Advances in Mathematics **168**, 133–192 (2002)
doi:10.1006/aima.2001.2045

Trace Functionals on Noncommutative Deformations of Moduli Spaces of Flat Connections

Philippe Roche[1]

Laboratoire de Physique Mathématique et Théorique, Université Montpellier 2, 34000 Montpellier, France
E-mail: roche@lpm.univ-montp2.fr

and

András Szenes[2]

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Pavel Etingof

Received May 9, 2001; accepted August 16, 2001

We describe an efficient construction of a canonical noncommutative deformation of the algebraic functions on the moduli spaces of flat connections on a Riemann surface. The resulting algebra is a variant of the quantum moduli algebra introduced by Alekseev, Grosse, and Schomerus and Buffenoir and Roche. We construct a natural trace functional on this algebra and show that it is related to the canonical trace in the formal index theory of Fedosov and Nest and Tsygan via Verlinde's formula.

Key Words: formal deformations; quantum groups; moduli spaces.

Contents.

[1] The first author was supported by the CNRS.
[2] The second author was supported by NSF Grant DMS-9870053 and NSA Grant 6800900.

0001-8708/02 $35.00

0. INTRODUCTION

Let G be a compact connected and simply connected Lie group and Σ be a compact topological Riemann surface with a point p marked on it. These data determine a space denoted by $\mathfrak{M}^G = \mathfrak{M}^G[\Sigma_p]$, the moduli space of flat G-connections on the punctured Riemann surface $\Sigma \setminus \{p\}$. This space decomposes into a union $\mathfrak{M}^G = \bigcup_\sigma \mathfrak{M}^G_\sigma$ of moduli spaces of flat connections with the holonomy around the point p fixed in some conjugacy class σ of G. For generic σ the space $\mathfrak{M}^G_\sigma$ is a smooth real algebraic orbifold (manifold if $G = SU(n)$), which comes equipped with a canonical algebraic symplectic structure.

In this paper, we study noncommutative deformations of the spaces of functions on $\mathfrak{M}^G$ and $\mathfrak{M}^G_\sigma$. We bring together two theories of deformation:

• the algebraic deformation of the moduli spaces initiated and developed in the works of Fock and Rosly [20], Alekseev *et al.* [3], and Buffenoir and Roche [11], and

• the theory of star products or formal deformations of symplectic manifolds, more specifically, the index theorem of Fedosov and Nest and Tsygan (F-NT) [14, 19, 29].

Our first result is a greatly simplified, invariant construction of the *quantum moduli algebra*. We fix a generic conjugacy class σ. The quantum moduli algebra is a canonically defined noncommutative algebra A^q, which is finitely generated over a subring $D(q)$ of rational functions in q. One may recover $F(\mathfrak{M}^G)$, the algebraic functions on $\mathfrak{M}^G$, from A^q by setting $q = 1$. Variants of this algebra were defined in [3] and [11] using generators and relations, and a more geometric construction was given in [9]. One may pass to a canonical quotient A^q_σ of A^q, which serves as a deformation of the algebraic functions on $\mathfrak{M}^G_\sigma$.

Next we would like to interpret the algebra A^q_σ in the framework of formal deformation theory of symplectic manifolds. Substituting $q = e^{2\pi i\hbar}$ one obtains an algebra $A^\hbar_\sigma$ over the formal power series $\mathbb{C}[\![\hbar]\!]$. Under certain assumptions and after appropriate normalizations and completions, this algebra can be seen to be isomorphic to a star product algebra defined on $C^\infty(\mathfrak{M}^G_\sigma)$. Note that we have a natural injective map $\iota\colon A^q_\sigma \to A^\hbar_\sigma$.

One of the most natural objects for which one may search when studying a noncommutative algebra A is a cyclic (or trace) functional: a functional $T\colon A \to K$ with value in a module is *cyclic* if $T(a\cdot b) = T(b\cdot a)$. Let M^{2n} be a symplectic manifold. The index theory of Fedosov and Nest and Tsygan applies to a local noncommutative deformation of $C^\infty(M)$ over the ring $\mathbb{C}[\![\hbar]\!]$. As the algebra $A^\hbar_\sigma$ obtained above is an example of such a deformation (this will be proved elsewhere), one may apply the theory to it. The central object of this formal index theory is a cyclic functional $\mathrm{Tr}_{\mathrm{can}}\colon A^\hbar_\sigma \to \hbar^{-n}\mathbb{C}[\![\hbar]\!]$, called the *canonical trace*. The elements of $A^\hbar_\sigma$ are interpreted as formal elliptic operators in this theory, and the canonical trace plays the role of the index functional.

The main focus of the present work is shown in the following commuting diagram:

$$\begin{tikzcd} A^q \arrow[r, "\iota"] \arrow[d, "\mathrm{Tr}^q"'] & A^\hbar_\sigma \arrow[d, "\mathrm{Tr}_{\mathrm{can}}"] \\ K \arrow[r, "\mathrm{ev}"'] & \frac{1}{\hbar^n}\mathbb{C}[\![\hbar]\!] \end{tikzcd}$$

We are searching for a natural module K defined over $D(q)$, a trace functional Tr^q and a map ev which complete the diagram. In other words, we seek to lift the canonical trace from the infinitesimal world to the global q world.

A natural solution to this lifting problem seems to be setting K to be $D(q)$ and $\mathrm{ev}\colon D(q) \to \mathbb{C}[\![\hbar]\!]$ to be the Laurent expansion at $q = 1$. Surprisingly, cyclic functionals on A^q_σ with these parameters do not exist. The natural trace Tr^q on A^q_σ takes values in the space of holomorphic functions on the unit disc $\{|q| < 1\}$. While there is no reasonable functional from this function space to $\mathbb{C}[\![\hbar]\!]$, there is a special class of holomorphic functions: the ones with an *asymptotic expansion* at $q = 1$; the values $\{\mathrm{Tr}^q(a) \mid a \in A^q_\sigma\}$ happen to land here. The map ev then is the asymptotic expansion in powers of $\hbar$.

Our results are subject to various restrictions and assumption which we have omitted in the discussion so far for the sake of clarity. These will be discussed below and in the main body of the paper.

Contents of the paper. In Section 1 we recall the volume formula of Witten and Verlinde's formula. We compute the asymptotics of the q-volume series that we introduce by analogy. In Section 2 we recall the necessary background about the topology of the spaces $\mathfrak{M}^G$ and $\mathfrak{M}^G_\sigma$. We give a short introduction to the theory of formal deformations in Section 3 and explain how a variation of the theory can be applied to the case of algebraic manifolds. We start the study of the algebraic functions on the moduli spaces in Section 4 by giving a construction of the Poisson structure and the Poisson trace using a graphical technique. This technique is mostly based on the ideas and constructions of [4, 20, 30], although it contains some novel elements. We quantize this construction in Section 5 by introducing a modification of the invariants of Reshetikhin and Turaev [31]. The algebra we obtain is analogous to those in [3, 11], but our construction is transparent, geometric, and computationally much more efficient. Our main result is contained in Section 6, where we prove the asymptotic correspondence of the traces to which we alluded above. The proof is complete only in the case of $G = SU(2)$, because a crucial ingredient, an aspect of the dynamical theory of quantum groups is completely understood in this case only. We review our results and formulate the main conjecture which served as a motivation for this article in Section 7. In order to avoid crowding the main body of the paper, and in a hope to make this work accessible for the reader unfamiliar with the theory of quantum groups, we included an introduction to a relevant part of the subject in Appendix A. This section also contains a few explicit calculations that have not appeared elsewhere.

The goal of our work is carrying out the program outlined in Section 3.2 for the algebras we construct in Section 5. This involves completing the analysis of Section 6 for the higher rank groups and proving the statements listed in Theorem 4 in full generality. These problems will be the subjects of further study. With these problems out of the way, one could try to approach Conjecture 1 about the characteristic class of A^q_σ.

We should note that the idea of approaching the Verlinde formulas via formal index theory was raised by Nest and Tsygan.

1. THREE SERIES

In this section we introduce three families of series associated to compact Lie groups. They are all related to the topology of the moduli spaces of flat connections on Riemann surfaces. The first two, the Witten series and the Verlinde sums, are well known. We will sketch their geometric significance in the next section. The last one is the main object of our study, and its connection to the moduli spaces will be explained in the subsequent parts of the paper.

Before we proceed, however, we need to fix some notation.

1.1. *Lie Theory, Notation, and Preliminaries*

The notions of Lie theory will play a major role in this article. Here we set the relevant conventions and notation.

- We will assume that G is a compact, simple, connected, and simply connected Lie group with complexified Lie algebra $\mathfrak{g}$. For most of the paper we also assume that G is simply laced.
- Fix a maximal torus $T \subset G$ with Lie algebra $\mathfrak{t}$, whose complexification is denoted by $\mathfrak{h}$. The pairing between $\mathfrak{t}$ and its dual $\mathfrak{t}^*$ will be denoted by $\langle\, , \rangle$.
- Let $\exp\colon \mathfrak{t} \to T$ be the exponential map. Then $\Lambda = \frac{1}{2\pi} \exp^{-1}(e) \subset \mathfrak{t}$ is called the *unit lattice* and Ω, its integral dual in $\mathfrak{t}^*$, the *weight lattice*. The set of roots will be denoted by $\Delta \subset \Omega$ and the coroot corresponding to a root $\alpha \in \Delta$ by $\check{\alpha}$. For a weight λ we write e_λ for the corresponding character of T, and denote $\tilde{x} = \exp(x)$.
- The Lie algebra $\mathfrak{g}$ has a symmetric bilinear form, which is invariant under the adjoint action of G. Such a form is unique up to multiplication by a constant and induces an inner product on $\mathfrak{g}^*$. The normalization of this inner product is usually fixed in such a way that the long roots (and coroots) have square length 2. The form thus normalized is called the *basic* inner product and will be denoted by $(.,.)$. It induces a linear isomorphism between $\mathfrak{t}$ and $\mathfrak{t}^*$; the element corresponding to x under this correspondence is denoted by x^*.
- The Weyl group W_G acts on T and $\mathfrak{t}$ effectively. Assume that a Weyl chamber in $\mathfrak{t}$ has been chosen. This induces a split of the roots into positive and negative: $\Delta = \Delta^+ \cup \Delta^-$ and the choice of the dominant weights Ω^+. As usual, $\rho = \frac{1}{2} \sum_{\alpha \in \Delta^+} \alpha$, and θ_G is the highest root of G. The dual Coxeter number h_G is the integer defined by $h_G = (\theta_G, \rho) + 1$.
- Denote the set of conjugacy classes of G by $\mathrm{Conj}(G)$; we will write $\sigma(t)$ for the conjugacy class of an element $t \in T$. Let T_{reg} be the set of regular elements in T and $\mathrm{Conj}_{\mathrm{reg}}(G)$ be the set of regular conjugacy classes. Denote by $\mathfrak{a}$ the open simplex in the chosen chamber of $\mathfrak{t}$, which has a vertex at 0 and maps in a one-to-one manner to a chamber of regular elements of T via the exponential mapping. We have

$$\mathrm{Conj}(G) \cong T/W_G \qquad \text{and} \qquad \mathrm{Conj}_{\mathrm{reg}}(G) \cong \mathfrak{a}.$$

Denote by $\mathfrak{a}^*$ the simplex in the dominant chamber corresponding to $\mathfrak{a}$ under the identification * defined above. Then

$$\mathfrak{a}^* = \{\gamma \in \mathfrak{t}^* \,|\, (\gamma, \theta_G) < 1, (\gamma, \alpha) > 0, \alpha \in \Delta^+\}, \tag{1.1}$$

where θ_G is the highest root. The quantity (λ, θ_G) is called the *height* of λ. Thus $\mathfrak{a}^*$ is the set of dominant weights of height less than 1.

- For a dominant integral weight λ denote the irreducible representation with highest weight λ by V_λ and its character by χ_λ. The representation ring $R(G)$ has an integral basis $\mathrm{Irrep}(G) = \{\chi_\lambda\}_{\lambda \in \Omega^+}$.
- Define a partial ordering on the weights by setting $\lambda \geqslant \mu$ if their difference is a sum of positive roots; i.e., $\lambda - \mu \in \mathbb{Z}^{\geqslant 0}\Delta^+$. This partial order may be extended to $\mathfrak{h}^*$, the set of all complex weights.

Other conventions:

- Given a nondegenerate pairing $Q: V \otimes W \to \mathbb{C}$ denote by δ_Q the *diagonal element* $\sum v^i \otimes w_i \in V \otimes W$, where $\{v^i, w_j\}$ are a pair of dual bases of V and W correspondingly; i.e., $Q(v^i, w_j) = \delta^i_j$. In particular, $\delta(V) \in V^* \otimes V$ is the diagonal element with respect to the canonical pairing between V^* and V.
- Underlining a symbol will mean multiplication by $2\pi i$.

For most of this section we assume that G is simply laced. The formulas for the non-simply-laced cases require minor modifications of the ones given.

1.2. *The Rational Case*

For a positive integer g, consider the Λ-periodic distribution on $\mathfrak{t}$ given by the series

$$\tilde{E}^G_g(x) = \sum_{\lambda \in \Omega^+} \frac{\chi_\lambda(\tilde{x})}{(\dim V_\lambda)^{2g-1}}. \tag{1.2}$$

Clearly, $\tilde{E}^G_g(x)$ depends on $\sigma(\tilde{x})$ only. Up to a normalization, to be discussed below, this is the volume series of Witten [43] (cf. (2.17)).

Recall the Weyl character and Weyl dimension formulas

$$\chi_\lambda = \frac{1}{\delta} \sum_{w \in W_G} \mathrm{sign}(w)\, e_{w(\lambda+\rho)} \qquad \text{and} \qquad \dim V_\lambda = \prod_{\alpha \in \Delta^+} \frac{(\lambda+\rho, \alpha)}{(\rho, \alpha)}, \tag{1.3}$$

where $\mathrm{sign}: W_G \to \pm 1$ is the antisymmetric character of W_G and $\delta = e_\rho \prod_{\alpha \in \Delta^+} (1 - e_{-\alpha})$ is the fundamental antisymmetric character.

Multiplying $\tilde{E}^G_g(x)$ by $\delta(\tilde{x})$ and an appropriate x-independent constant one obtains the expression

$$E^G_g(x) = c(g, G)\, \delta(\tilde{x})\, \tilde{E}^G_g(x) = (-1)^{(g-1)|\Delta^+|} |Z_G|^g B^G_{2g-1}(x), \tag{1.4}$$

where

$$B_m^G(x) = \sum_{\lambda \in \Omega_{\text{reg}}} \frac{e_\lambda(\tilde{x})}{\prod_{\alpha \in \Delta^+} (\underline{\lambda}, \alpha)^m}. \tag{1.5}$$

Here Ω_{reg} is the set of regular, not necessarily dominant weights, Z_G is the center of G, and underlining means multiplication by $2\pi i$. The constant $c(g, G)$ is defined by (1.4). The function (1.5) is a multidimensional Fourier series, a higher dimensional analog of the classical Bernoulli polynomials. Such series were studied in a more general context in [33] where, in particular, it was proved that $B_m^G(x)$ is a piecewise polynomial function with rational coefficients. When $G = SU(2)$, up to a normalization factor, one recovers the classical Bernoulli polynomials

$$B_m(x) = \sum_{n \neq 0} \frac{e^{x\underline{n}}}{\underline{n}^m}. \tag{1.6}$$

Here and in the rest of this paper, we will omit the subscript or superscript G when $G = SU(2)$ if this causes no confusion. To compute the coefficients of these polynomials one can apply the following simple lemma [33]:

LEMMA 1. *Let f be a rational function of degree $\leqslant -2$ on $\mathbb{C}$ and let* Pole(f) *be the set of its poles. Then for each x, $0 \leqslant x < 1$,*

$$\sum_{n \in \mathbb{Z},\, \underline{n} \notin \mathrm{Pole}(f)} e^{\underline{n}x} f(\underline{n}) = \sum_{p \in \mathrm{Pole}(f)} \operatorname*{Res}_{u=p} \frac{e^{xu}\, du}{1 - e^u} f(u). \tag{1.7}$$

In the case $f(u) = u^{-m}$, $g > 1$, one recovers the well-known formula

$$B_m(x) = \operatorname*{Res}_{u=0} \frac{1}{u^m} \frac{e^{\{x\}\, u}\, du}{(1 - e^u)},$$

where $\{\ \}$ stands for the fractional part of a real number.

Remark 1. The lemma can be extended to the case $\deg(f) > -2$, by assuming $0 < x < 1$ (cf. [33]).

