah [4]. As we will see, Costello's work is also a close cousin of the relationship [5] between Chern-Simons gauge theory in three dimensions and the Jones polynomial of a knot.[2]

---

[1]Another and superficially quite different relation of some of the same models to gauge theory was discovered earlier by Nekrasov and Shatashvili [2, 3].

[2]Costello has actually described two related approaches to this subject, one in terms of a four-dimensional cousin of Chern-Simons gauge theory and one in terms

There are several different kinds of integrable systems in one space or two spacetime dimensions, including classical nonlinear PDE's, continuum quantum field theories, quantum spin chains, and classical lattice systems. However, these different types of model turn out to be closely related. (Among the many excellent sources where one can find different perspectives are [7–10].) Even though our main topic today will be lattice statistical mechanics in two dimensions, to motivate the ideas I will begin with integrable models of continuum quantum field theory in $1 + 1$ dimensions. The ideas that I will sketch were first used as a tool to directly construct a relativistic S-matrix by A. and Al. Zamolodchikov [11].

Fig. 1 is meant to be a spacetime picture of elastic scattering of two particles in $1 + 1$ dimensions. A particle of constant velocity is represented by a straight line, with a slope depending on the velocity. Because of conservation of energy and momentum, the outgoing particles go off at the same slope (same velocity) as the incoming particles. There are time delays that I have not tried to draw. The time delays mean that the outgoing lines are parallel to the ingoing ones, but displaced slightly inwards. This will not really affect our discussion.

In a typical relativistic quantum field theory, there are also particle production processes, which are a large part of what makes quantum field theory interesting. An example with two particles going to three is sketched in Fig. 2. The symmetries of typical relativistic field theories allow such processes and they happen all the time in the real world. In a $2 \to 3$ scattering event, in a massive theory, the incoming and outgoing lines can be assumed to all end or begin at a common point in spacetime, to within an error that depends on the range of the particle interactions and is reflected in the time delays.

However, in two spacetime dimensions, there are "integrable" field theories that have extra symmetries that commute with the velocity or momentum but move a particle in space by an amount that depends on its velocity. Then particle production is not possible. Starting with a spacetime history in which the incoming and outgoing lines meet at a common point in spacetime, a symmetry that moves the incoming and outgoing lines by a velocity-dependent amount will create a history such as that of Fig. 3 in which the outgoing particles could have had no common origin in spacetime.

---

of a twisted version of four-dimensional $\mathcal{N} = 1$ super Yang-Mills theory. Roughly speaking, the first approach, which we follow here, comes by integrating out some variables from the second. For a review of the second approach, see [6].
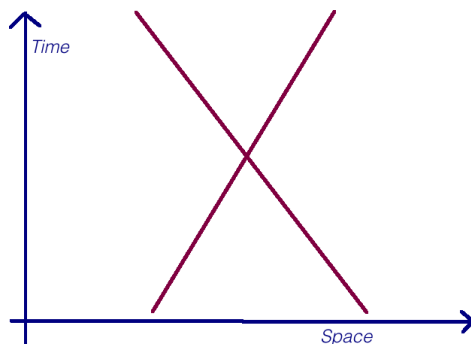
Figure 1: A spacetime picture of two-body elastic scattering in $1 + 1$ dimensions.
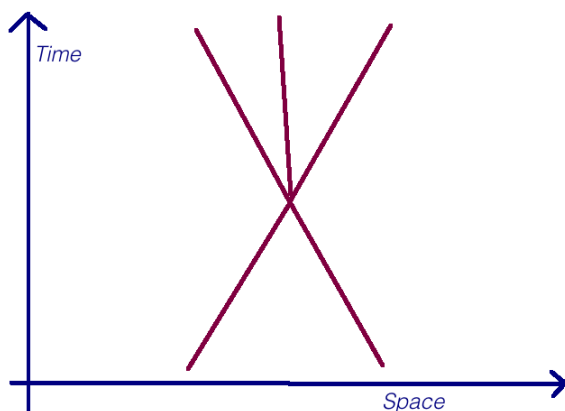

Figure 2: Inelastic scattering in $1 + 1$ dimensions with particle production.

By contrast, two particle scattering does happen even in integrable systems, since two lines in the plane do generically intersect, as in Fig. 4. But in an integrable theory, two particle scattering is purely *elastic*, in the sense that the initial and final particles have the same masses (this is true separately for the left- and right-moving particles). Otherwise, the initial and final velocities are different and consideration of a symmetry that moves particles in a velocity-dependent way again leads to a contradiction.
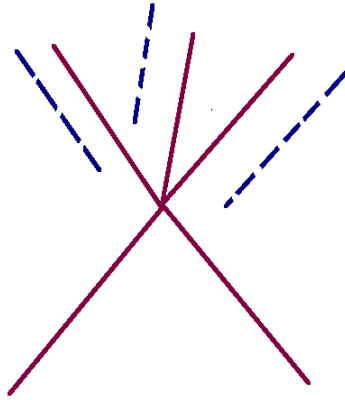
Figure 3: An inelastic $2 \to 3$ scattering process is sketched by the solid lines. In a theory that has a symmetry that moves particles by an amount that depends on their momentum, the application of such a symmetry would move the outgoing lines parallel to themselves, leaving their slope unchanged. (By combining the symmetry in question with an ordinary spacetime translation, we can assume that the incoming lines are unaffected, as drawn here.) After such a transformation, the outgoing lines – now showed as dotted lines – have no common point of intersection in the plane, let along a common point of intersection at which they also meet the incoming lines. This means that a $2 \to 3$ process is not possible in a theory with such a symmetry.
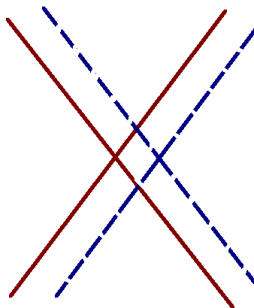

Figure 4: Two lines in the plane generically intersect, a statement that is not affected by translating them parallel to themselves. So two-body elastic scattering can be nontrivial even in an integrable field theory.
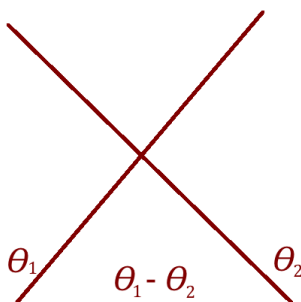
Figure 5: Elastic scattering depends only on the rapidity difference.