1.3. *The Trigonometric Case*

Denote by Ω_Δ the root lattice, which is the integral dual of the center lattice $\exp^{-1}(Z_G)$. For positive integers k, g and a dominant weight $\lambda \in \Omega_\Delta^+$, consider the finite sum

$$V_g^G(\lambda; k) = \frac{1}{|W_G|} \sum_{t \in Z_G[k+h_G] \cap T_{\text{reg}}} \chi_\lambda(t) \left(\frac{|Z_G|\, (k+h_G)^{\text{rank}(G)}}{\prod_{\alpha \in \Delta} (1 - e_\alpha(t))} \right)^{g-1}, \tag{1.8}$$

where $Z_G[k]=\{t\in T \mid t^k\in Z_G\}$, and $h_G=(\theta_G,\rho)+1$. This sum, first written down by Verlinde [39], turns out to be an integer valued function, whose dependence on λ and k is again piecewise (quasi-)polynomial. Note that the denominator of the fraction is the Weyl density function which can be written as $\delta\bar{\delta}=(-1)^{|\Delta+|}\delta^2$. Using this and taking advantage of the Weyl character formula again, we arrive at

$$V_g^G(\lambda;k)=((-1)^{|\Delta^+|}\,|Z_G|\,(k+h_G)^{\operatorname{rank}(G)})^{g-1}\sum_{t\in Z_G[k+h_G]\cap T_{\mathrm{reg}}}\frac{e_{\lambda+\rho}(t)}{\delta(t)^{2g-1}}. \tag{1.9}$$

For $G=SU(2)$ we obtain

$$V_g(l;k)=i(2(k+2))^{g-1}\sum_{j=1}^{2k+3}\frac{e^{\underline{j(l+1)/(2(k+2))}}}{(2\sin(j\pi/(k+2)))^{2g-1}},\qquad j\neq k+2,$$

where l is an even number. Again, we can use residue techniques to evaluate this sum:

Lemma 2. *Let $f(z)$ be a rational function on $\mathbb{C}$ with a set of poles* Pole(f), *and let m be a positive integer such that $f(z)\,dz/(z(1-z^m))$ is regular at 0 and at ∞. Then*

$$\sum_{z^m=1,\,z\notin \mathrm{Pole}(f)} f(z)=\sum_{p\in\mathrm{Pole}(f)}\operatorname*{Res}_{z=p}\frac{m\,dz}{z(1-z^m)}\,f(z).$$

Applying the lemma to our sum with $f(z)=z^{l+1}(z-1/z)^{1-2g}$ and $m=2(k+2)$, we obtain

$$V_g(l,k)=-2(-2(k+2))^g\operatorname*{Res}_{z=1}\frac{z^{l'}\,dz}{(1-z^{2(k+2)})(z-1/z)^{2g-1}\,z}, \tag{1.10}$$

where $l'=l+1 \bmod 2k+4$ and $0\leqslant l'<2k+4$. In principle, we also need the residue at $z=-1$, but by symmetry this is equal to the one at $z=1$ (l' is odd), hence the extra factor of 2.

Finally, we can make the substitution $z=e^{u/(2k+4)}$ to arrive at

$$V_g(l,k)=2(-2(k+2))^{g-1}\,p_{2g-1}(\{(l+1)/2(k+2)\},k+2), \tag{1.11}$$

where

$$p_m(x,k)=\operatorname*{Res}_{u=0}w_m(x,k;u),\qquad w_m(x,k;u)=\frac{e^{xu}\,du}{(1-e^u)(e^{u/2k}-e^{-u/2k})^m}. \tag{1.12}$$

It is easy to see that p_m is a polynomial of both of its arguments.

1.4. *The q-Rational Case*

Now we turn to the main object of our study. Define the q-integers by $[n]_q = (q^n - q^{-n})/(q - q^{-1}) \in \mathbb{Z}[q, q^{-1}]$ for every $n \in \mathbb{Z}^{>0}$. The q-dimension of V_λ is defined in analogy with (1.3) by

$$q\dim V_\lambda = \prod_{\alpha \in \Delta^+} \frac{[(\lambda+\rho, \alpha)]_q}{[(\rho, \alpha)]_q} \in \mathbb{Z}[q, q^{-1}].$$

By replacing the classical dimension by the q-dimension, we can write down the q-version of the $\widetilde{E}_g^G$ series:

$$\widetilde{T}_g^G(x; q) = \sum_{\lambda \in \Omega^+} \frac{\chi_\lambda(\tilde{x})}{(q\dim V_\lambda)^{2g-1}}.$$

The function $q\dim V_\lambda$, as a function of λ, has the same symmetry properties with respect to the Weyl group as the usual dimension, so we can use the same trick as above to arrive at the analog of the E_g^G series: $T_g^G(x; q) = c(g, G, q)\, \delta(\tilde{x})\, \widetilde{T}_g^G(x; q)$,

$$T_g^G(x; q) = (-1)^{(g-1)\,|\Delta^+|}\, |Z_G|^g \sum_{\lambda \in \Omega_{\mathrm{reg}}} \frac{e_\lambda(\tilde{x})}{(\prod_{\alpha \in \Delta^+} q^{(\alpha, \lambda)} - q^{-(\alpha, \lambda)})^{2g-1}}.$$

Note that while $c(g, G, q)$ is an analog of $c(g, G)$, we do not have $c(g, G, 1) = c(g, G)$.

When $G = SU(2)$, the formula reads

$$T_g(x; q) = (-1)^{g-1}\, 2^g \sum_{n \neq 0} \frac{e^{\underline{n}x}}{(q^n - q^{-n})^{2g-1}}. \tag{1.13}$$

Observe that for $g \geqslant 1$ this series converges to a holomorphic function on the unit disc in the complex q-plane. It is difficult to evaluate such a sum (although, cf. Remark 3). Instead, we will study the behavior of $T_g(x; q)$ as $q \to 1$. More precisely, we will compute the asymptotics of $T_g(x; e^{\pi i \hbar})$ as $\hbar \to i0^+$, i.e., as $\hbar$ approaches 0 along the ray of purely imaginary numbers in the upper half plane. To this end, consider the poles of the form $w_{2g-1}(x, \hbar^{-1}; u)$ in the complex u-plane for a fixed $\hbar \in i\mathbb{R}^+$. We divide the set of poles into three parts as follows:

1. $\mathrm{Pole}_1 = \{\underline{n} \mid n \in \mathbb{Z}, n \neq 0\}$;
2. $\mathrm{Pole}_2 = \{\underline{n}/\hbar \mid n \in \mathbb{Z}, n \neq 0\}$;
3. $u = 0$.

Clearly, we have

$$T_g(x; e^{\pi i\hbar}) = (-1)^g\, 2^g \sum_{p\in \mathrm{Pole}_1} \operatorname*{Res}_{u=p} w_{2g-1}(x, \hbar^{-1}; u).$$

Now consider $\sum_{p\in\mathrm{Pole}_2} \mathrm{Res}_{u=p}\, w_{2g-1}(x, \hbar^{-1}; u)$ and assume $x \notin \mathbb{Z}$. Since $e^{\{x\}\,u}/(1-e^u)$ and its first $2g$ derivatives vanish exponentially as $u \to \pm\infty$ on the real line and $(e^{\hbar u/2} - e^{-\hbar u/2})^{2g-1}$ is antiperiodic with period $2\pi/\hbar$, we have

$$\left|\operatorname*{Res}_{u=n/\hbar} w_{2g-1}(x, \hbar^{-1}; u)\right| < ce^{-\tau\,|n|/|\hbar|}$$

for some positive constants c and τ, independent from n. Then by summing the geometric series we obtain that

$$\left|\sum_{p\in\mathrm{Pole}_2} \operatorname*{Res}_{u=p} w_{2g-1}(x, \hbar^{-1}; u)\right| < 2ce^{-\tau/|\hbar|}$$

as $\hbar \to i0^+$. Finally, since outside a small neighborhood of its poles the function $e^{\{x\}\,u}/(1-e^u)$ vanishes exponentially as $|\mathrm{Re}(u)| \to \infty$, and so does $(e^{i\hbar u/2} - e^{-i\hbar u/2})^{1-2g}$ as $|\mathrm{Im}(u)| \to \infty$, we see that the line integral of $w_{2g-1}(x, \hbar^{-1}; u)$ over a sequence of appropriately chosen contours, e.g., the boundary of the rectangles

$$\mathrm{Rect}_L = \{|\hbar\,\mathrm{Re}(u)|, |\mathrm{Im}(u)| \leqslant (2L+1)\,\pi\}, \qquad L \in \mathbb{N},$$

goes to 0 as $L \to \infty$. We can summarize what we have found as follows:

PROPOSITION 1. *Given $x \notin \mathbb{Z}$, there are positive constants c and τ and a two-variable polynomial with rational coefficients $p(x, y)$, defined in* (1.12), *such that*

$$|T_g(x; e^{\pi i\hbar}) - p_{2g-1}(x, \hbar^{-1})| < ce^{-\tau/|\hbar|} \tag{1.14}$$

for sufficiently small $\hbar \in i\mathbb{R}^+$.

Remark 2. When x is an integer, the asymptotic expansion still exists, but the coefficients are not manifestly rational.

Remark 3. Note that the existence of such an expansion around $q=1$ for a holomorphic function of q on the unit disc is a rather rare occurrence. It strongly suggests that the function is related to modular forms. In fact, there are examples of such relations [42],

$$T_1(x, q) = \frac{d}{du} \log \theta_4(u/2, \tau), \qquad q = e^{i\pi\tau},$$

but we will not explore this connection in this paper. From the point of view of modular forms, the asymptotic behavior of T appears as a defect of sorts, a measurement of its failure to be modular.

We finish this section with an observation which will be central to our main result. Clearly, there is a formal analogy between $V_g^G(\lambda; k)$ and $T_g^G(x; \hbar)$, both of them being a trigonometric deformation of $E_g^G(x)$. Our residue calculations quantify this analogy in the case of $G = SU(2)$, as follows.

First, by shifting the variables we can rewrite (1.11) as

$$V_g(l-1; k-2) = 2(-2k)^{g-1}\, p_{2g-1}(\{l/2k\}, k), \tag{1.15}$$

for $l/2, k \in \mathbb{Z}^{\geqslant 0}$. On the other hand, (1.14) implies that asymptotically

$$T_g(x; e^{\pi i \hbar}) \sim 2(-2)^{g-1}\, p_{2g-1}(\{x\}, 1/\hbar) \qquad \text{as} \qquad \hbar \to i0^+. \tag{1.16}$$

A similar equality holds in the higher rank case. This will be shown in a later publication.

1.5. *Several Punctures*

Let P be a finite set. We can extend the results of this section as follows. We can write down a function of $\mathbf{x}\colon P \to \mathfrak{t}$

$$\widetilde{E}_g^G(\mathbf{x}) = \sum_{\lambda \in \Omega^+} \frac{\prod_{p \in P} \chi_\lambda(\tilde{\mathbf{x}}(p))}{(\dim V_\lambda)^{2g-2+|P|}} \tag{1.2P}$$

and the series

$$\begin{aligned} E_g^G(\mathbf{x}) &= c_{|P|}(g, G)\left(\prod_{p \in P} \delta(\tilde{\mathbf{x}}(p))\right) \widetilde{E}_g^G(\mathbf{x}) \\ &= \frac{(-1)^{(g-1)\,|\Delta^+|}\,|Z_G|^g}{|W_G|} \sum_{\mathbf{w}\colon P \to W_G} \operatorname{sign}(\mathbf{w})\, B^G_{2g-2+|P|}(\mathbf{w} \cdot \mathbf{x}), \end{aligned} \tag{1.5P}$$

where $\operatorname{sign}(\mathbf{w}) = \prod_{p \in P} \operatorname{sign}(\mathbf{w}(p))$ and $\mathbf{w} \cdot \mathbf{x} = \sum_{p \in P} \mathbf{w}(p)(\mathbf{x}(p))$. For $G = SU(2)$ we obtain the formula

$$E_g(\mathbf{x}) = (-2)^{g-1} \sum_{\varepsilon\colon P \to \pm 1} \operatorname{sign}(\varepsilon)\, B_{2g-2+|P|}(\varepsilon \cdot \mathbf{x}), \tag{1.6P}$$

where $\operatorname{sign}(\varepsilon) = \prod_p \varepsilon(p)$.

In the trigonometric case, one replaces $\lambda \in \Omega_\Delta$ by $\boldsymbol{\lambda}: P \to \Omega^+$ such that $\sum_{p \in P} \boldsymbol{\lambda}(p) \in \Omega_\Delta$ and writes a formula for $V_g(\boldsymbol{\lambda}; k)$ by replacing χ_λ in (1.8) by $\prod_{p \in P} \chi_{\boldsymbol{\lambda}(p)}(t)$. Now the formulas for $\tilde{T}_g^G(\mathbf{x}; q)$ and $T_g^G(\mathbf{x}; q)$ can be written down by analogy. We can summarize the final result as follows,

Proposition 2. *The Verlinde sums and the q-volume series for $G = SU(2)$ are both related to the same polynomial function as follows,*

$$V_g(\boldsymbol{\lambda} - \mathbf{1}, k-2) = (-2k)^{g-1} \sum_{\boldsymbol{\varepsilon}} \operatorname{sign}(\boldsymbol{\varepsilon})\, p_{2g-2+|P|}(\{\boldsymbol{\varepsilon} \cdot \boldsymbol{\lambda}/2k\}, k), \tag{1.15P}$$

where $\boldsymbol{\lambda} - \mathbf{1}$ is the $|P|$-tuple of integers, each component of which is smaller than the corresponding component of the vector $\boldsymbol{\lambda}$ by 1. Also, if for every choice of signs $\boldsymbol{\varepsilon}: P \to \pm 1$ the condition $\boldsymbol{\varepsilon} \cdot \mathbf{x} \notin \mathbb{Z}$ holds then

$$T_g(\mathbf{x}, e^{i\pi\hbar}) \sim (-2)^{g-1} \sum_{\boldsymbol{\varepsilon}} \operatorname{sign}(\boldsymbol{\varepsilon})\, p_{2g-2+|P|}(\{\boldsymbol{\varepsilon} \cdot \mathbf{x}\}, \hbar^{-1}), \tag{1.16P}$$

where $\sim$ means asymptotic equality as used above.

Remark 4. As we mentioned at the beginning of this section, the Bernoulli functions $B_m^G(x)$ restrict to polynomials on the complement of a Λ-periodic hyperplane arrangement on $\mathfrak{t}$. If $x \in \mathfrak{t}$ is on one of these hyperplanes, we will call it *special*. Because of the Λ-periodicity we may define an element $t \in T$ *special* if it is the exponential of a special element of $\mathfrak{t}$. Consistently with (1.6P) one may define $\mathbf{x}$, an l-tuple of elements of $\mathfrak{t}$, *special* if $\mathbf{w} \cdot \mathbf{x}$ is special for some l-tuple $\mathbf{w}$ of elements of W_G. We will give a more concrete definition of this property in the next section.

Note that according to Proposition 2, the asymptotic behavior of the q-volume sums also seems to be sensitive to the argument x being special. While we do not develop this idea further in this paper, it seems to be an interesting question to describe the behavior of these sums at special values of the argument.

2. MODULI SPACES OF FLAT CONNECTIONS

This section serves as a quick introduction to the topology of the moduli spaces of flat connections on Riemann surfaces. For more details [8] is a good reference. The first part, Section 2.1, is not essential for following the rest of the paper. We provide this material to familiarize the reader with the moduli spaces and to point out the relationship between the singularities of the Witten sums and those of the corresponding moduli spaces.

Keeping the notation of the previous section, let again G be a compact, simple, simply connected Lie group. Let Σ be a topological Riemann surface and $P \subset \Sigma$ be a finite nonempty set of points. We will use the shorthand $\Sigma_P = \Sigma \setminus P$; when P consists of a single point p, we will write Σ_p for $\Sigma \setminus P$. The moduli space of flat G-connections $\mathfrak{M}^G = \mathfrak{M}^G[\Sigma_P]$ modulo gauge transformations can be represented as the quotient of the space of representations of the fundamental group of Σ_P in G modulo the conjugations:

$$\mathfrak{M}^G = \mathrm{Hom}(\pi_1(\Sigma_P), G)/\mathrm{Ad}\, G.$$

We discuss the case of one puncture, $P = \{p\}$ first, and explain how the results generalize to several punctures at the end of the section.

There is a natural map $\mathrm{Hol}_p : \mathfrak{M}^G[\Sigma_p] \to \mathrm{Conj}(G)$ which assigns to a flat connection the conjugacy class of the holonomy around the puncture p. Denote by $\mathfrak{M}^G_\sigma$ the space $\mathrm{Hol}_p^{-1}(\sigma)$, where $\sigma \in \mathrm{Conj}(G)$. We will be mainly interested in the case of regular conjugacy classes, i.e., adjoint orbits of maximal dimension. In this case, $\mathfrak{M}^G_\sigma$ is (a possibly singular) manifold of dimension $(2g-1) \dim G - \mathrm{rank}(G)$.

2.1. *Topology and Singularities*

In this paragraph, we describe the family of spaces $\mathfrak{M}^G_\sigma$ as σ varies in $\mathrm{Conj}(G)$. We will pay particular attention to the question of smoothness, since we will discuss the Riemann–Roch calculus on these spaces later. This material is necessarily incomplete and somewhat informal. For complete details cf. [8, 28].

Choosing the standard presentation of $\pi_1(\Sigma_p)$, we can represent $\mathfrak{M}^G_\sigma$ as

$$\{[A_1, B_1] \cdots [A_g, B_g] \in \sigma \mid A_i, B_i \in G\}/\mathrm{Ad}\, G,$$

where $[A, B] = ABA^{-1}B^{-1}$. As in the previous section, sometimes we will replace the regular orbit σ by a representative $t \in T_{\mathrm{reg}}$. Then we have

$$\mathfrak{M}^G_{\sigma(t)} \cong \{[A_1, B_1] \cdots [A_g, B_g] = t \mid A_i, B_i \in G\}/\mathrm{Ad}\, T.$$

Since T is compact, the space $\mathfrak{M}^G_{\sigma(t)}$ is Hausdorff. Given an element $\mu \in \mathfrak{M}^G_{\sigma(t)}$ considered to be a representation of $\pi_1(\Sigma_p)$, denote by $\mathrm{Im}'(\mu)$ the commutator subgroup of the image of μ in G. The smoothness of $\mathfrak{M}^G_{\sigma(t)}$, not surprisingly, is determined by the set of possible groups, which arise as centralizers $Z(\mathrm{Im}'(\mu))$ of these subgroups in G.

Barring some degenerate cases, for a generic μ, we have $\mathrm{Im}'(\mu) = G$. The space $\mathfrak{M}^G_{\sigma(t)}$ is singular for some $t \in T_{\mathrm{reg}}$ whenever $Z(\mathrm{Im}'(\mu))$ is strictly greater than Z_G. Fix $t \in T_{\mathrm{reg}}$; there are two cases:

1. There is a $\mu \in \mathrm{Hol}_p^{-1}(t)$ such that $\dim Z(\mathrm{Im}'(\mu)) > 0$. This means that $\mathrm{Im}'(\mu)$ is a semisimple subgroup of G of strictly smaller rank than that of G. As $t \in \mathrm{Im}'(\mu)$, we may assume that t is in a maximal torus of $\mathrm{Im}'(\mu)$. Since the coroots of $\mathrm{Im}'(\mu)$ will be a subset of the coroots of G, we are led to the following definition: an element $x \in \mathfrak{t}$ is *special* if $x - \lambda$ is a linear combination of at most $\mathrm{rank}(G) - 1$ coroots of G for some $\lambda \in \Lambda$. Then it is natural to call $t \in T$ special if it is the exponential of some special element of $\mathfrak{t}$ (cf. Remark 4). It is easy to see that $\mu \in \mathrm{Hol}_p^{-1}(t)$ with the above properties exists if and only if t is special. We will call the singularities which appear in this case *serious*.

2. There is a μ such that $Z(\mathrm{Im}'(\mu))$ is finite, but strictly larger than Z_G. The existence of such μ does not depend on t, but on the group G only. More precisely, one needs a nontrivial element $z \in Z(\mathrm{Im}'(\mu))/Z_G$, i.e., an element which lies on $\mathrm{rank}(G)$ singular subsets $U_\alpha = \{e_\alpha = 1\} \subset T$, $\alpha \in \Delta$, but not in the intersection of all of the U_αs, which is Z_G. The existence of such an element z produces orbifold singularities. Note that since such an element does not exist for $G = SU(n)$, in this case, for a nonspecial t the space $\mathfrak{M}^G_{\sigma(\tilde{x})}$ is smooth.

Recall from (1.1) that the set of regular orbits in G is represented by an alcove $\mathfrak{a} \subset \mathfrak{t}$, which is in one-to-one correspondence with the set $\mathfrak{a}^* \subset \mathfrak{t}^*$ of dominant weights of height at most 1. By the above discussion, the set of those $x \in \mathfrak{a}$, for which $\mathfrak{M}^G_{\sigma(\tilde{x})}$ has no serious singularities, is the intersection of $\mathfrak{a}$ with the complement of a hyperplane arrangement; this has a rather complicated chamber structure. The spaces $\mathfrak{M}^G_{\sigma(\tilde{x})}$, where x varies in one of these chambers in $\mathfrak{a}$, are all isomorphic and form a trivial family. The spaces $\mathfrak{M}^G_{\sigma(\tilde{x})}$ and $\mathfrak{M}^G_{\sigma(\tilde{x}')}$ corresponding to two neighboring chambers differ by a set of high codimension and, in general, are not isomorphic. For details see [7, 8, 23, 36].

2.2. *Line Bundles and the Symplectic Form*

A useful approach to studying the topology of $\mathfrak{M}^G_{\sigma(\tilde{x})}$ is to represent it as a quotient via infinite dimensional symplectic reduction [1, 8, 20, 28]. In particular, such a representation induces a natural symplectic form ω_x on $\mathfrak{M}^G_{\sigma(\tilde{x})}$, whose normalization depends only on the normalization of the invariant symmetric bilinear form on $\mathfrak{g}$. We summarize the necessary facts about line bundles over the spaces $\mathfrak{M}^G_\sigma$ in the following proposition. For details consult [8, 28].

Proposition 3. *Let $x \in \mathfrak{a}$ be a nonspecial element. Then*

- *there is an identification $\eta\colon \mathfrak{t}^* \oplus \mathbb{R} \cong H^2(\mathfrak{M}^G_{\sigma(\tilde{x})}, \mathbb{R})$, such that*
- *the cohomology class of ω_x is $\eta(x^*, 1)$;*
- *for a weight $\lambda \in \Omega$ and $k \in \mathbb{Z}$, the cohomology class $\eta(\lambda, k)$ is the Chern class of an (orbifold-) line bundle $L_{\lambda,k}$ over $\mathfrak{M}^G_{\sigma(\tilde{x})}$;*
- *the following formula holds: $c_1(\mathfrak{M}^G_{\sigma(\tilde{x})}) = 2\eta(\rho, (\theta_G, \rho) + 1)$, independent from x.*

Now we can describe the connection of the formulas of the previous section with the topology of the moduli spaces. The first formula, proved by Witten [27, 43], says that the functions $E^G_g(x)$ give the volume of the moduli spaces with respect to the canonical symplectic structure:

$$E^G_g(x) = \int_{\mathfrak{M}^G_{\sigma(\tilde{x})}} e^{\omega_x}. \tag{2.17}$$

It follows from the proposition that the right hand side depends polynomially on x, as long as x stays in a chamber of nonspecial elements in $\mathfrak{a}$. The analogous statement about the left hand side was proved in [33].

The second formula computes the Riemann–Roch number of the line bundle $L_{\lambda,k}$ on the space $\mathfrak{M}^G_{\sigma(\tilde{x})}$, where $\lambda/k = x^*$ with x nonspecial:

$$V^G_g(\lambda; k) = \int_{\mathfrak{M}^G_{\sigma(\tilde{x})}} e^{\eta(\lambda,k)}\,\mathrm{Todd}(\mathfrak{M}^G_{\sigma(\tilde{x})}). \tag{2.18}$$

There are a number of approaches to the proof of this formula. One of them, initiated in [32], is based on establishing an efficient calculus of characteristic numbers of the moduli spaces using residues; Jeffrey and Kirwan [23] proved the formula for some weights when $G = SU(n)$, and Bismut and Labourie [8, 34] cover almost all cases via a similar approach. There is a complete treatment by Meinrenken and Woodward [28] using the notion of group-valued moment map. Finally, note that the original algebro-geometric interpretation $V^G_g(\lambda; k)$ was the dimension of the space of sections of the line bundle $L_{\lambda,k}$ endowed with an appropriate holomorphic structure (cf. [6, 8, 28]):

$$V^G_g(\lambda; k) = \dim H^0(\mathfrak{M}^G(t), L_{\lambda,k}). \tag{2.19}$$

One may conclude (2.18) then if armed with an appropriate vanishing theorem (cf. [35] for a related result).

Consider now the case of $G = SU(2)$. Using the fact that $\mathrm{Todd}(M) = \hat{A}(M)\, e^{c_1(M)/2}$ and comparing the formulas for ω_x and $c_1(M)$ given in Proposition 3 with (1.15), we obtain

$$2(-2)^{g-1}\, p_{2g-1}(x, k) = \int_{\mathfrak{M}^G_{\sigma(\tilde{x})}} e^{k\omega_x} \hat{A}(\mathfrak{M}^G_{\sigma(\tilde{x})}), \tag{2.20}$$

for $0 < x < 1/2$.

A similar formula holds for general G as well [8, 34].

2.3. *Several Punctures*

The case of several punctures is entirely analogous, so we will be very brief. In this case, the moduli space $\mathfrak{M}^G[\Sigma_P]$ is a union of the fibers $\mathfrak{M}^G_{\boldsymbol{\sigma}}$ of the map $\mathrm{Hol}_P : \mathfrak{M}^G[\Sigma_P] \to \mathrm{Conj}(G)^{\times |P|}$, which associates to each flat connection on Σ_P the $|P|$-tuple of conjugacy classes of holonomies around the punctures. The moduli space $\mathfrak{M}^G_{\boldsymbol{\sigma}(\tilde{\mathbf{x}})}$, where $\mathbf{x}: P \to \mathfrak{t}$, has serious singularities whenever for some $|P|$-tuple of Weyl group elements $\mathbf{w}: P \to W_G$ the sum $\sum_{p \in P} \mathbf{w} \cdot \mathbf{x}$ is a special element of $\mathfrak{t}$ (as defined above). Again, there is a canonical symplectic form $\omega_{\mathbf{x}}$ on $\mathfrak{M}^G_{\boldsymbol{\sigma}(\tilde{\mathbf{x}})}$ and an isomorphism $\eta: \mathfrak{t}^{\oplus |P|} \oplus \mathbb{R} \cong H^2(\mathfrak{M}^G_{\boldsymbol{\sigma}(\tilde{\mathbf{x}})}, \mathbb{R})$. The multiple puncture version of the formula for $G = SU(2)$ from above reads

$$(-2)^{g-1} \sum_{\boldsymbol{\varepsilon}: P \to \pm} \mathrm{sign}(\boldsymbol{\varepsilon})\, P_{2g-2+|P|}(\{\boldsymbol{\varepsilon} \cdot \mathbf{x}\}, k) = \int_{\mathfrak{M}_{\boldsymbol{\sigma}(\tilde{\mathbf{x}})}} e^{k\omega_{\mathbf{x}}} \hat{A}(\mathfrak{M}_{\boldsymbol{\sigma}(\tilde{\mathbf{x}})}), \tag{2.20P}$$

where again $0 < \mathbf{x}(p) < 1/2$ for all $p \in P$.

3. FORMAL DEFORMATIONS AND ALGEBRAIC MANIFOLDS

3.1. *Formal Deformations of Symplectic Manifolds*

Here we review the formal deformation theory of symplectic manifolds. A reference for this material is [19, 40].

For a complex vector space V, denote by $V[\![\hbar]\!]$ the space of formal power series in $\hbar$ with coefficients in V. A *formal deformation* (the terms *star product* or *deformation quantization* are also used) of a manifold M is a bilinear map "$\cdot$": $C^\infty(M) \otimes C^\infty(M) \to C^\infty(M)[\![\hbar]\!]$, which, when extended linearly to $C^\infty(M)[\![\hbar]\!]^{\otimes 2}$, is

- associative: $f \cdot (g \cdot h) = (f \cdot g) \cdot h$;
- local: $f \cdot g = fg + \sum_{n=1}^{\infty} B_n(f, g)\, \underline{\hbar}^n$, where B_n is a bidifferential operator.

It is easy to see that $B_1(f, g) - B_1(g, f) = \{f, g\}$ is a Poisson bracket on $C^\infty(M)$. When (M^{2n}, ω) is a symplectic manifold, and the deformation is such that the induced Poisson bracket is the symplectic one, then we speak of a formal deformation of this symplectic manifold.

Remark 5. Note the factor of $2\pi i$ in front of $\hbar$, marked by underlining, inside the expansion in the definition of locality. This is at variance with most conventions in the literature. We chose it because it is consistent with $\hbar$ being real and with the symplectic form ω being integral in our applications.

The associativity condition induces an infinite set of complicated nonlinear equations on the B_ns, which were explicitly solved in local coordinates by Kontsevich [26]. The existence of the solution in the symplectic case was proved in [14, 21] (cf. [40] for references).

The group of formal base changes (gauge transformations)

$$\mathscr{G}(M) = \left\{ \gamma \colon f \mapsto f + \sum_{n=1}^{\infty} \gamma_n(f)\, \underline{\hbar}^n \mid \gamma_n \text{ is a differential operator} \right\} \tag{3.21}$$

acts on the space of star products via $f \cdot_\gamma g = \gamma(\gamma^{-1} f \cdot \gamma^{-1} g)$. A natural question is the classification of formal deformations up to this action.

From here on, we only study the symplectic case (M^{2n}, ω). In this case, we have a complete classification:

THEOREM 1 [13, 14, 24, 29, 41]. *The orbits of the action* (3.21) *on the set of formal products are in a one-to-one correspondence with* $H^2(M)[\![\hbar]\!]$.

It turns out to be natural to associate to each formal deformation a *characteristic class* $\theta(M, \cdot)$ equal to $\omega/\hbar$ plus the element of $H^2(M)[\![\hbar]\!]$ mentioned in the theorem.

Another interesting related object is the *canonical trace* on the algebra $(C^\infty(M)[\![\hbar]\!], \cdot)$. A functional T on a noncommutative ring is called *cyclic* or a *trace* if $T(a \cdot b) = T(b \cdot a)$. Such functionals on $(C^\infty(M)[\![\hbar]\!], \cdot)$ form a 1-dimensional free module over $\mathbb{C}[\![\hbar]\!]$ with a distinguished element. To understand this we recall the local theory of formal deformations of symplectic manifolds.

Let $M = \mathbb{R}^{2n}$ with a translation invariant symplectic form ω. Then the famous Moyal product is given by

$$f \cdot_{\mathrm{Mp}} g = m(e^{\underline{\hbar} \pi_\omega} f \otimes g),$$

where $\pi_\omega = \omega^{-1}$ is the translation invariant Poisson bivector field induced by ω and m is the ordinary commutative product on $C^\infty(\mathbb{R}^{2n})[\![\hbar]\!]^{\otimes 2}$. Denote the resulting algebra $(C^\infty(\mathbb{R}^{2n})[\![\hbar]\!], \cdot_{\mathrm{Mp}})$ by A_{Mp}. We collect the highlights of the local theory in the theorem below.

THEOREM 2 [19, 29]. 1. *The Moyal product is an associative star product on the symplectic space* $(\mathbb{R}^{2n}, \omega)$.

2. *Any formal deformation of* $(\mathbb{R}^{2n}, \omega)$ *is gauge-equivalent to* A_{Mp}.

3. *For every derivation* D *of* A_{Mp}, *where* $D = \sum_{n=1}^{\infty} D_n \hbar^n$ *and* $D_n \in$ Diff$(\mathbb{R}^{2n})$, *there exists an element* $f \in A_{\mathrm{Mp}}$ *such that*

$$Dg = f \cdot_{\mathrm{Mp}} g - g \cdot_{\mathrm{Mp}} f$$

for every $g \in A_{\mathrm{Mp}}$. *Similarly, any local automorphism of* A_{Mp} *of the form* $1 + \sum_{n=1}^{\infty} A_n \hbar^n$ *can be obtained by exponentiating such a derivation.*

4. *Integration against the density* $\omega^n/\hbar^n$ *defines a trace functional on* A_{Mp}, *which is unique up to an* $\hbar$*-dependent constant.*

The particular choice of normalization $f \to \int f \omega^n/\hbar^n$ is called the *canonical trace* on the Moyal product. By the above theorem, this notion can be extended to an arbitrary formal deformation of $(\mathbb{R}^{2n}, \omega)$ via the isomorphism from statement (2), and this notion is well defined by statement (3). Finally, using Darboux's theorem, we can define the canonical trace on a formal deformation of an arbitrary symplectic manifold (M^{2n}, ω) by requiring that the pull-back of such a functional tr: $C^\infty(M)[\![\hbar]\!] \to \hbar^{-n}\mathbb{C}[\![\hbar]\!]$ with respect to a symplectic embedding of an open subset of $\mathbb{R}^{2n}$ into M is a canonical trace on (an open subset of) $\mathbb{R}^{2n}$.