How do we characterize a particle? A particle has a velocity, or better, in relativistic terms, a "rapidity" θ. The energy E and momentum p are expressed in terms of θ and the particle mass m by

$$(1.1) \qquad \begin{pmatrix} E \\ p \end{pmatrix} = m \begin{pmatrix} \cosh\theta \\ \sinh\theta \end{pmatrix}$$

A Lorentz boost adds a constant to the rapidity, so scattering of two particles with rapidities θ_1 and θ_2 depends only on the rapidity difference $\theta = \theta_1 - \theta_2$ (Fig. 5).

But the amplitude for scattering of two particles of rapidities θ_1 and θ_2 is in general not only a function of the rapidity difference $\theta = \theta_1 - \theta_2$ because there may be several different "types" of particles of the same mass. An obvious reason for this, although not the most general possibility, is that the theory might have a symmetry group G and the particles may be in an irreducible representation ρ of G.

The picture is then more like Fig. 6. Here i, j, k, l can be understood to represent basis vectors in the representation ρ. We write $R_{ij,kl}(\theta)$ for the quantum mechanical "amplitude" that describes this process. It is usually called the R-matrix.

The real fun comes when we consider three particles in the initial and final state. Since we can move them relative to each other, leaving their slopes (or rapidities) fixed, we can assume that there are only pairwise collisions. But there are two ways to do this, as in Fig. 7, and they must give equivalent results. Concretely equivalence of these pictures leads to the celebrated
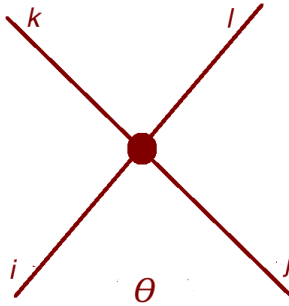
Figure 6: Elastic scattering of particles that carry labels i, j, k, l, with rapidity difference θ.
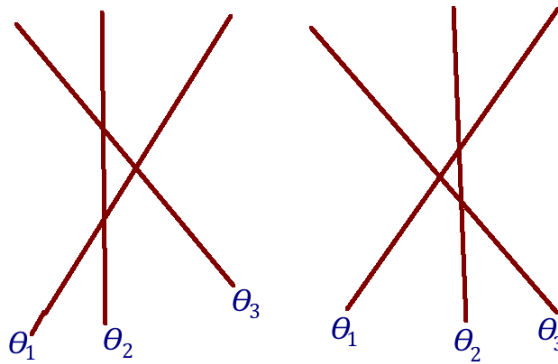
Figure 7: There are two ways to express a $3 \to 3$ scattering process as a succession of three $2 \to 2$ scattering processes.

"Yang-Baxter equation," which schematically reads

(1.2) $$R_{23}R_{13}R_{12} = R_{12}R_{13}R_{23}.$$

The Yang-Baxter equation is a good example of a relationship that is much more transparent in terms of a picture (Fig. 8) than by writing out an algebraic formula in detail.
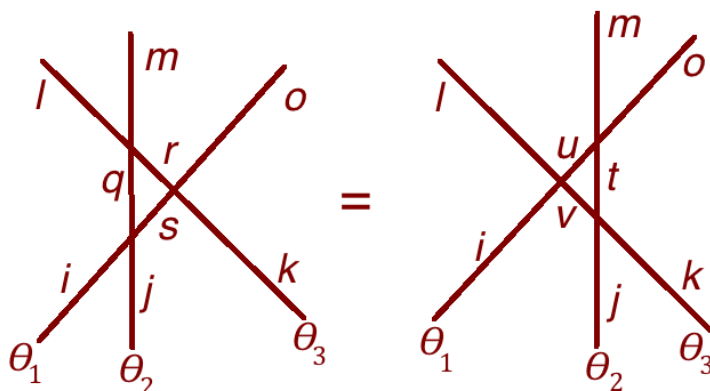
Figure 8: In more detail, the equivalence of the two pictures in Fig. 7 leads to this relation, called the Yang-Baxter equation. A sum over internal particle types (q, r, and s on the left and u, t, and v on the right) is understood.

Actually, there is a subtle but important difference between the R-matrix that solves the Yang-Baxter equation and the S-matrix that describes particle scattering in an integrable relativistic field theory. The reason for this is that the Yang-Baxter equation is not sensitive to an overall c-number factor, that is it is invariant under the transformation $R_{ijkl}(\theta) \to F(\theta)R_{ijkl}(\theta)$, for any scalar function $F(\theta)$. (This invariance is manifest in Fig. 8; on both the left and the right, one has a succession of three scattering events with rapidity differences $\theta_i - \theta_j$, $1 \le i < j \le 3$. So overall factors $F(\theta_i - \theta_j)$ will cancel out.) In applications of the Yang-Baxter equation to classical statistical mechanics, such an overall factor is of little importance and one usually simply picks a convenient normalization of the solution of the Yang-Baxter equation. But in S-matrix theory, the prefactor $F(\theta)$ is very important. In simple examples, as pioneered in [11], the prefactor is determined by conditions of unitarity and crossing symmetry and some knowledge of whether the S-matrix should have bound state poles. It is possible to have two different models (for example the nonlinear σ-model with target a sphere S^{N-1} and corresponding $O(N)$ symmetry and an N-component Gross-Neveu model with the same symmetry) that are governed by the same R-matrix but with

different choices of the prefactor. In fact, recently it has been argued [12] that integrable relativistic field theories in $1 + 1$ dimensions have continuous deformations whose S-matrices differ only by the choice of $F(\theta)$. These deformations depend on infinitely many parameters that are irrelevant in the renormalization group sense.

The traditional solutions of the Yang-Baxter equation – as discovered by Bethe, Lieb, Yang, Baxter, Fadde'ev, Belavin, Drinfeld and others – are classified by the choice of a simple Lie group G and a representation ρ, subject to (1) some restrictions on ρ, and (2) the curious fact that in many important cases (like the 6-vertex model of Lieb and the 8-vertex model of Baxter) a solution of the Yang-Baxter equation associated to a given group G does not actually have G symmetry. In fact, there are three broad classes of solutions of Yang-Baxter, which are called rational, trigonometric, and elliptic depending on whether the R-matrix is a rational, trigonometric, or elliptic function of θ. Only the rational solutions of Yang-Baxter have G symmetry. Prior to the work of Costello, it was, at least to me, a longstanding puzzle to understand better why there are solutions of the Yang-Baxter equation that are in some sense associated to the Lie group G but do not have G symmetry. (The question can be restated in terms of quantum groups, but to me this merely raises the equivalent question of "why" these exist appropriate quantum group deformations.)