PROPOSITION 4 [29]. *For every formal deformation* $(C^\infty(M)[\![\hbar]\!], \cdot)$ *of a symplectic manifold* (M, ω) *the canonical trace exists and is unique.*

We denote the canonical trace by $\mathrm{Tr}_{\mathrm{can}}$. Now we can formulate the Fedosov–Nest–Tsygan index theorem [19, 29] for a compact symplectic manifold M. It relates the two objects defined above:

$$\mathrm{Tr}_{\mathrm{can}}(1) = \int_M e^\theta \hat{A}(M). \tag{3.22}$$

This is an analog of the Hirzebruch–Riemann–Roch formula, which we have already used in the previous section. For a compact symplectic manifold (M, ω) with integral symplectic form, let L be a line bundle whose first Chern class is ω. If M has a compatible Kähler structure and L

is endowed with an appropriate holomorphic structure, then, for large k, the HRR theorem gives the following expression for the dimension of the space of sections of L^k:

$$\dim H^0(L^k) = \int_M e^{kc_1(L)}\,\mathrm{Todd}(M). \tag{3.23}$$

This expression is an integer valued polynomial which only depends on the symplectic structure. Note that the RHS, which in the projective algebraic case is also known as the Hilbert polynomial, can be defined for any symplectic manifold.

An alternative expression of the same type can be given using the $\hat{A}$-genus:

$$\dim H^0(L^k) = \int_M e^{kc_1(L)+c_1(M)/2}\hat{A}(M).$$

Define the shifted Hilbert polynomial as

$$P_L(k) = \int_M e^{kc_1(L)}\hat{A}(M). \tag{3.24}$$

Now define a deformation $(M, \cdot)$ of a symplectic manifold *basic* if $\theta(M, \cdot) = \omega/\hbar$. Then, combining the above formulas, we obtain that given a basic deformation of a compact symplectic manifold (M^{2n}, ω) and a line bundle L on M with $c_1(L) = \omega$, we have [19, 29]

$$\mathrm{Tr}_{\mathrm{can}}(1) = P_L(\hbar^{-1}). \tag{3.25}$$

3.2. *Algebraic Manifolds and Local Deformations*

For the purposes of our paper we need to reformulate the theory we just outlined.

Let A_0 be a subalgebra of $C^\infty(M)$ which separates points and is not in the kernel of any nonzero complex differential operator on M. Let $A^\hbar$ be an $\hbar$-adically complete associative algebra over $\mathbb{C}[[\hbar]]$, such that $A^\hbar/\hbar A^\hbar \cong A_0$. Assume that there exists a section $s\colon A_0 \to A^\hbar$ of the natural map $A^\hbar \to A_0$ which, when extended to $A_0[[\hbar]]$, gives an isomorphism $s_\hbar\colon A_0[[\hbar]] \simeq A^\hbar$. Then the formula $f \cdot_s g = s^{-1}(s(f)\, s(g))$ defines a product on $A_0[[\hbar]]$. If this product is local (see the definition at the start of this section), then we say that s is a *local* section.

DEFINITION 1. An algebra $A^\hbar$ which *has* a local section is called a *local deformation* of A_0.

As is clear from the construction above, the data $(A^\hbar, s)$ of an algebra with a local section such that $A_0 = C^\infty(M)$ are equivalent to that of a formal deformation of the manifold M.

LEMMA 3. *The action of the gauge group defined in* (3.21) *corresponds to the action* $s \to s_\gamma = s \circ \gamma^{-1}$.

The proof is clear. This means that the classification of formal deformations of symplectic manifolds up to gauge transformations is equivalent to the classification of local deformations of the algebra $C^\infty(M)$. Then the following statement is immediate:

PROPOSITION 5. *The Poisson structure, the characteristic class, and the canonical trace* $A^\hbar \to \mathbb{C}[\![\hbar]\!]$ *induced by a pair* $(A^\hbar, s)$ *is independent of* s.

The interest in this statement lies in the possibility of studying these invariants in examples of local deformations $A^\hbar$ which do not have natural sections.

Such examples could arise in the following setting. Suppose that (M, ω) is a smooth *real* affine algebraic manifold with a symplectic Poisson structure on the space of algebraic functions. Denote by A_0 the space of *complex* algebraic functions on M. While this is somewhat arbitrary, we choose our deformation ring to be

$$D(q) = \{\text{Rational functions in } q, \text{ with no poles at } q = 1 \text{ and } 0 < |q| < 1\}. \tag{3.26}$$

A noncommutative algebra A^q over $D(q)$, which is given by a finite number of generators and relations, such that $A^q/(q-1)\, A^q \cong A_0$ is often called a *q-deformation* of A_0. It is a *local q-deformation* if, in addition, the $\hbar$-adic completion $A^\hbar = A^q \mathbin{\widehat{\otimes}}_{D(q)} \mathbb{C}[\![\hbar]\!]$, where $q = e^{\frac{\hbar}{\cdot}}$, is a local deformation of A_0 in the sense defined above. Assuming that the Poisson structure derived from $A^\hbar$ is the given symplectic one, we can ask the following questions:

1. What is the characteristic class of A^q, which is defined as the characteristic class of $A^\hbar$?
2. Can the canonical trace be defined on A^q?

The first question is clear, although such computations are very difficult. The second question requires some comment. One could hope to start with a cyclic functional $\mathrm{Tr}^q\colon A^q \to D(q)$ and by Taylor expansion obtain a functional $\mathrm{Tr}^\hbar\colon A^\hbar \to \mathbb{C}[\![\hbar]\!]$. Then, after choosing a local section s, one could arrive at a functional $\mathrm{Tr}^\hbar\colon A_0[\![\hbar]\!] \to \mathbb{C}[\![\hbar]\!]$. Next, one needs to show that such a functional extends as a cyclic functional to $C^\infty(M)[\![\hbar]\!]$ and then finally, one could see if $\hbar^{-n}\,\mathrm{Tr}^\hbar$ is the canonical trace or not. Of course, it could very well happen that A^q does not have any cyclic functionals at all.

AN EXAMPLE. The simplest nontrivial example with which we are familiar is the quantum torus. The 2-dimensional torus can be written as an affine variety with generators $\{U^\pm, V^\pm\}$ and relations $\{U^+U^- = V^+V^- = 1\}$. It has a canonical translation invariant symplectic form ω, which we normalize so that $\int \omega = 1$. The noncommutative deformation has the same, but now noncommutative, generators and relations, and an additional relation: $UV = qVU$. Clearly, the monomials $\{U^mV^n \mid m, n \in \mathbb{Z}\}$ form a basis of this algebra and this gives the necessary section.

PROPOSITION 6. 1. *This deformation is local, and the derived Poisson structure is the one, corresponding to ω.*

2. *The functional* $\mathrm{Tr}^q(U^mV^n) = \delta_m\delta_n$ *is cyclic.*

3. *When passing to* $\mathbb{C}[\![\hbar]\!]$ *via* $q = e^\hbar$, *this functional is given by integration against ω and it is $\hbar$ times the canonical trace.*

4. *The characteristic class of this deformation is $\omega/\hbar$; thus the deformation is basic.*

We leave the proof of the proposition as an exercise to the reader. The first two statements are easy, and the third one is doable. Note that statement (3) and the F-NT index theorem imply statement (4).

This setup turns out to be insufficient for our purposes in this paper. As explained in the Introduction, one needs to generalize it by allowing the trace Tr^q to take values in a module K defined over $D(q)$ and endowed with a structure map $\mathrm{ev}\colon K \to \mathbb{C}[\![\hbar]\!]$. The appropriate module K in our case turns out to be the space of meromorphic functions on the unit disc $\{|q| < 1\}$, which have an asymptotic expansion in $q-1$ as q approaches 1 along the real axis. The function $T_g(x; q)$ introduced in (1.13) is an example of such a function. Then one can pass to values in $\mathbb{C}[\![\hbar]\!]$ by setting $q = e^{\underline{\hbar}}$ and taking the asymptotic expansion at $h = 0$ along the imaginary axis.

In this paper, we construct q-deformations of the moduli spaces of flat connections $\mathfrak{M}_\sigma^G[\Sigma_P]$ and study the questions raised above for these noncommutative algebras. We review our results in Section 7.

4. G-COLORED GRAPHS

4.1. *Functions on the Moduli Space*

In this section we introduce the notion of *G-colored graphs*, which is a graphical method to represent the set $F(\mathfrak{M}^G)$ of algebraic functions on the moduli space $\mathfrak{M}^G$. This is based on the notion of *graph connections* of Fock and Rosly [20], which were introduced as a discretization of gauge theory on Riemann surfaces. Our version is substantially similar but technically more flexible than the original one. A similar construction can be found in [4], where ordinary circular holonomies are used and an almost identical notion appears in [30]. Still there are some technical differences and the notion of equivalence seems to be new. Our proofs will be brief since they are similar to the original ones.

We keep the notation of Section 2.

A *G-colored graph* f, or G-graph for short, on $\Sigma_P = \Sigma \setminus P$ consists of

- an oriented, not necessarily connected graph $\Gamma = \Gamma(f)$, immersed into the punctured Riemann surface: $\rho: \Gamma \to \Sigma_P$. We will often informally identify the Γ with the image $\rho(\Gamma)$. Denote the set of edges of $\Gamma(f)$ by $E_\Gamma(f)$ and the set of its vertices by $V_\Gamma(f)$.
- a coloring of each edge $e \in E_\Gamma(f)$ by a representation $C_f(e)$ of the group G.
- a coloring of each vertex $v \in V_\Gamma(f)$ by an invariant

$$\phi_f(v) \in \left(\bigotimes_{e \to v} C_f(e)^* \bigotimes_{e \leftarrow v} C_f(e) \right)^G,$$

where the tensor products are taken over the incoming and outgoing edges correspondingly.

Remark 6. Note that the definition of the coloring of the vertices given above is somewhat imprecise, since we did not specify the order in which the tensor products are taken. We are taking advantage of the fact that the space of invariants of a tensor product of representations of G are naturally isomorphic to the space of invariants of the same representations tensored in a different order. In fact, using the orientation of the surface, we have a natural *cyclic orientation* of the edges adjacent to a vertex, but this still does not provide us with an ordering. This might seem like hair splitting here, but we will have to return to this question in the next section, where it becomes essential.

Given an immersed graph Γ in Σ_P with colored edges $e \mapsto C(e)$, and a connection ∇ on the trivial G-bundle over Σ_P, one can construct an element

$$\nabla_\Gamma \in \bigotimes_{e \in E_\Gamma} \mathrm{Hom}(C(e), C(e)) \cong \left(\bigotimes_{e \to v} C_f(e) \bigotimes_{e \leftarrow v} C_f(e)^* \right)$$

by taking the parallel transports of ∇ along the oriented edges of Γ. Then given a G-colored graph f, we obtain a number $f(\nabla)$ by pairing ∇_Γ with $\bigotimes_{v \in V_\Gamma(f)} \phi_f(v)$.

LEMMA 4. 1. *The number $f(\nabla)$ does not change if ∇ is replaced by a gauge-equivalent connection.*

2. *Let f and g be G-graphs and define $f \cup g$ to be the G-colored graph with $\Gamma_{f \cup g} = \Gamma(f) \cup \Gamma(g)$ and with coloring inherited from f and g. Then $(f \cup g)(\nabla) = f(\nabla)\, g(\nabla)$.*

3. *If ∇ is flat then $f(\nabla)$ is invariant under homotopic changes of the immersion of $\Gamma(f)$.*

The proofs are straightforward and will be omitted.

Denote by $F(G, \Sigma_P)$ the free complex vector space generated by all G-colored graphs modulo smooth homotopy of immersion $\rho\colon \Gamma \to \Sigma_P$. Then taking unions of the underlying graphs as in the lemma endows this space with an algebra structure. According to statement (1) of the lemma, every element of this space defines a function on $\mathfrak{M}^G[\Sigma_P]$. These functions are clearly algebraic; it is easy to see that there is a surjective homomorphism $F(G, \Sigma_P) \to F(\mathfrak{M}^G[\Sigma_P])$ to the space of algebraic functions on $\mathfrak{M}^G[\Sigma_P]$.

To describe the kernel of this map, fix an open disc $D \subset \Sigma_P$ with boundary ∂D a smooth embedded circle. Consider a G-graph f in generic position with respect to D. This means that every edge can intersect ∂D only transversally and only finitely many times. Then we can define a new, contracted G-colored graph $f\,|\,D$ with underlying immersed graph the quotient graph $\Gamma/(D \cap \Gamma)$ obtained by contracting to a single vertex all vertices and all edges which are entirely inside D. This has a single vertex in D and the edges adjacent to this vertex correspond to the points of intersection of the edges of Γ with ∂D. The colorings of the edges and vertices outside D are inherited from Γ, while the coloring of the new vertex can be obtained by contracting each edge e in D using the canonical diagonal element $\delta(C(e))$ in $C(e)^\star \otimes C(e)$. We will call two G-colored graphs, related by such a contraction, or a sequence of such contractions, *equivalent*. We denote this relation of equivalence by $\sim$; thus, in particular, we have $f \sim f \mid D$. Note that since we are free to move the graphs homotopically, the position of the disc D does not play any role.

Below we list a few important special cases of equivalence among G-colored graphs. The proofs are simple exercises and will be omitted.

EXAMPLE 1. 1. A G-graph with an edge colored by the trivial representation is equivalent to the same G-graph with this edge erased.

2. We can obtain an equivalent G-graph by placing a 2-valent vertex colored by the diagonal element $\delta(V)$ on any edge colored by a representation V.

3. For two crossing edges, we can place a vertex at the intersection, colored by the permutation $P_{VW}: V \otimes W \to W \otimes V$.

4. If the graph $\Gamma(f)$ is contractible on the surface Σ_P, then f is equivalent to a number. In particular, a contractible loop colored by the representation V_λ is equivalent to the number $\dim V_\lambda$. Note that a loop maybe converted into a graph using (2).

5. Fix an embedded interval $I \subset \Sigma_P$ and assume that $\Gamma(f)$ is in generic position with respect to I. For simplicity, also assume that the m edges intersecting I are similarly oriented with respect to I; denote their colorings by $V_1, ..., V_m$. Then f is equivalent to $f \dagger I$ where $\Gamma(f \dagger I)$ is obtained from $\Gamma(f)$ by introducing two new $m+1$-valent vertices S and E joined by an edge e, as shown in Fig. 1. (The interval I is represented by a thick line.) The element $f \dagger I \in F(G, \Sigma_P)$ is a sum of colored graphs which retain the colorings of those edges and vertices which are common with f and thus have the form $g(f, V, \phi(S), \phi(E))$. Here $V = C(e)$, $\phi(S) \in \mathrm{Hom}(V_1 \otimes \cdots \otimes V_m, V)$ and $\phi(E) \in \mathrm{Hom}(V, V_1 \otimes \cdots \otimes V_m)$. These two

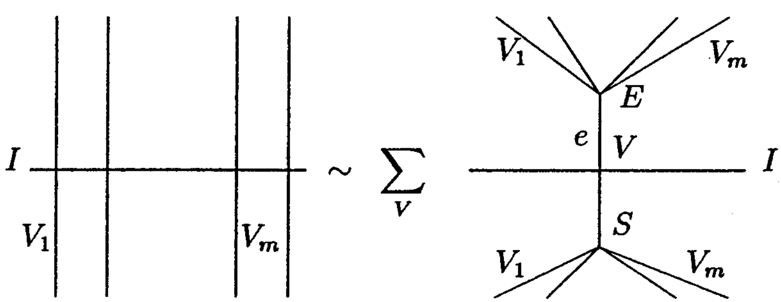

FIG. 1. f and $f \dagger I$.

spaces of invariants are naturally paired to each other via the formula $\mathrm{Tr}_V(\phi(S)\,\phi(E))$; thus we can define $f \dagger I$ by the finite sum

$$f \dagger I = \sum_{V \in \mathrm{Irrep}(G)} \sum_i g(f, V, \psi^i, \psi_i), \tag{4.27}$$

where $\sum_i \psi^i \otimes \psi_i = \delta_{\mathrm{Tr}}(\mathrm{Hom}(V_1 \otimes \cdots \otimes V_m, V))$ is the diagonal element induced by this pairing.

Remark 7. Note that the pairing $\mathrm{Tr}_V(\phi(S)\,\phi(E))$ is somewhat redundant, since by Schur's lemma we have

$$\phi(S)\,\phi(E) = \frac{\mathrm{Tr}_V(\phi(S)\,\phi(E))}{\dim V}\,\mathrm{id}_V. \tag{4.28}$$

Denote by $\mathrm{Col}(\Gamma)$ the set of all possible colorings of the edges of a fixed immersed graph Γ by *irreducible* representations and by $\Phi(\Gamma, C)$ the linear space of colorings of the vertices of a graph Γ with a fixed coloring $C \in \mathrm{Col}(\Gamma)$ of its edges. According to the following proposition the kernel of the map discussed above is generated by our notion of equivalence.

PROPOSITION 7. 1. *Two equivalent G-colored graphs f and $f \mid D$ take the same values on any flat connection ∇.*

2. *Let $\Gamma \subset \Sigma_P$ be an* embedded *graph such that each* face *of Γ is contractible and contains exactly one puncture. Then*

$$F(\mathfrak{M}^G[\Sigma_P]) = \bigoplus_{C \in \mathrm{Col}(\Gamma)} \Phi(\Gamma, C).$$

3. *The kernel of the map $F(G, \Sigma_P) \to F(\mathfrak{M}^G[\Sigma_P])$ is linearly generated by equivalence.*

DEFINITION 2. Define the graphs satisfying the condition in the 2nd statement *exact*. Such graphs always exist as long as there is at least one puncture. For an exact graph $\Gamma \subset \Sigma_P$ define the *dual graph* $\check{\Gamma}$ to be a graph embedded into Σ with vertices at the punctures (the set P) and faces containing exactly one vertex of Γ each.

Proof. (1) If we trivialize ∇ over D, then the parallel transports of ∇ along the edges will be all equal to the identity element of G. Then performing the partial contractions in the definition of $f(\nabla)$ by contracting along the edges in D only, we arrive at $f \mid D(\nabla)$.

(2) This was pointed out in [20]. The statement follows from the Peter–Weyl theorem since $\mathfrak{M}^G$ is simply a product of groups divided by the diagonal adjoint action. Note that coloring each edge by the trivial representation and the vertices by the trivial invariants gives the unit element of the algebra.

(3) It follows from (2) that it is sufficient to prove that given an exact graph Γ, any element $f \in F(G, \Sigma_P)$ is equivalent to a sum of G-colored graphs with underlying graph Γ. Let $\check{\Gamma}$ be the dual graph as described in the definition above. Put the given G-colored graph f into general position with respect to $\check{\Gamma}$. Then by performing the $\dagger$ operation on f with respect to each edge of $\check{\Gamma}$ we obtain a new G-colored graph $f_{\check{\Gamma}}$ (or a sum of such) which intersect each edge of $\check{\Gamma}$ exactly once and which is equivalent to f. Finally, using equivalence again, we can replace each of the vertices in each face of $\check{\Gamma}$ by a single vertex, thus obtaining a sum of G-colored graphs, each having exactly one vertex in each face of $\check{\Gamma}$. ∎

For each puncture $p \in P$ and representation V define an element $c^p_V \in F(G, \Sigma_P)$ with $\Gamma(c^p_V)$ a small counterclockwise oriented circle around p colored by V. We will also use the notation c^p_λ when $V = V_\lambda$. The graph underlying the product $c^p_V c^p_W$, is the union of two small concentric circles around p. The colored graph $(c^p_V c^p_W) \dagger I$, where I is an interval which intersects the two circles transversally, has four vertices. Contracting the two pairs of vertices joined by two edges, it is easy to see that $c^p_V c^p_W$ is equivalent to $c^p_{V \otimes W}$. Thus the correspondence $V \mapsto c^p_V(G)$ is, in fact, a homomorphism of algebras from $R(G)$ to $F(G, \Sigma_P)$.

4.2. *The Poisson Structure*

Now we define a Poisson structure on the space of G-colored graphs [4, 20, 30]. Fix an element $t \in (\mathrm{Sym}^2(\mathfrak{g}))^G$. Then for two G-colored graphs f and g in general position, and a point $m \in \Gamma(f) \cap \Gamma(g)$, we can define a new G-colored graph $f \cup^t_m g$ which is obtained from $f \cup g$ by placing a vertex at m, colored by

$$P_{12} \circ t : C(e_f(m)) \otimes C(e_g(m)) \to C(e_g(m)) \otimes C(e_f(m)),$$

where $e_f(m)$ and $e_g(m)$ are the two edges containing m and P_{12} is the permutation operator.

Now define

$$\{f, g\} = \sum_{m \in \Gamma(f) \cap \Gamma(g)} \mathrm{sign}(e_f(m), e_g(m))\, f \cup^t_m g, \tag{4.29}$$

where the sign is obtained by comparing the orientations of the ordered pair $(e_f(m), e_g(m))$ to the orientation of Σ.

PROPOSITION 8. 1. *The operation* $f, g \mapsto \{f, g\}$ *is well defined on* $F(\mathfrak{M}^G[\Sigma_P])$*; i.e., it is compatible with homotopy and equivalence.*

2. *The operation* $\{f, g\}$ *is a Poisson bracket on* $F(\mathfrak{M}^G[\Sigma_P])$.

3. *The elements* c_V^p, $p \in P$ *generate a subalgebra in the Poisson center of* $F(\mathfrak{M}^G[\Sigma_P])$ *in the sense that* $\{c_V^p, f\} = 0$ *for all* $f \in F(\mathfrak{M}^G[\Sigma_P])$.

4. *The spaces* $\mathfrak{M}^G_{\sigma(\tilde{x})} \subset \mathfrak{M}^G[\Sigma_P]$ *are symplectic leaves of this Poisson structure for generic* x*, and the induced symplectic form is exactly* $\omega_{\tilde{x}}$ *(cf. Section 2).*

Proof. (1) To prove compatibility with homotopy, it is sufficient to show the property shown in Fig. 2, where vertices colored by t are marked by circles. This easily follows from the identity

$$t_{13} + t_{23} = (\Delta_0 \otimes \mathrm{id})(t) \in U(\mathfrak{g})^{\otimes 3},$$

where Δ_0 is the coproduct in the universal enveloping algebra $U(\mathfrak{g})$. The indices, as usual, mark the embedding of $V^{\otimes 2}$ into $V^{\otimes 3}$; e.g., $t_{13} = P_{23}(t \otimes \mathrm{id})$. The compatibility with equivalence follows from this because using homotopy, one can always arrange that there be no intersection points in D.

(2) This statement follows from simple combinatorics of the intersection points (cf. [4, 30]).

(3) Again, using homotopy we can arrange that an arbitrary graph does not intersect a small circle around p.

(4) This statement is one of the main results of [20]. ∎

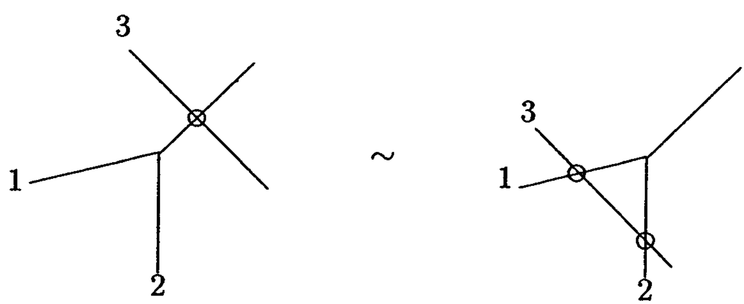

FIGURE 2

4.3. *The Poisson Trace*

The second statement in Proposition 7 allows us to define an augmentation $H_\Gamma\colon F(\mathfrak{M}^G[\Sigma_P]) \to \mathbb{C}$ by projecting onto the exact graph Γ colored by trivial representations and invariants. We also give a more constructive formula for $H_\Gamma(f)$ for a general $f \in F(G, \Sigma_P)$ which will be useful for computations later on. First, we define a variant of the operation $\dagger I$, denoted by $\dagger_0 I$ and called *cutting*, which is similar to $\dagger$ with the difference that V is allowed to be the trivial representation only. Thus (4.27) is modified as follows,

$$f \dagger_0 I = \sum_i g(f, \mathbb{C}, \psi^i, \psi_i), \tag{4.30}$$

where $\sum_i \psi^i \otimes \psi_i = \delta_{\mathrm{Tr}}(\mathrm{Hom}(V_1 \otimes \cdots \otimes V_m, \mathbb{C}))$.

Remark 8. We present a schematic picture of the cutting operation in Fig. 3. Note that now we can erase the edge between S and E, since it is colored by the trivial representation. However, as a mnemonic for the diagonal element δ_{Tr} that is inserted, we join the two vertices by a dashed line (chord).

Now assume that Γ is exact, and suppose that the graph $\Gamma(f)$ is in a generic position with respect to the dual graph $\check{\Gamma}$. By applying the cutting operation with respect to each edge of $\check{\Gamma}$ we obtain a new G-colored graph

$$f \dagger_0 \check{\Gamma} = f \prod_{e \in E_{\check{\Gamma}(f)}} \dagger_0 e.$$

The graph underlying $f \dagger_0 \check{\Gamma}$ is a union of disjoint pieces, each located on some contractible face. Thus $f \dagger_0 \check{\Gamma}$ is equivalent to a number.

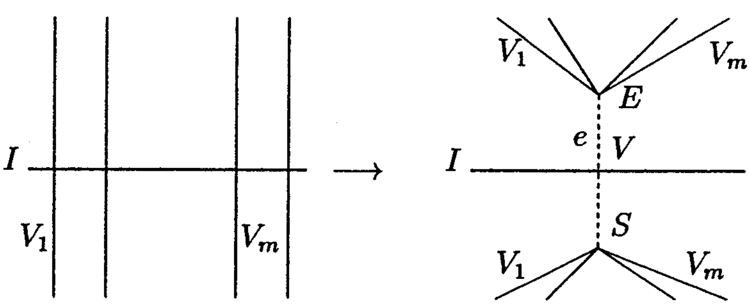

FIG. 3. Cutting along I.

PROPOSITION 9. 1. *The operation* $f \dagger_0 \check{\Gamma}$ *is well defined on* $F(\mathfrak{M}^G[\Sigma_P])$; *i.e., it is compatible with homotopy and equivalence.*

2. *The G-colored graph* $f \dagger_0 \check{\Gamma}$ *is equivalent to the number* $H_\Gamma(f)$.

3. $H_\Gamma(f)$ *does not depend on the choice of the exact graph* Γ.

4. $H_\Gamma(f)$ *is given by integration against a smooth top form on the smooth part of* $\mathfrak{M}^G[\Sigma_P]$.

5. *For any* f, g, $H(\{f, g\}) = 0$.

Proof. (1) Clearly compatibility with equivalence holds with respect to any disc D which lies entirely in one of the faces of $\check{\Gamma}$. Using this, we may assume that each face contains only one vertex. Then the only relevant homotopy relation is moving one of these vertices across one of the cutting edges. We leave proving this case as an exercise to the reader.

(2) Because of part (1) we may assume that $\Gamma(f) = \Gamma$, in which case the statement is obvious.

(3) It is easy to see that any two exact graphs are related by the operation of contracting a single edge in some exact graph Γ. This corresponds to removing an edge in $\check{\Gamma}$, and the statement now follows from compatibility with equivalence. Hence from now on we can omit the index Γ in H_Γ.

(4) Observe that there is another simple way to define H_Γ. It is clear from our earlier discussion that the space $\mathfrak{M}^G$ is a quotient of a product of copies of the group G, corresponding to the edges of Γ by the action of a product of copies of the group G corresponding to the vertices of Γ. Since the only matrix coefficient on a compact Lie group whose integral does not vanish is that of the trivial representation, we see that the measure induced on $\mathfrak{M}^G$ by the operation H_Γ is simply the push-forward of the Haar measure on the product of the groups. This push-forward is clearly smooth whenever the action is locally trivial.

(5) The proof of this statement is analogous to that of Proposition 12 (cf. Remark 10). It follows from the fact that a self-intersecting edge colored by V_λ, with the tensor $P_{12} \circ t$ inserted at the intersection point, is equivalent to an edge without self-intersection, colored the same way and multiplied by $-2C(\lambda)$, where $C(\lambda)$ is the value of the Casimir operator, normalized using t, on V_λ. This, in turn, follows from the equality $t = \Delta C - 1 \otimes C - C \otimes 1$. There is a somewhat exotic proof of this statement in [30]. ∎

Consider now the following general situation. Assume that $\pi: M \to N$ is a fibration between two compact, smooth manifolds, and assume that the manifolds are endowed with smooth volume forms μ_M and μ_N such that the volume of both manifolds is 1. Then one can define the push-forward operation on continuous functions $\pi_*: C^0(M) \to C^0(N)$ by integrating with respect to the natural measure $\mu_M / \pi^* \mu_N$ along the fibers. The proof of the following formulas will be omitted:

LEMMA 5. *Let $\{c_i, c^i\}_{i=0}^{\infty}$ be dual bases of functions on N; i.e., assume that $\int_N c^i c_j \, \mu_N = \delta_{ij}$ and the functions $\{c_i\}$ are complete in $L^2(N, \mu_N)$. Then*

1.

$$\pi_*(f) = \sum_i \left(\int_M f \pi^*(c^i) \, \mu_M \right) c_i \tag{4.31}$$

2. *The permanence equation*

$$\pi_*(f \pi^*(g)) = \pi_*(f) \, g \tag{4.32}$$

holds for any $f \in C^0(M)$ and $g \in C^0(N)$.

For notational simplicity we will concentrate on the $|P| = 1$ case from here on. Set $N = \mathrm{Conj}_{\mathrm{reg}}(G)$ with measure induced by the Haar measure on G, $\pi = \mathrm{Hol}_p$ and $M \subset \mathfrak{M}^G[\Sigma_p]$ the smooth part of $\mathrm{Hol}_p^{-1}(N)$ with the measure defined by H above. While the technical conditions of the lemma do not hold, the conclusions do (cf. [8, Theorem 4.2]); thus we have

PROPOSITION 10. *Define a functional* Tr *on $F(\mathfrak{M}^G[\Sigma_p])$ with values in adjoint invariant functions on G by the series*

$$\mathrm{Tr}(f) = \sum_{\lambda \in \Omega^+} H(f c^\lambda) \, \bar{\chi}_\lambda. \tag{4.33}$$

1. *The series converges pointwise at every regular orbit and takes values in continuous functions on* $\mathrm{Conj}_{\mathrm{reg}}(G)$.

2. *The $\mathbb{C}$-valued functional $\mathrm{Tr}_\sigma(f)$ obtained by evaluation at a non-special conjugacy class σ can be obtained by integration along $\mathfrak{M}^G_\sigma$ with respect to a smooth measure μ_σ.*

3. *We have*

$$\mathrm{Tr}(f c_\lambda) = \mathrm{Tr}(f) \, \chi_\lambda.$$

Recall that for a nonspecial $x \in \mathfrak{t}$ we have a symplectic form ω_x on $\mathfrak{M}^G_{\sigma(\tilde{x})}$ which also produces a smooth volume form. The following equality is due to Witten [8, 27, 43]

$$c(g, G)\,\delta(\tilde{x})\,\mu_{\sigma(\tilde{x})} = \frac{\omega_x^{\dim \mathfrak{M}^G_{\sigma(\tilde{x})}}}{\dim \mathfrak{M}^G_{\sigma(\tilde{x})}!},$$

where $c(g, G)$ is defined by (1.4).