Now finally we come to the lattice models. I have motivated the Yang-Baxter equation by talking about relativistic scattering, but solutions of Yang-Baxter can be used to construct several different kinds of integrable model. Today we will focus on the integrable lattice systems of classical statistical mechanics. They are constructed directly from a solution of the Yang-Baxter equation.[3] To understand how this is done, consider Fig. 9. In this this rather busy picture, the vertical and horizontal lines are labeled by rapidities. I have labeled the vertical lines by the same rapidity θ (though this restriction is not necessary) and the horizontal lines by different rapidities θ_i. In addition, each line segment in the figure is labeled by a basis vector $i_s, j_s, k_s, \ldots$ in some matrix-valued solution of the Yang-Baxter equation. These labels will be the "spins" of our lattice model. For the Boltzmann weights, we include for each crossing in the figure a factor of the appropriate R-matrix element. Thus this is a model in which the spins

---

[3]For example, see [9]. The solution of Yang-Baxter used in this construction need not necessarily be rational, so for this application one considers more general solutions of Yang-Baxter than the ones that arise in the integrable relativistic field theories.
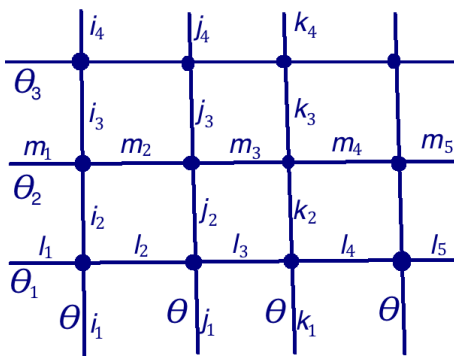
Figure 9: An integrable lattice model associated to a solution of the Yang-Baxter equation. Horizontal and vertical lines are labeled by rapidities. (In constructing the standard integrable models and proving their integrability, one often labels vertical lines by a common rapidity θ and horizontal lines by varying rapidities θ_i, so this case is sketched here.) Line segments between two vertices are labeled by basis vectors i_s, j_s, k_s, etc., in some solution of the Yang-Baxter equation; these are the "spins" of the integrable lattice model. At each vertex, there is a four-spin interaction given by the relevant matrix element of the R-matrix. In some special cases, the four-spin interactions collapse to two-spin interactions, leading to more simple-looking integrable models such as the Ising model.

live on the edges of a square lattice, and there is an interaction among every four spins that live on edges that meet at a common vertex.[4] It turns out that models constructed in this way are integrable, because "the transfer matrices commute," which means that (using the Yang-Baxter equation) the horizontal lines can be moved up and down past each other. Conversely, familiar integrable lattice models can be put in this form.

2. The Yang-Baxter equation and gauge theory

Perhaps the most obvious question about the Yang-Baxter equation is "why" solutions of this highly overdetermined equation exist. There is another area in which one finds something a lot like the Yang-Baxter equations. This is

---

[4]To make a physically sensible model of classical statistical mechanics, one chooses the rapidities so that the relevant matrix elements of the R-matrix are all real and positive.
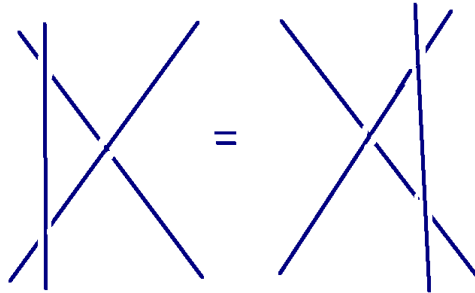
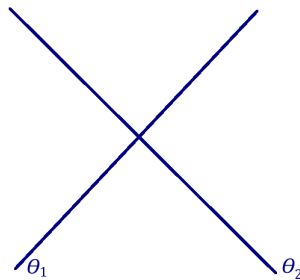Figure 10: A Reidemeister move of knot theory, with an obvious resemblance to the Yang-Baxter equation.

Figure 11: In knot theory, one knot passes "over" or "under" the other, as in the last figure or the next one. But in Yang-Baxter theory, two lines simply cross in the plane, as sketched here.

the theory of knots in three dimensions. Knots are often described in terms of projections to the plane; two knot projections describe the same knot if they can be related by certain moves that are known as Reidemeister moves. The most important Reidemeister move is shown in Fig. 10. The resemblance to the Yang-Baxter equation is obvious, but there is also a conspicuous difference. In knot theory, one strand passes "over" or "under" the other, while Yang-Baxter theory is a purely two-dimensional theory in which lines simply cross, with no "over" or "under," as sketched in Fig. 11. Likewise, rough analogs of the other Reidemeister moves exist in Yang-Baxter theory

(see figs. 4 and 5 of [13]), but they do not really have the same content as in knot theory because of the lack of a distinction between "over" and "under."

In this respect, knot theory has structure that is absent in Yang-Baxter theory. But there is also an important difference in the opposite direction: in Yang-Baxter theory, the spectral parameter is crucial, but it has no analog in knot theory.