To demonstrate the power of the cutting calculus described above, we finish this section with the computation of Tr(1) for an arbitrary number of punctures. First note that the generalization of (4.33) reads as follows:

$$\mathrm{Tr}(f) = \sum_{\lambda\colon P \to \Omega^+} H\left(f \prod_{p \in P} c^p_{\lambda(p)}\right) \prod_{p \in P} \bar{\chi}_{\lambda(p)}.$$

Thus computing Tr(1) involves computing $H(\prod_{p \in P} c^p_{\lambda(p)})$. Choose an exact graph Γ and cut the graph under consideration by the edges of $\check{\Gamma}$. Then every edge e of $\check{\Gamma}$ will cut through two edges with opposite orientations. In order for them to give a nonzero contribution, they have to have the same coloring. Thus $H(\prod_{p \in P} c_{\lambda(p)}) = 0$, unless $\lambda(p) = \lambda(p') = \lambda$ for $p, p' \in P$. Then according to (4.28) the contribution at each cutting edge is a factor of $(\dim V_\lambda)^{-1}$, while the remaining graph consists of a union of loops colored by λ, one on each face of $\check{\Gamma}$. These faces correspond to the vertices of Γ and each contributes a factor of $\dim V_\lambda$ (cf. Example 1 (2)). Thus we obtain

$$H\left(\prod_{p \in P} c^p_\lambda\right) = (\dim V_\lambda)^{|V_\Gamma| - |E_\Gamma|} = (\dim V_\lambda)^{2-2g-|P|}.$$

Here we used that the Euler characteristic of Σ is $2-2g$ and that the faces of Γ are in one-to-one correspondence with the punctures. Thus we have

$$\mathrm{Tr}_\sigma(1) = \sum_{\lambda \in \Omega^+} \frac{\prod_{p \in P} \chi_\lambda(\sigma(p))}{(\dim V_\lambda)^{2g-2+|P|}}. \tag{4.34}$$

5. RIBBON GRAPHS AND THE MODULI ALGEBRA

In this section we construct a noncommutative q-deformation of the algebra $F(\mathfrak{M}^G[\Sigma_P])$, based on the representation theory of quantum groups. This algebra is equivalent to the "moduli algebra" constructed in [3] and [11] (cf. also [5, 9, 38]). Our construction is more geometric and transparent, however, and the calculations are much simpler. We clarify the relation of this algebra to the ribbon categories of Reshetikhin and Turaev [31], which simplifies the construction a great deal.

Since we will need to pass to concrete values of q, instead of the standard quantum group $\mathfrak{U}_q(\mathfrak{g})$, we will work over a smaller, "nonrestricted" algebra defined over the ring $D(q)$, which we introduced in Section 3 (cf. Appendix A for some details). We will use the generic symbol $\mathfrak{U}$ for this algebra.

We start with a technical prelude. In defining the quantum analog of the G-colored graphs, we need a more invariant version of the coloring of the vertices than the one used in [31]. The notion of a *cyclic invariant* that we introduce allows us to define an algebra of the correct size.

5.1. *The Reshetikhin–Turaev Map*

Recall the construction of ribbon categories of Reshetikhin and Turaev [31].

Define a *band* as a rectangle embedded in oriented 3-space, which has its sides marked as follows: the starting edge, the ending edge, the left edge, and the right edge. Alternatively, one can think of an arrow drawn on one side (the "marked side") of the rectangle parallel to one pair of edges. In particular, the band and its boundary segments are oriented. By attaching an edge e of a band to an oriented segment I, we mean that restricting the embedding of the band to e is an orientation-reversing embedding of e into I.

Let $\mathfrak{U}$ be an appropriate nonrestricted ribbon Hopf algebra [12, 31] (cf. Section A.1 and remark above). Fix L_s and L_e, two parallel oriented lines in $\mathbb{R}^3$. A *ribbon configuration* is a union of bands of two types, *ribbons* and *coupons*, which projects into the strip between the two lines and such that the lines, ribbons, and coupons are all disjoint except that the starting and ending edges of each ribbon are attached either to a line or to the starting or ending edge of a coupon.

Let $\mathbf{C}$ be a ribbon configuration which has an irreducible representation of $\mathfrak{U}$ associated to each of its ribbons. Then each coupon $\mathbf{c}$ of the configuration acquires a *type*, $(S(\mathbf{c}), E(\mathbf{c}))$, which is simply the list of the colorings of the ribbons at the starting and the ending edges of the coupon, with the direction of the arrows recorded as follows: we record a $+$ if the arrow points toward the coupon and a $-$ if it points away from the coupon. Thus $(S(\mathbf{c}), E(\mathbf{c}))$ has the form $([(V_1, \varepsilon_1), \ldots, (V_k, \varepsilon_k)], [(W_1, \rho_1), \ldots, (W_l, \rho_l)])$, where $\varepsilon_i = \pm 1$, $\rho_i = \pm 1$, and $V_i, W_j \in \mathrm{Rep}(\mathfrak{U})$. A type $(S(\mathbf{C}), E(\mathbf{C}))$ maybe similarly associated to the whole configuration $\mathbf{C}$.

A $\mathfrak{U}$-*ribbon configuration* is one where each ribbon is colored by a representation of $\mathfrak{U}$ and each coupon by a $\mathfrak{U}$-invariant map

$$V_s(\mathbf{c}) = V_1^{\varepsilon_1} \otimes \cdots \otimes V_k^{\varepsilon_k} \to W_1^{\rho_1} \otimes \cdots \otimes W_l^{\rho_l} = V_e(\mathbf{c}). \tag{5.35}$$

Here we used the following convention: if the coloring is V, then on the starting end $V^+ = V$ and $V^- = V^*$, where the latter is the left representation of $\mathfrak{U}$ defined on the dual space to V using the antipode; on the other end it is the other way around: $W^- = W$ and $W^+ = W^*$.

A fundamental result of [31] is that each $\mathfrak{U}$-ribbon configuration defines a morphism of $\mathfrak{U}$-representations $V_s(\mathbf{C}) \to V_e(\mathbf{C})$, which depends on the ribbon configuration up to homotopy only. Denote by $I_{(S,E)}$ the linear space of intertwining maps (5.35), which we will call invariants of type (S, E). Then the result may be summarized by saying that a ribbon configuration $\mathbf{C}$ with colored edges induces a well-defined product

$$\mathrm{RT}_{\mathbf{C}}\colon \bigotimes_{\mathbf{c}} I_{(S(\mathbf{c}),\, E(\mathbf{c}))} \to I_{(S(\mathbf{C}),\, E(\mathbf{C}))}. \tag{5.36}$$

We will call this product the *Reshetikhin–Turaev* map.

5.2. *Cyclic Invariants and Colored Ribbon Graphs*

The improvement that we suggest is the following: define a *cyclic type* $(T)^{\mathrm{cycl}}$ to be a cyclically ordered set of representations of $\mathfrak{U}$, each marked with a $+$ or a $-$. Clearly, each ordinary type induces a cyclic type by assigning a cyclic order to the V's and W's the obvious way, listing the representations counterclockwise, ignoring which ones are attached to the lower edge and which ones to the upper edge. If $(T)^{\mathrm{cycl}}$ has m representations, there will be $m(m+1)$ different types corresponding to $(T)^{\mathrm{cycl}}$. Indeed, one may reduce a cyclic ordering on m elements to an ordinary ordering in m different ways, and then one can divide an ordered set of m elements into two ordered sequences in $m+1$ ways. Denote the cyclic type derived from an ordinary type (S, E) by $(S, E)^{\mathrm{cycl}}$.

There are operations of bending an edge (or "attaching a candy cane"), which map spaces of invariants of the same cyclic type into each other. Given a type

$$(S, E) = ([(V_1, \varepsilon_1), \ldots, (V_k, \varepsilon_k)], [(W_1, \rho_1), \ldots, (W_l, \rho_l)]),$$

and an intertwiner $\phi \in I_{(S,E)}$, let

$$A_{er}\phi = (\mathrm{id} \otimes \mathrm{coinv}(W_l, \rho_l))(\phi \otimes \mathrm{id}_{W_l^{-\rho_l}}),$$

where $\mathrm{coinv}(W, \rho)\colon W^\rho \otimes W^{-\rho} \to \mathbb{C}$ is the standard coinvariant (cf. end of Section A.1). Then $A_{er}\phi$ is an intertwiner of type

$$([(V_1, \varepsilon_1), \ldots, (V_k, \varepsilon_k), (W_l, \rho_l)], [(W_1, \rho_1), \ldots, (W_{l-1}, \rho_{l-1})]).$$

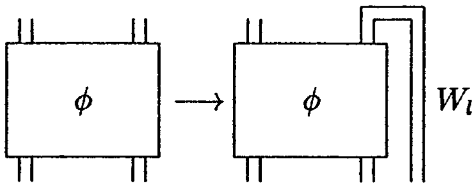

FIG. 4. Operation A_{er}.

A pictorial representation of this operation is shown on Fig. 4. We can define three other operations A_{el}, A_{sr}, and A_{sl} which bend the edges W_1, V_k, and V_1 respectively. Clearly, we have $A_{er}A_{sr} = \mathrm{id}$ and $A_{el}A_{sl} = \mathrm{id}$. Iterating these maps we get various maps between the resulting spaces of invariants.

LEMMA 6. *The spaces of invariants of the same fixed cyclic type*

$$\{I_{(S,E)} \mid (S,E)^{\mathrm{cycl}} = (T)^{\mathrm{cycl}}\}$$

are canonically isomorphic under these maps.

Proof. This is implicit in the original paper [31]. One needs to check a generalization of the relation (Rel_{13}) of [31]. An instance of such a relation is

$$\phi(\mu v_1, \mu v_2, ..., \mu v_l) = \phi(v_1, v_2, ..., v_l)$$

for an invariant $\phi \in \mathrm{Hom}(V_1 \otimes V_2 \otimes \cdots \otimes V_l, \mathbb{C})$. This follows from the fact that μ is a group-like element, and $\epsilon(\mu) = 1$. ▮

We will call invariants related by the above maps *cyclically equivalent*. The lemma allows us to define the space $I_{(T)^{\mathrm{cycl}}}$ of *cyclic* $\mathfrak{U}$*-invariants* of type $(T)^{\mathrm{cycl}}$ as equivalence classes of intertwiners of the same cyclic type related by the above maps.

EXAMPLE 2. The simplest example of a cyclic invariant is the trivial invariant of type $([(V,+)], [(V,-)])^{\mathrm{cycl}}$. This can be interpreted as the six different invariant maps, which correspond to the diagram (Fig. 1) of [31], or equivalently to the six maps of [12, p. 167] denoted $\iota_V^+, \iota_V^-, \alpha_V^+, \alpha_V^-, \beta_V^+, \beta_V^-$.

Remark 9. Since μ equals the identity when $q = 1$, cyclic equivalence of quantum invariants corresponds to ordinary cyclic permutations and contractions of the factors in the classical case.

Now we can define a noncommutative generalization of G-colored graphs.

DEFINITION 3. A $\mathfrak{U}$-*colored ribbon graph* $\hat{f}$ consists of

• a *ribbon graph* $R(\hat{f})$, which is a union of oriented discs (vertices) and ribbons embedded into $\mathbb{R}^3$. The discs and ribbons are disjoint except that each end of each ribbon is attached to the boundary of a disc, as always, respecting the orientations. We denote the set of ribbons of $R(\hat{f})$ by $E_{R(\hat{f})}$ and the set of vertices by $V_{R(\hat{f})}$.

• a coloring $C_{\hat{f}}: E_{R(\hat{f})} \to \mathrm{Rep}(\mathfrak{U})$ of the ribbons by representations of $\mathfrak{U}$. Clearly, such a coloring assigns a cyclic type $(T(v))^{\mathrm{cycl}}$ to each vertex $v \in V_{R(\hat{f})}$.

• A coloring of each vertex $v \in V_R$ by a cyclic invariant $\phi(v)$ of type $(T(v))^{\mathrm{cycl}}$.

The key point is that the Reshetikhin–Turaev map given in (5.36) is compatible with cyclic equivalence. This allows us to define a variant of the equivalence operation which we used in the Poisson case.

LEMMA 7. *Let $\hat{f}$ be a $\mathfrak{U}$-colored ribbon graph with m free edges embedded in the interior of a cylindrical surface CS (e.g., in $\{x^2+y^2<1\}$) in $\mathbb{R}^3$ in such a way that the free edges are attached to a fixed circle (e.g., $\{x^2+y^2=1, z=0\}$) on CS. A choice of type compatible with the cyclic type, induced by the colorings and the cyclic ordering of the m free edges, as well as a choice of a type for each vertex, give rise to a $\mathfrak{U}$-ribbon configuration. The RT map* (5.36) *induced by this configuration applied to the colorings of the vertices give rise to an invariant map between the appropriate tensor products of the colorings of the free edges. Then all the invariant maps thus obtained are cyclically equivalent.*

The proof of this lemma is left to the reader as an exercise. Note that this operation of replacing a ribbon graph by a single cyclic vertex is defined over $D(q)$ since the R-matrix, the central element v and μ are.

We can pair the free edges against a vertex and obtain the following:

COROLLARY 1. *The algorithm described in the lemma associates a well-defined number to every $\mathfrak{U}$-colored ribbon graph (without free edges).*

This is our version of the Reshetikhin–Turaev invariant. A special case is

COROLLARY 2. *The natural pairing between* $\mathrm{Hom}_{\mathfrak{U}}(\mathbb{C}, V_1 \otimes \cdots \otimes V_l)$ *and* $\mathrm{Hom}_{\mathfrak{U}}(V_1 \otimes \cdots \otimes V_l, \mathbb{C})$ *given by composition is cyclically invariant.*

One could also say that there is a canonical pairing

$$\langle\, , \rangle : I_{(T)^{\mathrm{cycl}}} \otimes I_{(T^*)^{\mathrm{cycl}}} \to \mathbb{C}$$

between cyclic invariants of type $(T)^{\mathrm{cycl}}$ and those of the dual type $(T^*)^{\mathrm{cycl}}$, which changes the cyclic orientation to the opposite and changes each representation to its dual.

5.3. *Quantization of Moduli Spaces*

The notions of cyclic invariants and colored ribbon graphs permit us to q-deform the constructions of the previous section. The constructions below are completely parallel to the Poisson case.

Let Σ be a compact Riemann surface embedded into 3-space, oriented outward, with a set of marked points P, as before. Denote by Σ^{ε} a small open neighborhood of Σ and by π a projection $\pi\colon \Sigma^{\varepsilon} \to \Sigma$; let $\Sigma_P^{\varepsilon} = \pi^{-1}(\Sigma \setminus P)$. Then define $F^q(\mathfrak{U}, \Sigma_P)$ to be the free $D(q)$-module, generated by the homotopy classes of embeddings of colored ribbon graphs into $\Sigma_P^{\varepsilon} \subset \mathbb{R}^3$. We can define a *product* of two colored ribbon graphs $\hat{f}, \hat{g} \in F^q(\mathfrak{U}, \Sigma_P)$ as the disjoint union of the two ribbon graphs $\hat{f}'$ and $\hat{g}'$, where $\hat{f}'$ is in the interior of Σ and is homotopic to $\hat{f}$, while $\hat{g}'$ is in the exterior of Σ and is homotopic to $\hat{g}$. Note that this is an associative but, generally, noncommutative product. An analogous operation was used earlier by Turaev in [38] (also cf. [5]).

Now we define the notion of *equivalence*. We fix a disc $D \subset \Sigma_P$ and a $\mathfrak{U}$-colored ribbon graph $\hat{f}$ in Σ_P^{ε} such that $R(\hat{f}) \cap \pi^{-1}(\partial D) \subset \Sigma$. This means that the ribbons intersecting $\pi^{-1}(\partial D)$ are effectively attached to ∂D. This is a version of the notion of *generic position* with respect to D. Then using Lemma 7, we can construct a new graph $\hat{f} \mid D \in F^q(\mathfrak{U}, \Sigma_P)$, replacing the part $\pi^{-1}D \cap R(\hat{f})$ by a single vertex, which can be arranged to lie entirely in Σ. The colorings of the edges and vertices outside D are inherited from $\hat{f}$, while the coloring of the new vertex is obtained Lemma 7. Now define the algebra $F^q(\mathfrak{M}^G[\Sigma_P])$ to be the quotient of $F^q(\mathfrak{U}, \Sigma_P)$ by the $D(q)$ linear subspace generated by this notion of equivalence; this subspace is clearly also an ideal. Thus $F^q(\mathfrak{M}^G[\Sigma_P])$ is endowed with a natural associative algebra structure over $D(q)$.

Again we have the basic examples of equivalence:

EXAMPLE 3. 1. We can erase an edge which is colored by the trivial representation.

2. We can divide every ribbon into two pieces joined by a vertex colored by the trivial cyclic invariant (cf. Example 2).

3. A small ring-like ribbon contractible in Σ_P^{ϵ} and colored by V_λ is equivalent to the number $q\dim V_\lambda$ introduced in Section 1.4.

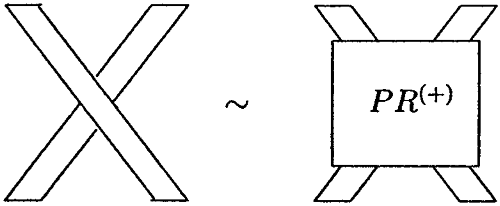

FIGURE 5

4. We can replace two overcrossing (resp. undercrossing) ribbons colored by V and W, by four ribbons attached to a coupon colored by $P_{VW}R^{(+)}_{VW}$ (resp. $P_{VW}R^{(-)}_{VW}$), where as usual $R^{(+)}$ denotes the R matrix and $R^{(-)}$ denotes R_{21}^{-1} (Fig. 5).