Despite these differences, the analogy between the Yang-Baxter equation and the first Reidemeister move of knot theory is rather conspicuous, so let us pursue this a little bit. The usual solutions of Yang-Baxter depend, as I have said, on the choice of a simple Lie group[5] G and a representation ρ. There are knot invariants that depend on the same data. To define them at least formally, let M be a three-manifold, $E \to M$ a G-bundle, and A a connection on G. Then one has the *Chern-Simons function*

$$(2.1) \qquad \mathrm{CS}(A) = \frac{1}{4\pi} \int_M \mathrm{Tr}\left(A\mathrm{d}A + \frac{2}{3}A \wedge A \wedge A\right),$$

which was introduced in quantum field theory in [14, 15]. I have normalized $\mathrm{CS}(A)$ so that, if $G = \mathrm{SU}(N)$ and Tr is the trace in the N-dimenisonal representation, it is gauge-invariant mod $2\pi\mathbb{Z}$. For any G, we define Tr as an invariant quadratic form on the Lie algebra of G normalized so that $\mathrm{CS}(A)$ is gauge-invariant predcisely mod $2\pi\mathbb{Z}$. In quantum mechanics, the "action" I must be well-defined mod $2\pi\mathbb{Z}$, so we can take

$$(2.2) \qquad I = k\mathrm{CS}(A), \quad k \in \mathbb{Z}.$$

A quantum field theory with this action is a "topological quantum field theory" (modulo some subtleties involving a framing anomaly), since there is no metric tensor in sight. Let us just take the three-manifold M to be $\mathbb{R}^3$, and let $K \subset \mathbb{R}^3$ be an embedded knot, such as the trefoil knot of Fig. 12. We pick a representation ρ of G, and let

$$(2.3) \qquad W_\rho(K) = \mathrm{Tr}_\rho P \exp\left(\oint_K A\right)$$

be the Wilson loop operator (the trace of the holonomy) in the representation ρ.

---

[5]In the following gauge theory construction, we take the compact form of G. Also, we slightly simplify the exposition by assuming that G is connected and simply-connected. This implies in particular that a G-bundle $E \to M$, where M is a three-manifold, is in fact trivial. Hence a connection A on E can be understood as a Lie algebra valued 1-form, and $\mathrm{CS}(A)$ can be defined by the naive formula (2.1).
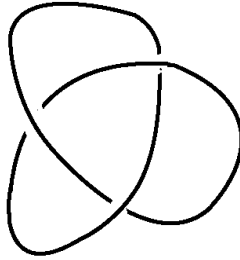
Figure 12: The trefoil knot, projected to the plane.

The usual "quantum knot invariants," of which the prototype is the Jones polynomial of a knot, can be defined via the expectation value of the Wilson operator, $\langle W_\rho(K)\rangle = \langle \mathrm{Tr}_\rho P \exp(\oint_K A)\rangle$. One can make knot invariants this way, but from them one cannot really extract the usual solutions of the Yang-Baxter equation since one is missing the spectral parameter. However (see [9], Eqn. (6.11)), in a sense from these knot invariants one can extract a special case of the trigonometric solutions of the Yang-Baxter equation in which the spectral parameter is taken to $i\infty$.

How can we modify or generalize Chern-Simons gauge theory to include the spectral parameter? A naive idea is to replace the finite-dimensional gauge group G with its loop group $\mathcal{L}G$. We parametrize the loop by an angle θ. The loop group has "evaluation" representations that "live" at a particular value $\theta = \theta_0$ along the loop.[6] We hope that this will be the spectral parameter label θ_0 carried by a particle in the solution of the Yang-Baxter equation.

Taking the gauge group to be a loop group means that the gauge field $A = \sum_i A_i(x)\mathrm{d}x^i$ now depends also on θ and so is $A = \sum_i A_i(x,\theta)\mathrm{d}x^i$. Note that there is no $\mathrm{d}\theta$ term so this is not a full four-dimensional gauge field.

---

[6] It is important here that by $\mathcal{L}G$ we mean the loop group, not its central extension that is encountered in conformal field theory. $\mathcal{L}G$, unlike its central extension, has homomorphisms to G that map a loop $g(\theta)$ to its value $g(\theta_0)$ at some given θ_0. By composing this with an ordinary representation of G, we get a representation of $\mathcal{L}G$ that informally "lives" at the point $\theta = \theta_0$. These representations have no close analog for the centrally extended loop group.
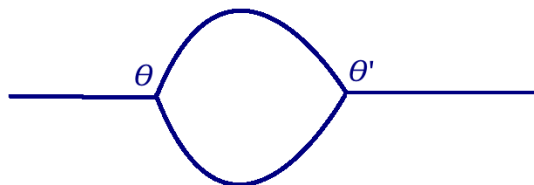
Figure 13: In a naive Chern-Simons theory with gauge group $\mathcal{L}G$, this loop will come with a factor $\delta(\theta - \theta')^2 = \delta(\theta - \theta')\delta(0)$ because of the two propagators connecting the same two vertices.

The Chern-Simons action has a generalization to this situation:

$$(2.4) \qquad I = \frac{k}{4\pi} \int_{M \times S^1} \mathrm{d}\theta \, \mathrm{Tr} \left(A\mathrm{d}A + \frac{2}{3} A \wedge A \wedge A \right).$$

This is perfectly gauge-invariant.

What goes wrong is that because there is no $\partial/\partial\theta$ in the action, the "kinetic energy" of A is not elliptic and the perturbative expansion is not well-behaved. The propagator is

$$(2.5) \qquad \left\langle A_i(\vec{x}, \theta) A_j(\vec{x}', \theta') \right\rangle = \frac{2\pi}{k} \frac{\epsilon_{ijk}(x - x')^k}{|\vec{x} - \vec{x}'|^2} \delta(\theta - \theta')$$

with a delta function because of the missing $\partial/\partial\theta$. Because of the delta function, every loop diagram will acquire a factor $\delta(0)$ (Fig. 13).

Costello cured this problem via a very simple deformation. Take our three-manifold to be $\mathbb{R}^3$, and write x, y, t for the three coordinates of $\mathbb{R}^3$, so overall we have x, y, t, and θ. Costello combined t and θ into a complex variable

$$(2.6) \qquad z = \varepsilon t + \mathrm{i}\theta.$$

Here ε is a real parameter. The theory will reduce to the bad case that I just described if $\varepsilon = 0$. As soon as $\varepsilon \neq 0$, its value does not matter, since it can be eliminated by rescaling t or θ, and one can set $\varepsilon = 1$. The purpose of including ε in Eqn. (2.6) was simply to explain in what sense we will be making an infinitesimal deformation away from the ill-defined Chern-Simons theory of the loop group.