5. The operation $\dagger I$, as well as the operation of *cutting* $\dagger_0 I$, are defined as in the previous section. The pairing $\mathrm{Tr}_V(\phi(S)\,\phi(E))$ needs to be replaced by $\mathrm{Tr}_V(\mu\phi(S)\,\phi(E))$ as is natural in the theory of quantum groups (cf. Section A.1). Just as in the definition of equivalence, the condition of generic position is $\pi^{-1}(I) \cap R(\hat{f}\,) \in I$.

6. Assume that a ribbon $\hat{e}$ of a $\mathfrak{U}$-colored ribbon graph $\hat{f}$ is colored by a representation V_λ. Then we have $\hat{g} = v(\lambda)\,\hat{f}$, where $\hat{g}$ is $\hat{f}$ with the ribbon $\hat{e}$ twisted according to the orientation of 3-space by 360°, and $v(\lambda)$ is the value of the (central) ribbon element v in the representation V_λ (cf. [31], Section A.1). Twisting in the opposite direction induces multiplication by $v(\lambda)^{-1}$.

For an embedded graph $\Gamma \subset \Sigma_P$, there is a well-defined ribbon graph $R_\Gamma \subset \Sigma_P$ obtained by thickening Γ in Σ. In particular, one can associate canonical elements $\hat{c}^p_\lambda$ to the Poisson central elements c^p_λ defined in the previous section.

According to Remark 9, there is a well-defined classical limit of a cyclic invariant to a classical invariant, covering the evaluation at $q=1$ of the coefficient rings: ev: $D(q) \to \mathbb{C}$.

PROPOSITION 11. 1. *Define a map over* $\mathbb{C}$

$$\mathrm{red}_F : F^q(\mathfrak{M}^G[\Sigma_P]) \to F(\mathfrak{M}^G[\Sigma_P]),$$

by shrinking the width of the ribbons to 0, projecting them onto Σ_P, *and applying the above-mentioned reduction of the cyclic invariants to the classical ones. This map is an algebra homomorphism over* $\mathbb{C}$.

2. *The space of all colorings of an exact graph* Γ *span* $F^q(\mathfrak{M}^G[\Sigma_P])$.

3. *The elements* $\{\hat{c}^p_\lambda \mid p \in P, \lambda \in \Omega^+\}$ *are in the center of* $F^q(\mathfrak{M}^G[\Sigma_P])$. *They span an algebra isomorphic to* $R(G)^{\oplus |P|} \otimes D(q)$.

Proof. (1) This amounts to checking that the relations in the quantum case reduce to the classical ones when $q=1$. This is straightforward.

(2) The proof is similar to that of Proposition 7 (2). By iterating the $\dagger e$ operation with respect to the edges to the dual graph $\check{\Gamma}$, one can write any element of $F^q(\mathfrak{M}^G[\Sigma_P])$ as a sum of elements with underlying graph Γ.

(3) The statement is clear since one can move a small circle around the line $\pi^{-1}(p)$ for some puncture p, the graph underlying $\hat{c}^p_V$, past any other ribbon graph using a homotopy. Moreover, using the same argument as in the classical case, in $F^q(\mathfrak{M}^G[\Sigma_P])$ we have $\hat{c}^p_{V\otimes W} = \hat{c}^p_V \hat{c}^p_W = \hat{c}^p_W \hat{c}^p_V$. ∎

The definition of the analog of the operation H_Γ is a bit more subtle, because out of the three definitions we gave in the commutative case (projection, cutting, integration) only the cutting operation is clearly well defined.

PROPOSITION 12. 1. *The functional* H_Γ *defined by the cutting operation does not depend on* Γ. *Thus we have a well-defined functional* $H\colon F^q(\mathfrak{M}^G[\Sigma_P]) \to D(q)$. *Also, the functionals H in the classical and quantum cases are compatible with the reduction map* red_F; *i.e., we have* $H(\mathrm{red}_F(\hat{f})) = \mathrm{ev}(H(\hat{f}))$.

2. *The Poisson structure induced on* $F(\mathfrak{M}^G)$ *by the evaluation map* red_F *and the relation* $q = e^{\pi i \hbar}$ *coincides with the one defined in Section* 4.2.

3. *For* $\hat{f}, \hat{g} \in F^q(\mathfrak{M}^G[\Sigma_P])$, *we have* $H(\hat{f}\hat{g}) = H(\hat{g}\hat{f})$.

Proof. (1) The proof is the same as in the classical case.

(2) This can be derived from the results [20]. We will not give the details here.

(3) It is sufficient to prove the statement for the case when $R(\hat{f})$ is the thickening of an exact graph Γ and $R(\hat{g})$ is the same graph, but with edges oriented in the opposite way. Choose an ordering of the edges at each vertex of Γ compatible with the cyclic order. Every edge of the dual graph $\check{\Gamma}$ is crossed by two edges of $\hat{f}\hat{g}$ and $\hat{g}\hat{f}$ which have the same coloring but opposite orientation. To perform the cutting operation, we move these two edges side by side, so that we achieve the condition

$$\pi^{-1}(\check{\Gamma}) \cap R(\hat{f}\hat{g}) = \check{\Gamma} \cap R(\hat{f}\hat{g}),$$

and we do the same for $\hat{g}\hat{f}$. Note that this involves making a noncanonical choice, but we make the same choice in both cases. After performing the cutting with respect to the edges of $\check{\Gamma}$ we obtain that

$$H(\hat{f}\hat{g}) = \prod_{e \in E_\Gamma} (q\dim C(e))^{-1} \prod_{v \in V_\Gamma} \langle \phi_{\hat{f}}(v), \phi_{\hat{g}}(v) \rangle_1,$$

where $\langle\, ,\, \rangle_1$ is a certain pairing between the corresponding invariants which depends on the particular choice that we made at each edge of $\check{\Gamma}$. The formula for $H(\hat{g}\hat{f})$ is the same, but with a pairing $\langle\, ,\, \rangle_2$ replacing $\langle\, ,\, \rangle_1$. The difference between the two pairings is that in the first case the ribbon graph $R(\hat{g})$ is above $R(\hat{f})$ and in the second case it is below. The two cases are related by the homotopic move of rotating by 360° the piece of $R(\hat{g})$ remaining on a particular face after the cutting, so that it ends up under the corresponding piece of $R(\hat{f})$. Thus the difference between $H(\hat{f}\hat{g})$ and $H(\hat{g}\hat{f})$ will be a twist of $\pm 360°$ at every edge of $\check{\Gamma}$, wherever the two pieces of graphs are joined. At an edge of $\check{\Gamma}$, which we assume to be colored by V_λ, this twist contributes a factor of $v^{\pm}(\lambda)$ on one side and $v^{\mp}(\lambda)$ on the other, which cancel each other. This ends the proof. ∎

Remark 10. Parts (2) and (3) together imply Proposition 9 (5), but the proof of (3) given above has a simple semiclassical version giving a proof of the Poisson trace property of H. Instead of the 360° twists, in that case one encounters self-intersecting edges with the operator $P_{12} \circ t$ (cf. Section 4.2) inserted at the point of self-intersection.

6. THE TRACE

Now we are ready to quantize the fixed holonomy moduli spaces. Again, to avoid complicating our notation any further, we will assume that G is simply laced.

Fix a set of regular conjugacy classes $\boldsymbol{\sigma}: P \to \mathrm{Conj}_{\mathrm{reg}}(G)$ and define the quotient by the ideal

$$F^q(\mathfrak{M}^G_{\boldsymbol{\sigma}}) = F^q(\mathfrak{M}^G[\Sigma_P]) / \langle \{ \hat{c}^p_\lambda = \chi_\lambda(\boldsymbol{\sigma}(p)) \mid p \in P, \lambda \in \Omega^+ \} \rangle.$$

Define a functional Tr^q on $F^q(\mathfrak{M}^G[\Sigma_P])$ by the series

$$\mathrm{Tr}^q_{\boldsymbol{\sigma}}(\hat{f}) = \sum_{\lambda: P \to \Omega^+} H\left(\hat{f} \prod_{p \in P} \hat{c}^p_{\lambda(p)} \right) \prod_{p \in P} \bar{\chi}_{\lambda(p)}(\boldsymbol{\sigma}(p)). \tag{6.37}$$

Formally, the series takes values in functions in q and $|P|$ copies of (a completion of) $R(G)$.

Our main result is formulated in the theorem below. As we mentioned earlier, we are only proving this statement for $G = SU(2)$ in this paper, although several partial results are proved in the general case.

We will use the term "convergence" in the punctured unit disc to mean absolute and uniform convergence of holomorphic functions on each ringlike domain $\{\epsilon \leqslant |q| \leqslant 1-\epsilon\}$.

THEOREM 3. *Let $\hat{f} \in F^q(\mathfrak{M}^G[\Sigma_P])$, $G = SU(2)$, and $\boldsymbol{\sigma}$ as above.*

1. *Then the series* (6.37) *defining* $\mathrm{Tr}^q_{\boldsymbol{\sigma}}(\hat{f})$ *converges to a holomorphic function on the punctured unit disc.*

2. *For every $p \in P$, $\lambda \in \Omega^+$*

$$\mathrm{Tr}^q(\hat{f}\hat{c}^p_\lambda) = \mathrm{Tr}^q(\hat{f})\, \chi_\lambda;$$

thus the evaluation $\mathrm{Tr}^q_{\boldsymbol{\sigma}}$ *descends to the quotient* $F^q(\mathfrak{M}^G_{\boldsymbol{\sigma}})$.

3. *If $\boldsymbol{\sigma}$ is not special (cf. Section* 1.5*), then* $\mathrm{Tr}^q_{\boldsymbol{\sigma}}(\hat{f})$ *has an asymptotic expansion at $q = 1$. More precisely, there is a function* $\mathrm{Tr}^\hbar(\hat{f})$*, analytic in a neighborhood of 0, and positive constants τ, C such that*

$$|\mathrm{Tr}^q(\hat{f}) - \mathrm{Tr}^\hbar(\hat{f})| < Ce^{-\frac{\tau}{|\hbar|}},$$

where $q = e^{\pi i \hbar}$ and $\hbar \in i\mathbb{R}^+$ is sufficiently small. Moreover, $\mathrm{Tr}^\hbar(\hat{f})$ *is a rational function in q and $\hbar$.*

Remark 11. The notation $\mathrm{Tr}^\hbar(\hat{f})$ is somewhat inconsistent. We will sometimes use $\mathrm{aexp}(\mathrm{Tr}^q(\hat{f}))$ instead.

Remark 12. A statement analogous to Part (1) of the theorem was independently proved in [10].

The proofs of parts (1) and (3) are given after Proposition 15.

Proof of part (2). Consider the diagonal element $\sum_\lambda \chi_\lambda \otimes \bar{\chi}_\lambda$ with respect to the standard quadratic form on $R(G)$: $(\alpha, \beta) = H_G(\alpha\beta)$, where H_G is the projection onto the trivial character. This form is manifestly invariant, $(\alpha\gamma, \beta) = (\alpha, \beta\gamma)$; thus the diagonal element has a similar property:

$$\sum_\lambda \chi_\lambda \chi_\mu \otimes \bar{\chi}_\lambda = \sum_\lambda \chi_\lambda \otimes \chi_\mu \bar{\chi}_\lambda.$$

This implies the statement since the algebra spanned by $\{\hat{c}_\lambda\}$ is simply another copy of the algebra $R(G)$. ∎

6.1. *The One Puncture Case*

Again, first we consider the $|P|=1$ case. Let p be a marked point on a surface Σ of genus g, and fix a conjugacy class σ of the group G.

Our goal is to study the infinite series

$$\mathrm{Tr}_{\sigma}^{q}(\hat{f}) = \sum_{\lambda \in P^+} H(\hat{f}\hat{c}_{\lambda}^{p})\, \bar{\chi}_{\lambda}(\sigma).$$

We first express $H(\hat{f}\hat{c}_{\lambda}^{p})$ in terms of intertwiners of irreducible finite dimensional representations of $\mathfrak{U}$.

Let Γ be the usual exact graph with vertex o, $2g$ edges, $\{e_i\}_{i=1}^{2g}$, and a single face containing the puncture p. Then the ribbon graph $R(\hat{c}_{\lambda}^{p})$ is the thickening of a small circle around p, which is homotopic to the product $e_1 e_2 e_1^{-1} e_2^{-1} \cdots e_{2g-1} e_{2g} e_{2g-1}^{-1} e_{2g}^{-1}$ taken in $\pi_1(\Sigma_p)$.

By Proposition 11 (2), any colored ribbon graph $\hat{f} \in F^q(\mathfrak{M}^G[\Sigma_P])$ is equivalent to one with underlying ribbon graph R_Γ, the thickening of the graph Γ; thus we can assume that, in fact, $R(\hat{f}) = R_\Gamma$. Then $\hat{f}$ is given by the colorings of the edges: $C_{\hat{f}}(e_{2i-1}) = V_{\mu(i)}$, $C_{\hat{f}}(e_{2i}) = V_{\nu(i)}$, $i=1, \ldots, g$, and a coinvariant $\phi_{\hat{f}} \in \mathrm{Hom}_{\mathfrak{U}}(\mathscr{V}, D(q))$, where we used the notation $\mathscr{V} = \bigotimes_{i=g}^{1} \mathscr{V}_i$, with $\mathscr{V}_i = V_{\nu(i)}^{*} \otimes V_{\mu(i)}^{*} \otimes V_{\nu(i)} \otimes V_{\mu(i)}$.

The dual graph $\check{\Gamma}$ is isomorphic to Γ, has its vertex at p, and each of its $2g$ edges intersect exactly one edge of Γ. Denote the points of intersection by $\{a_i\}_{i=1}^{2g}$, correspondingly. We can represent $H(\hat{f}\hat{c}_{\lambda}^{p})$ by performing the cutting operation with respect to the edges of $\check{\Gamma}$. The resulting colored ribbon graph will have the form of a cartwheel lying entirely above the face of $\check{\Gamma}$. If we represent the insertion of the diagonal element by a dotted line, or *chord*, between the two relevant vertices of $H(\hat{f}\hat{c}_{\lambda}^{p})$, then we obtain the "cartwheel with a snow chain" diagram with $4g+1$ vertices $\{o, a_i^{\pm}\}$ and with a_i^{+} and a_i^{-} joined by a chord. We have drawn a schematic picture of the resulting graph for the case $g=2$ on Fig. 6.

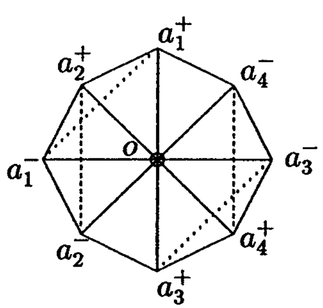

FIG. 6. Genus 2 case.

Next, for $\lambda, \mu \in P^+$ and a finite dimensional representation V, introduce the notation

$$I(V; \lambda, \mu) = \mathrm{Hom}(V_\lambda, V_\mu \otimes V) \qquad \text{and} \qquad I^*(V; \lambda, \mu) = \mathrm{Hom}(V_\mu, V_\lambda \otimes V^*);$$

denote by $I(V; \lambda)$ the space $I(V; \lambda, \lambda)$. This notation is appropriate since the spaces of intertwiners I and I^* are naturally dual to each other. Indeed, let $\phi \in I(V; \lambda, \mu)$ and $\psi \in I^*(V, \lambda, \mu)$. If we use the shorthand $\psi \circ \phi$ for the composition $(\psi \otimes \mathrm{id})\, \phi$ and $\langle\ \rangle_V$ for the natural invariant pairing on $V^* \otimes V$, then the expression $\langle \psi \circ \phi \rangle_V$ can be considered to be a constant, since it represents an intertwiner from the irreducible representation V_λ to itself. The resulting pairing is nondegenerate and gives rise to the diagonal element $\delta(V; \lambda, \mu) \in I^*(V; \lambda, \mu) \otimes I(V; \lambda, \mu)$ defined by $\delta(V; \lambda, \mu) = \sum \alpha^i \otimes \alpha_i$ where $\{\alpha_i\}$ and $\{\alpha^i\}$ are dual bases with respect to the pairing; i.e., $\langle \alpha_i \circ \alpha^j \rangle_V = \delta_i^j$.