One replaces $\mathrm{d}\theta$ (or $(\mathrm{i}k/4\pi)\mathrm{d}\theta$) in the naive theory (2.4) with $\mathrm{d}z$ (or $\mathrm{d}z/\hbar$, where $\hbar$ is a parameter[7] that plays the role of $1/k$ in the usual Chern-Simons theory) and one now regards A as a partial connection on $\mathbb{R}^3 \times S^1$ that is missing a $\mathrm{d}z$ term (rather than missing $\mathrm{d}\theta$, as before). The action is now

$$(2.7) \qquad I = \frac{1}{\hbar} \int_{\mathbb{R}^3 \times S^1} \mathrm{d}z \, \mathrm{Tr}\left(A\mathrm{d}A + \frac{2}{3}A \wedge A \wedge A \right).$$

We have lost the three-dimensional symmetry of standard Chern-Simons theory, because of splitting away one of the three coordinates of $\mathbb{R}^3$ and combining it with θ. We still have two-dimensional diffeomorphism symmetry. However, as we discussed in comparing Yang-Baxter theory to knot theory, Yang-Baxter theory does not have three-dimensional symmetry, but only two-dimensional symmetry. Modifying standard Chern-Simons theory in this fashion turns out to be just right to give Yang-Baxter theory rather than knot theory: the three-dimensional diffeomorphism invariance is reduced to two-dimensional diffeomorphism invariance, but on the other hand, now there is a complex variable z that will turn out to be the (complexified) spectral parameter.

I have described the action so far on $\mathbb{R}^2 \times \mathbb{C}^*$ where $\mathbb{C}^* = \mathbb{R} \times S^1$ (parametrized by $z = t + \mathrm{i}\theta$) is endowed with the complex 1-form $\mathrm{d}z$. The classical action makes sense more generally on[8] $\Sigma \times C$, where Σ is any smooth (oriented) two-manifold and C is a complex Riemann surface endowed with a holomorphic 1-form ω:

$$(2.8) \qquad I = \frac{1}{\hbar} \int_{\Sigma \times C} \omega \wedge \mathrm{Tr}\left(A \wedge \mathrm{d}A + \frac{2}{3}A \wedge A \wedge A \right).$$

---

[7]One should think of $\hbar$ as a formal complex parameter, since there is in general no natural reality condition that it might obey. We return to this point at the end of these notes.

[8] Quantum mechanically, one runs into an analog of the framing anomaly of Chern-Simons theory, such that Σ has to be framed (its tangent bundle has to be trivialized) in order to define the quantum theory. This is very restrictive, since a compact two-manifold that can be framed is topologically a two-torus T^2. The example $\Sigma = T^2$ with a periodic array of Wilson lines is used in [1] in constructing integrable lattice models with periodic boundary conditions. There is also an analog in this theory of the usual framing anomaly for knots in Chern-Simons theory; this anomaly markedly restricts the class of Wilson operators that can be considered. For example, a Wilson operator supported on a simple closed loop in $\Sigma = \mathbb{R}^2$, at a point in C, is anomalous. There is no problem for the "straight" Wilson lines that are used in constructing integrable lattice models.

It turns out, however, that to get a quantum theory, one wants ω to have no zeroes. Intuitively this is because a zero of ω is equivalent to a point at which $\hbar \to \infty$. (More technically, if ω has a zero, the kinetic energy of the theory is not elliptic and one runs into a difficulty in quantization somewhat similar to the problem with the naive theory of Eqn. (2.4). One can resolve the problem by modifying the theory near a zero of ω, but then two-dimensional symmetry is lost.) By contrast, there is no problem with poles of ω, where effectively $\hbar \to 0$.

So C has to be a complex Riemann surface that has a differential ω with possible poles, but with no zeroes. The only three options are $\mathbb{C}$, $\mathbb{C}/\mathbb{Z} \cong \mathbb{C}^*$, and $\mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$, which (with τ a complex number of positive imaginary part) is a Riemann surface of genus 1. It turns out that these three cases correspond to the three traditional classes of solutions of the Yang-Baxter equation – rational, trigonometric, and elliptic.

The first point is that this theory has a sensible propagator and a sensible perturbation expansion. The basic reason for a sensible propagator is that on $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R} \times S^1$ parametrized by t and θ, the operator $\partial/\partial t$ that appeared in the naive action (2.4) is not elliptic, but the operator $\partial/\partial\overline{z}$ that appears in the deformed version is elliptic. After a suitable gauge-fixing, the propagator (for the rational model, i.e. on $\mathbb{R}^2 \times C$ with $C = \mathbb{C} \cong \mathbb{R}^2$) is

$$
(2.9) \qquad \langle A_i(x, y, z) A_j(x', y', z') \rangle
$$
$$
= \hbar \varepsilon_{ijkz} g^{kl} \frac{\partial}{\partial x^l} \left(\frac{1}{(x - x')^2 + (z - z')^2 + |z - z'|^2} \right),
$$

where ε_{ijkm} is the four-dimensional antisymmetric tensor (but in the formula we set $m = z$, so that i, j, k take the values $x, y, \overline{z}$), and the metric on $\mathbb{R}^4 = \mathbb{R}^2 \times C$ is $\mathrm{d}x^2 + \mathrm{d}y^2 + |\mathrm{d}z|^2$.

With this propagator, the perturbative expansion is well-defined, as shown in [1]. This is a tricky point. The theory is actually unrenormalizable by power counting, so on that basis, one would not expect a well-behaved quantum theory. However, it has no possible counterterms, because all local gauge-invariant operators vanish by the classical equations of motion. Anyway, using a fairly elaborate algebraic machinery of BV quantization, it is shown in [1] that the theory has a well-defined perturbation expansion.

Now we consider Wilson operators, that is holonomy operators

$$
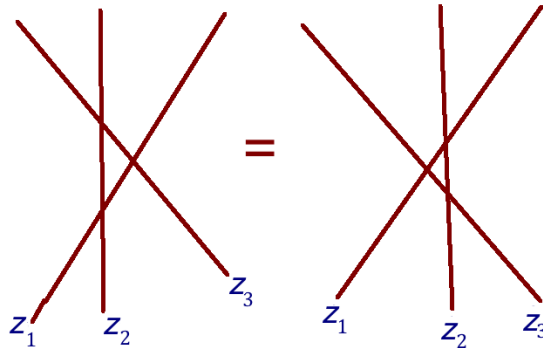(2.10) \qquad \mathrm{Tr}_\rho P \exp \oint_\ell A
$$

Figure 14: Three lines crossing in the plane, now labeled by the value of a complexified spectral parameter z. The equivalence of the two pictures is the Yang-Baxter equation.

where ℓ is a loop in $\Sigma \times C$. As before, Σ is a topological two-manifold, and C is a complex Riemann surface with the differential $\omega = \mathrm{d}z$. But we have only a partial connection

$$(2.11) \qquad\qquad A = A_x \mathrm{d}x + A_y \mathrm{d}y + A_{\bar{z}} d\bar{z}$$

so we would not know how to do any parallel transport in the z direction. (We cannot interpret A as a gauge field with $A_z = 0$ because this condition would not be gauge-invariant, and quantizing the theory requires gauge-invariance. We have to interpret it as a theory with A_z undefined, so we cannot do parallel transport in the z direction.) This means that we must take ℓ to be a loop that lies in Σ, at a particular value of z.

But this is what we wanted to explain Yang-Baxter theory. It means that ℓ is labeled by some constant value $z = z_0$ of the (complexified) spectral parameter z.

Now let us consider some lines that meet in Σ in the familiar configuration associated to the Yang-Baxter equation (Fig. 14). Two-dimensional diffeomorphism invariance means that we are free to move the lines around as long as we do not change the topology of the configuration. But assuming that z_1, z_2, and z_3 are all distinct, it is manifest that there is no discontinuity when we move the middle line from left to right even when we do
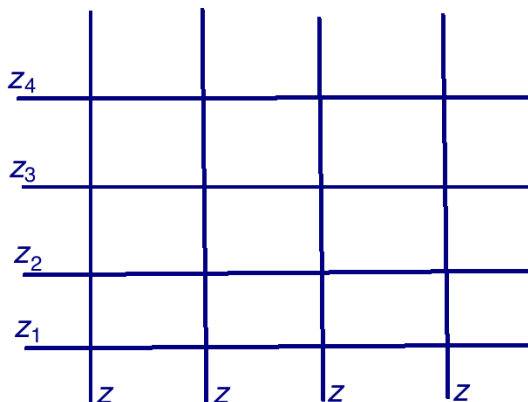
Figure 15: The configuration of Wilson lines associated to the integrable lattice models. The vertical and horizontal lines are labeled by points in C and by representations of G (the representations are not indicated in the figure).

cross between the two pictures, because the Feynman diagrams in four dimensions have no singularity when this occurs. Thus two configurations of Wilson operators that differ by what we might call a Yang-Baxter move are equivalent.

Likewise, in the configuration associated to integrable lattice models (Fig. 15), we can move the horizontal lines up and down at will.

But why is there as elementary a picture as in the usual integrable lattice models, where one can evaluate the path integral by labeling each line by a basis element of the representation ρ and each crossing by a local factor $R_{ij,kl}(z)$? This is a little tricky and depends on picking the right boundary conditions so that the only classical solution of the equations of motion is the trivial solution $A = 0$, and it has no automorphisms (unbroken gauge symmetries). There is a way to achieve this for each of the three choices of C, corresponding to rational, trigonometric, and elliptic solutions of the Yang-Baxter equation.

For simplicity I will here only explain the rational case, in which the Riemann surface is $C = \mathbb{C}$, the complex z-plane, with 1-form $\mathrm{d}z$. We require that the gauge field A on $\Sigma \times \mathbb{C}$ goes to 0 at infinity in the $\mathbb{C}$ direction, and likewise in quantizing we divide only by gauge transformations that approach 1 at infinity along $\mathbb{C}$. With these conditions, it is indeed true that $A = 0$ is the only classical solution, and that it has no unbroken gauge symmetries.

However, G still exists as a group of global symmetries (corresponding to gauge transformations that are constant at infinity along C, rather than equaling 1). That will explain, from this point of view, why the rational solutions of the Yang-Baxter equation, which arise from this construction, are G-invariant.[9]

When we expand around the trivial solution $A = 0$, the absence of deformations or automorphisms of this solution makes the perturbative expansion straightforward. Moreover, perturbation theory gives a simple answer because the theory is infrared-trivial, which is the flip side of the fact that it is unrenormalizable by power-counting. That means that effects at "long distances" in the topological space are negligible.

I put the phrase "long distances" in quotes because two-dimensional diffeomorphism invariance means that there is no natural notion of distance on the topological two-manifold Σ. A metric on $\Sigma \times \mathbb{C}$ entered only when we fixed the gauge to pick a propagator. Recall that we used the metric $\mathrm{d}x^2 + \mathrm{d}y^2 + |\mathrm{d}z|^2$. We could equally well scale up the metric along Σ by any factor and use instead $e^B(\mathrm{d}x^2 + \mathrm{d}y^2) + |\mathrm{d}z|^2$ for very large B.

That means that when one looks at the picture of Fig. 15 associated to the integrable lattice systems, one can consider the vertical lines and likewise the horizontal lines to be very far apart (compared to $z - z_i$ or $z_i - z_j$). In such a situation, in an infrared-free theory, effects that involve a gauge boson exchange between two nonintersecting lines, as in Fig. 16, are negligible.

One should worry about gauge boson exchange from one line to itself, as in Fig. 17, because in this case the distance $|a - b|$ need not be large. Such effects correspond roughly to "mass renormalization" in standard quantum field theory. In the present problem, in the case of a straight Wilson line such as those of Fig. 15, the symmetries do not allow any interesting effect analogous to mass renormalization. (For more general Wilson lines, the "mass renormalization" diagram leads to a framing anomaly that was mentioned in footnote 8.)

What about gauge boson exchange between two Wilson lines with distinct spectral parameters z, z' that do cross in Σ? The lowest order example with exchange of a single gauge boson is depicted in Fig. 18. To evaluate the contribution of this diagram, we have to integrate over the points a, b

---

[9]For the other choices of C, boundary conditions that allow a unique classical solution that is free of unbroken gauge symmetries do exist (for any G if $C = \mathbb{C}^*$ and for $G = PSU(N)$ if C is a curve of genus 1), but lack the full global G symmetry. This explains from the present point of view why the trigonometric and elliptic solutions of Yang-Baxter are not G-invariant, and why elliptic solutions only exist for the special linear group.
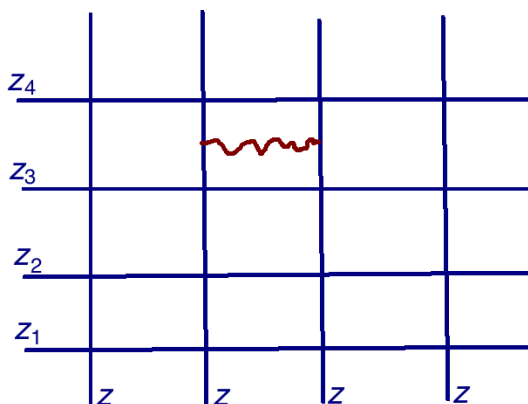
Figure 16: Because the deformed Chern-Simons theory is defined with no choice of metric on Σ, we can always use a gauge such that nonintersecting Wilson lines are very far apart. Gauge boson exchanges between noninter-secting Wilson lines, as depicted here, vanish in this limit.
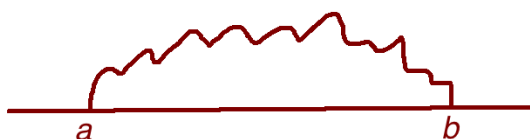

Figure 17: Exchange of a gluon from one Wilson line to itself would corre-spond to mass renormalization (of an external probe charge) in conventional quantum field theory. In the framework that leads to the integrable lattice models, the symmetries do not allow any nontrivial effect analogous to mass renormalization.
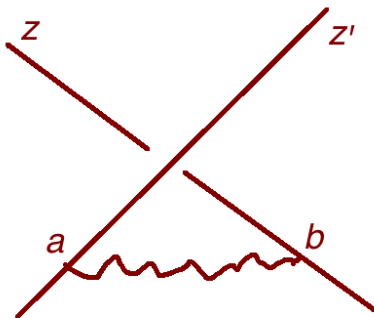
Figure 18: The lowest order diagram with gauge boson exchange between two lines that cross in Σ. The two lines have spectral parameters z, z'. To evaluate the contribution of this diagram, we have to integrate over a and b. The integral converges and is dominated by the region $|a|, |b| \lesssim |z - z'|$.

at which the gauge boson is attached to the two Wilson lines. The resulting integral converges and is dominated by the region $|a|, |b| \lesssim |z - z'|$. What it converges to will be discussed shortly.

Now when we study a general configuration such as the one related to the integrable lattice models we can draw very complicated diagrams, as in Fig. 19, but the complications are all localized near one crossing point or another. The diagrams localized near one crossing point simply build up an R-matrix associated to that crossing, and the discussion makes it obvious that the R-matrix obeys the Yang-Baxter relation of Fig. 14.

Moreover, this makes it clear that the path integral in the presence of the configuration of Wilson operators (Fig. 15) associated to the integrable lattice models can be evaluated by the standard rules: label each vertical or horizontal line segment by a basis vector of the representation ρ and include the appropriate R-matrix element at each crossing; then sum over all such labelings.

But why is the R-matrix obtained this way the standard rational solution of the Yang-Baxter equation? (And similarly, with a different choice of C, why would we get the standard trigonometric and elliptic solutions of Yang-Baxter?) In his paper [1], Costello explicitly evaluates the lowest order correction in

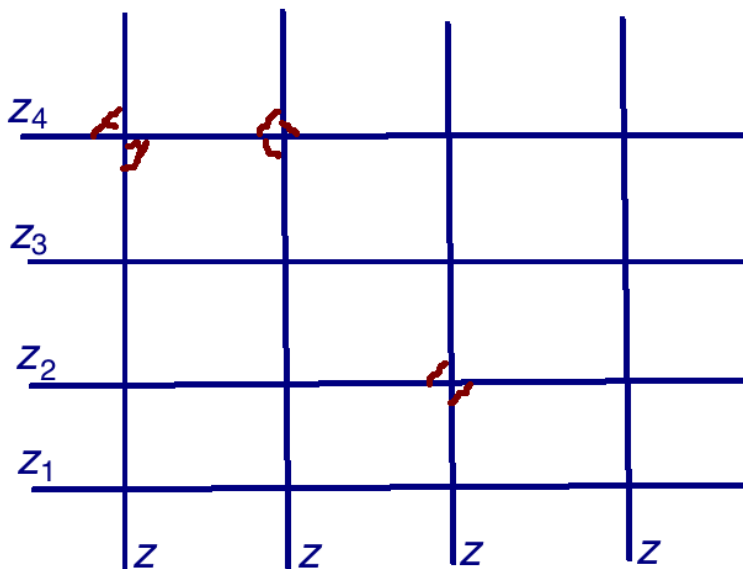$$(2.12) \qquad\qquad R = 1 + \hbar r + \mathcal{O}(\hbar^2)$$

Figure 19: Evaluating in perturbation theory the configuration of Wilson operators associated to the integrable lattice models, we run into very complicated Feynman diagrams. But after scaling up the metric of Σ, we can assume that these diagrams are built from decoupled diagrams each of which is associated to just one crossing point.

from the lowest order non-trivial diagram of Fig. 18, and and gets the standard answer

$$(2.13) \qquad\qquad r = \frac{\sum_a t_a t_a'}{z - z'}$$

(where t_a, t_a', $a = 1, \ldots, \dim G$ are the generators of the Lie algebra of G acting in the two representations). Once the first order deformation is known, the full answer for the R-matrix, up to a change of variables, follows from general arguments (see p. 814 in [16] and p. 418 in [17]).

I conclude with a few final comments. Costello's theorem is purely about perturbation theory, but his theorem shows that, in this particular theory (and rather exceptionally), perturbation theory converges. As a physicist, one would want to give an *a priori* "nonperturbative definition" of the theory, which would have the claimed perturbative expansion. At first sight,

there are some difficulties in doing this. The action of conventional Chern-Simons theory is gauge-invariant mod $2\pi\mathbb{Z}$ if the coupling parameter k is an integer, as one wishes for a quantum system that will be well-defined beyond perturbation theory. But there is no choice of the analogous parameter $\hbar$ in Eqn. (2.7) or (2.8) that makes the action gauge-invariant mod $2\pi\mathbb{Z}$. Also, the path integral of conventional Chern-Simons gauge theory is an oscillatory path integral, like that of any unitary quantum system, because the action is real.[10] The real and imaginary parts of the generalized action in Eqn. (2.7) or (2.8) are unbounded above and below, so naively the path integral is exponentially divergent, no matter what we assume for $\hbar$.

Because of such considerations, to get a convergent, nonperturbative path integral in this theory, one will have to consider an analytically continued version of the theory, in a sense that for conventional Chern-Simons theory was described in [18]. This analytic continuation is achieved by complexifying all variables and then constructing a different integration cycle on which the path integral converges. For ordinary Chern-Simons theory, such analytic continuation is optional; the theory at integer k is perfectly well-defined without any analytic continuation. But in the four-dimensional theory that is under discussion here, only an analytically-continued version of the path integral will make sense beyond perturbation theory. To construct this analytic continuation, one can consider the D4-NS5 system of string theory, and proceed via the same arguments that were used in [18] to relate ordinary Chern-Simons theory and the Jones polynomial to the D3-NS5 system.

An interesting goal for the future is to somehow link the story that has been reviewed here to the work of Nekrasov and Shatashvili, who developed a seemingly quite different relationship between integrable quantum spin systems and supersymmetric gauge theory [2, 3].

Another natural question is whether models such as the chiral Potts model (see for example [19–22]) can be placed in the gauge theory framework. The chiral Potts model is an integrable lattice model in which the spectral parameter takes values in a curve of genus greater than 1. Perhaps surprisingly (given the genus of its spectral curve), the chiral Potts model can be related to a trigonometric solution of Yang-Baxter for the group $SU(2)$ (or SL_2), and this viewpoint has also been used to develop an analog for $SU(N)$ [23, 24]. But such models are not yet understood in the framework described in the present lecture.

---

[10]With $\mathrm{CS}(A)$ defined as in Eqn. (2.1), the argument of the path integral is $\exp(ik\mathrm{CS}(A))$.

References

[1] K. Costello, *Supersymmetric gauge theory and the yangian*, arXiv:1303.2632.

[2] N. A. Nekrasov and S. L. Shatashvili, *Supersymmetric vacua and Bethe ansatz*, Nucl. Phys. Proc. Suppl. **192-3** (2009), 91–112. arXiv:0901.4744.

[3] N. A. Nekrasov and S. L. Shatashvili, *Quantization of integrable systems and four dimensional gauge theories*, in: Proceedings, 16th International Congress on Mathematical Physics (P. Exner, ed.), arXiv:0901.4748.

[4] M. F. Atiyah, *New invariants of* 3- *and* 4-*dimensional manifolds*, in: The Mathematical Heritage of Hermann Weyl (R. O. Wells, ed.), American Mathematical Society, 1988.

[5] E. Witten, *Quantum field theory and the Jones polynomial*, Commun. Math. Phys. **121** (1989), 351–399.

[6] K. Costello, *Integrable lattice models from four-dimensional field theory*, Proc. Symp. Pure Math. **88** (2014), 3–24. arXiv:1308.0370.

[7] L. D. Faddeev and L. A. Takhtajan, *Hamiltonian methods in the theory of solitons*, Springer-Verlag, 1980.

[8] M. Jimbo, ed., *Yang-Baxter equation in integrable systems*, World-Scientific, 1989.

[9] M. Wadati and Y. Akutsu, *From solitons to knots and links*, Prog. Theor. Phys. Suppl. **94** (1988), 1–41.

[10] J. H. H. Perk and H. Au-Yang, *Yang-Baxter Equation*, in: Encyclopedia of Mathematical Physics (J.-P. Françoise, G. L. Naber, and S. T. Tsou, eds.), Vol. 5, pp. 465–73, Oxford: Elsevier, 2006. Expanded version of arXiv:math-ph/0606053.

[11] A. B. Zamolodchikov and Al. B. Zamolodchikov, *Factorized S-matrices in two-dimensions as the exact solutions of certain relativistic field theories*, Annals Phys. **120** (1979), 253–291.

[12] F. A. Smirnov and A. B. Zamolodchikov, *On space of integrable quantum field theories*, arXiv:1608.05499.

[13] J. Perk and F. Y. Wu, *Graphical approach to the nonintersecting string model: Star-triangle equation, inversion relation, and exact solution*, Physica **138A** (1986), 100–124.

[14] A. S. Schwarz, *The partition function of degenerate quadratic functional and Ray-Singer invariants*, Lett. Math. Phys. **2** (1978), 247–252.

[15] S. Deser, R. Jackiw, and S. Templeton, *Topologically massive gauge theories*, Annals Pys. **140** (1982), 372–411.

[16] V. G. Drinfeld, *Quantum groups*, in: Proceedings of the International Congress of Mathematicians, Vols. 1, 2, American Mathematical Society, 1987.

[17] V. Chari and A. Pressley, *A guide to quantum groups*, Cambridge University Press, 1995.

[18] E. Witten, *Fivebranes and knots*, Quantum Topol. **3** (2012), 1–137. `arXiv:1101.3216`.

[19] H. Au-Yang, B. M. McCoy, J. H. H. Perk, S. Tang, and M.-L. Yan, *Commuting transfer matrices in the chiral Potts models: solutions of star triangle equations with genus > 1*, Phys. Lett. **A123** (1987), 219–223.

[20] R. J. Baxter, J. H. H. Perk, and H. Au-Yang, *New solutions of the star triangle relations for the chiral Potts model*, Phys. Lett. **A128** (1988), 138–142.

[21] V. V. Bazhanov and Yu. G. Stroganov, *Chiral Potts model as a descendant of the six vertex model*, J. Stat. Phys. **59** (1990), 799–817.

[22] R. Baxter, V. V. Bazhanov, and J. H. H. Perk, *Functional relations for transfer matrices of the chiral Potts model*, Int. J. Mod. Phys. **B4** (1990), 803–870.

[23] E. Date, M. Jimbo, K. Miki, and T. Miwa, *Generalized chiral Potts models and minimal cyclic representations of $U_q(\widehat{\mathfrak{gl}}(n, \mathrm{C}))$*, Commun. Math. Phys. **137** (1991), 133–147.

[24] V. V. Bazhanov, R. M. Kashaev, V. V. Mangazeev, and Yu. G. Stroganov, *$(Z_N \times)^{n-1}$ generalization of the chiral Potts model*, Commun. Math. Phys. **138** (1991), 393–408.

School of Natural Sciences, Institute for Advanced Study
Einstein Drive, Princeton, NJ 08540 USA
E-mail address: witten@ias.edu