If V, W are two finite dimensional $\mathfrak{U}$-modules, let

$$\xi_{V,W}(\lambda) = \sum_{j,k} \alpha^j \circ \beta^k \circ \alpha_j \circ \beta_k \in \mathrm{Hom}(V_\lambda, V_\lambda \otimes V^* \otimes W^* \otimes V \otimes W),$$

where $\sum_j \alpha^j \otimes \alpha_j = \delta(V; \lambda)$ and $\sum_k \alpha^k \otimes \alpha_k = \delta(W; \lambda)$. As usual, we will write $\xi_{\lambda\mu}$ instead of $\xi_{V_\lambda V_\mu}$.

Our new notation allows us to rewrite the formula for the trace we are trying to compute as follows.

Lemma 8. *We have*

$$H(\hat{f} \hat{c}_\lambda^p) = \frac{1}{(q\dim V_\lambda)^{2g-1}} (\mathrm{id}_{V_\lambda} \otimes \phi_{\hat{f}})\, \xi_{\nu(g)\,\mu(g)}(\lambda) \circ \cdots \circ \xi_{\nu(1)\,\mu(1)}(\lambda), \tag{6.38}$$

where again $\circ$ *stands for the composition of intertwiners as above, each acting on* V_λ *only.*

Proof. After cutting the cartwheel diagram between a_{2g}^- and a_1^+ we obtain the expression on the RHS exactly, with each chord contributing a factor $(q\dim V_\lambda)^{-1}$. When we make this cut, we lose a factor of $q\dim V_\lambda$, hence the exponent $1-2g$ in the expression. ∎

As a particular case of this expression, we can take $\hat{f} = 1$ and obtain

$$H(\widehat{c_\lambda^p}) = \frac{1}{(q\dim V_\lambda)^{2g-1}} \tag{6.39}$$

for the $|P|=1$ case. This formula can be found in [3, 11]. In general, this leads to the quantum version of (4.34):

$$\mathrm{Tr}_{\boldsymbol{\sigma}}^{q}(1)=\sum_{\lambda}\frac{\prod_{p\in P}\chi_{\lambda}(\boldsymbol{\sigma}(p))}{(q\dim V_{\lambda})^{2g-2+|P|}}. \tag{6.40}$$

For $G=SU(2)$ this is exactly the series $\widetilde{T}(\mathbf{x};q)$ introduced in Section 1.4, where $\mathbf{x}$ and $\boldsymbol{\sigma}$ are related by the exponential map as usual.

Next, we would like to study the behavior of (6.38) in λ. In order to track the λ-dependence of such an expression, one would like to identify the spaces $I(V;\lambda)$ as λ varies. This can be done using the notions of "expectation value" and "fusion matrices" introduced in [15, 16]. As we need a refinement of these objects, we give a small introduction to the subject in Section A.3, while only providing the basic definitions below.

Our plan is to identify the spaces $I(V;\lambda,\mu)$ and the corresponding intertwiner spaces for the Verma modules $\tilde{I}(V;\lambda,\mu)$ with $V[\lambda-\mu]$ via the *expectation value map* $\langle.\rangle$ (cf. Definition 4). This would identify, for example, all of the spaces $I(V;\lambda)$ with the same vector space $V[0]$, which would facilitate a universal treatment. To this end we need the expectation value map to be an isomorphism, and this turns out to be true most of the time. Indeed, if μ is sufficiently far from the walls of the Weyl chamber relative to v, i.e., μ is generic with respect to v (cf. Definition 5), then the expectation value map is an isomorphism between $I(V_v;\lambda,\mu)$ and $V_v[\lambda-\mu]$. In addition, in the Verma module case, the map $\langle.\rangle$ has a right inverse $v\mapsto\phi_{\lambda}^{v}$.

For any sufficiently generic λ, this allows one to introduce the *fusion matrices* $J_{VW}(\lambda)\in\mathrm{End}(V\otimes W)$, defined by (A.54), which represent the operation of composition of the intertwiners under the identification $\langle.\rangle$. This object has a natural generalization $J_{UVW\ldots}$ to tensor products with more than two factors. The operator $Q_W(\lambda)\in\mathrm{End}(W)$ is defined by contracting the vectors in the image of $J_{W^*W}(\lambda)$ (cf. (A.56)). Finally, the *dynamical Weyl group* elements $A_{V,w}(\lambda)\colon V[v]\to V[w(v)]$, defined for $\lambda\in\Omega^+$, $w\in W_G$, represent the standard inclusion of Verma modules $M_{w.\lambda}\hookrightarrow M_{\lambda}$ under the same identification.

All objects defined in the previous paragraph have a universal form: $J(\lambda)=J_{12}(\lambda)$, $J_{1,2,\ldots,n}$, $Q(\lambda)$, $A_w(\lambda)$ in an appropriate completion of tensor products of $\mathfrak{U}$. The following lemma is a simple consequence of the notation we introduced.

Lemma 9. *Let V be a finite dimensional representation, $\lambda,\mu\in\Omega^+$, and μ sufficiently generic. Then under the identification of $I(V;\lambda,\mu)$ with $V[\lambda-\mu]$*

and $I^*(V;\lambda,\mu)$ with $V[\lambda-\mu]^*$, the diagonal element $\delta(V;\lambda,\mu)$ corresponds to the element

$$\sum_i v^i \otimes Q_V(\lambda)^{-1}\, v_i \in V[\lambda-\mu]^* \otimes V[\lambda-\mu],$$

where $\sum_i v^i \otimes v_i = \delta(V[\lambda-\mu])$.

The lemma allows us to write down a formula for $H(\hat{f}\hat{c}^p_\lambda)$ in terms of fusion matrices only.

PROPOSITION 13. *For notational convenience, denote by $W(i)$, where $i = 1 \cdots 4g$, the spaces*

$$V^*_{\nu(g)}[0], V^*_{\mu(g)}[0], V_{\nu(g)}[0], V_{\mu(g)}[0], \ldots, V^*_{\nu(1)}[0], V^*_{\mu(1)}[0], V_{\nu(1)}[0], V_{\mu(1)}[0]$$

in that order. If $\lambda \in \Omega^+$ is sufficiently generic, then we have

$$H(\hat{f}\hat{c}^p_\lambda) = \frac{1}{(q\dim V_\lambda)^{2g-1}}\, \phi_{\hat{f}} J_{1,\ldots,4g}(\lambda) \prod_{i=1}^{g} Q^{-1}_{W(4i-1)}(\lambda)\, Q^{-1}_{W(4i)}(\lambda)\, \omega, \qquad (6.41)$$

where $\omega = \bigotimes_{i=1}^{g} \delta(W(4i-1))_{4i-3,\,4i-1} \otimes \delta(W(4i))_{4i-2,\,4i}$ is the diagonal element of $\bigotimes_{i=1}^{4g} W(i)$.

The key point is that while the LHS of (6.41) is defined for dominant integral weights λ, the RHS is meaningful for arbitrary generic λ. This allows us to interpret the RHS in representation theoretical terms for arbitrary λ.

PROPOSITION 14. *Let $\hat{f} \in F^q(\mathfrak{M}^G)$. Then there exists a function $\tilde{R}_{\hat{f}}$, which is rational of nonpositive degree in the variables q_α, $\alpha \in \Delta^+$ with coefficients in $D(q)$ and with poles along the "hyperplanes" $q_\alpha - q^m$, such that for sufficiently generic $\lambda \in \Omega^+$ one has*

$$H(\hat{f}\hat{c}^p_\lambda) = \frac{1}{(q\dim V_\lambda)^{2g-1}}\, \tilde{R}_{\hat{f}}\left(q_\alpha \mapsto \frac{q^{(\alpha,\lambda)}-1}{q-1}, \alpha \in \Delta^+\right), \qquad (6.42)$$

where $\mapsto$ means substitution.

Proof. We only give a proof in the case of $G = SU(2)$. This statement could be derived from results of [16], but in this paper, since we are mostly working with $SU(2)$, we chose to give explicit formulas, from which the statement is manifest. We are hoping that this will give a better idea of the complexity of the computations to the reader. No such formulas are known in the case of other groups.

For $G = SU(2)$, the operators J and Q are given in a concrete basis of each irreducible representation in (A.61) and (A.62). Comparing these with (6.41) and the formula (A.55) for $J_{1,\ldots,N}$, the statement of the proposition follows immediately. ▮

Denote the function obtained by substitution in the RHS of (6.42) by $R_{\hat{f}}(\lambda)$. Define the shifted Weyl group action on $\mathfrak{t}^*$ by $w.\lambda = w(\lambda+\rho)-\rho$ for all $w \in W_G$. We will now show that the function $R_{\hat{f}}(\lambda)$ is invariant under the shifted Weyl group action. We will use the properties of the operators $A_{w,V}$ defined above and in Section A.3.

PROPOSITION 15. *The function $R_{\hat{f}}(\lambda)$ is invariant under the shifted action of the Weyl group: $R_{\hat{f}}(\lambda) = R_{\hat{f}}(w.\lambda)$ for every $w \in W_G$.*

Proof. From the relation (A.58), we obtain

$$\begin{aligned} &\Delta^{(4g)}(A_w(\lambda))\, J_{1,\ldots,4g}(\lambda) \\ &\quad = J_{1,\ldots,4g}(w.\lambda)\, A_w^{(4g)}(\lambda)\, A_w^{(4g-1)}(\lambda - h^{(4g)}) \cdots A_w^{(1)}(\lambda - h^{(2)} - \cdots h^{(4g)}), \end{aligned} \tag{6.43}$$

where $\Delta^{(n)}$ is the iterated coproduct and $h^{(m)}$ means that the element h acts on the mth component of the tensor product. Since $\phi_{\hat{f}}$ is an intertwiner, we have $\phi_{\hat{f}} \Delta^{(4g)}(A_w(\lambda)) = \epsilon(A_w(\lambda))\, \phi_{\hat{f}} = \phi_{\hat{f}}$. Then we obtain:

$$R_{\hat{f}}(\lambda) = \phi_{\hat{f}} J_{1,\ldots,4g}(w.\lambda) \prod_{j=4g}^{1} A_{w,W(j)}(\lambda) \prod_{i=1}^{g} Q^{-1}_{W(4i-1)}(\lambda)\, Q^{-1}_{W(4i)}(\lambda)\, \omega. \tag{6.44}$$

As a result it is sufficient to show the following relation:

$$A_w^{(4)}(\lambda)\, A_w^{(3)}(\lambda)\, A_w^{(2)}(\lambda)\, A_w^{(1)}(\lambda)\, Q_2(\lambda)^{-1}\, Q_1(\lambda)^{-1}\, \omega = Q_2^{-1}(w.\lambda)\, Q_1^{-1}(w.\lambda)\, \omega.$$

This is equivalent to showing that

$$Q(w.\lambda)^{-1}|_{V[0]} = A_w(\lambda)\, Q^{-1}(\lambda)\, S(A_w(\lambda))|_{V[0]} \tag{6.45}$$

on the zero weight subspace. This follows from the relation (A.59). ▮

Remark 13. Note that in the case of $G = SU(2)$, the operator $J_{VW}(\lambda)$ restricted to $V[0] \otimes W[0]$ has simples poles of the type $(1-q^{2(\lambda-k)})^{-1}$, $k \in \mathbb{N}$. From (A.63), $Q_V(\lambda)^{-1}(\lambda)$ restricted to $V[0]$ has simple poles of the type: $(1-q^{2(-\lambda-2-k)})^{-1}$, $k \in \mathbb{N}$. This is a consistency check with the Weyl invariance of $R_{\hat{f}}(\lambda)$ and (6.41).

We have shown that far from the walls of the dominant chamber $H(\hat{f}\hat{c}^p_\lambda)$ has a certain form; this will be used to study $\mathrm{Tr}^q(\hat{f})$. The expression for

$\mathrm{Tr}^q(\hat{f})$, however, contains a sum over all dominant integral weights, and therefore we have to analyze the behavior of $H(\hat{f}\hat{c}^p_\lambda)$ on the hyperplanes of nongeneric weights as well. One needs to refine the results of [16] in order to be able to treat these cases, and we will not do this here.

Nevertheless, we can continue the study in the $G = SU(2)$ case where this problem can be circumvented. Indeed, as it was pointed out in Remark 15 (3), in this case, there are only finitely many nongeneric weights for every representation V.

Proof of parts (1) *and* (3) *of Theorem* 3. Denote by v the fundamental weight of $SU(2)$. We now return to the notation of Section 1: assume that a number $0 < x < 1/2$ represents the conjugacy class σ, so that $\exp(x) = \tilde{x} \in \sigma$. We can conclude from Propositions 14 and 15 that

$$\delta(\tilde{x})\,\mathrm{Tr}^q_\sigma(\hat{f}) - \frac{1}{2}\sum_{n\in\mathbb{Z}}{}^{*}\,\frac{e_{nv}(x)}{(q^n - q^{-n})^{2g-1}}\,R_{\hat{f}}(nv) \in D(q)\otimes R(T),$$

where $\sum{}^{*}$ means only summing finite values.

Now we can apply the arguments of Section 1.4 to the sum in the above formula. Indeed, since $\tilde{R}_{\hat{f}}$ is of nonpositive degree in q_v, the exponential convergence of the series is guaranteed. The residue calculations also go through: we apply the residue theorem to the form

$$w_{\hat{f}}(u, x; \hbar) = \frac{e^{\{x\}\,u}\,du}{1 - e^u}\,\frac{\tilde{R}_{\hat{f}}(q \mapsto e^{i\pi\hbar},\, q_v \mapsto e^{\hbar u/2})}{(e^{\hbar u/2} - e^{-\hbar u/2})^{2g-1}}. \tag{6.46}$$

Again, the poles break into three parts:

1. $\mathrm{Pole}_1 = \{\underline{m} \mid m \in \mathbb{Z},\, (q^m - q^{-m})^{1-2g}\, R_{\hat{f}}(mv) \neq \infty\}$;
2. $\mathrm{Pole}_2 \subset \{\underline{n}/\hbar + \underline{m} \mid n \in \mathbb{Z},\, n \neq 0\ |m| < M\}$ for some M;
3. $\mathrm{Pole}_3 = \{\underline{m} \mid m \in \mathbb{Z},\, (q^m - q^{-m})^{1-2g}\, R_{\hat{f}}(mv) = \infty\}$.

The set of poles Pole_1 contributes the infinite sum, Pole_2 contributes exponentially small corrections, and Pole_3 contributes the asymptotic expansion. The statements in part (3) then clearly follow. ∎

6.2. *The Case of Several Punctures*

We will use the notational conventions from Sections 1 and 2. Thus $P \subset \Sigma$ is a set of punctures and we fix a set of conjugacy classes $\boldsymbol{\sigma}\colon P \to \mathrm{Conj}_{\mathrm{reg}}(G)$ and corresponding elements $\mathbf{x}\colon P \to \mathfrak{t}$. We set $|P| = k$,

$P=\{p_1, \ldots, p_k\}$ and use the notation $\sigma_i=\boldsymbol{\sigma}(p_i)$, etc. Our goal is the study of the series (6.37):

$$\operatorname{Tr}_q(\hat{f})=\sum_{\lambda_1, \ldots, \lambda_k \in \Omega^+} H\left(\hat{f} \prod_{i=1}^k \hat{c}_{\lambda_i}^{p_i}\right) \prod_{i=1}^k \bar{\chi}_{\lambda_i}(\sigma_i). \tag{6.47}$$

The analysis is similar to the one-puncture case; thus we only highlight the differences here. The difficulty is that this sum formally has k parameters instead of 1. Our plan is to reduce the computations to the $k=1$ case. Again, we first state the results which can be obtained for any G and then analyze the case $G=SU(2)$ in more detail.

We first express $H(\hat{f} \prod_{i=1}^k \hat{c}_{\lambda_i}^{p_i})$ in terms of intertwiners of irreducible finite dimensional representations of $\mathfrak{U}$. Let Γ be the exact graph with vertex o and with $2g+k-1$ edges: $\{e_i, i=1, \ldots, 2g, m_j, j=1, \ldots, k-1\}$. The ribbon graph $\Gamma(\hat{c}_{\lambda_i}^{p_i})$ is homotopic to m_i in $\pi_1(\Sigma_P)$, whereas the ribbon graph $\Gamma(\hat{c}_\lambda^{p_k})$ is homotopic to $e_1 e_2 e_1^{-1} e_2^{-1} \cdots e_{2g-1} e_{2g} e_{2g-1}^{-1} e_{2g}^{-1} m_1 \cdots m_{k-1}$. Again, we can compute $H(\hat{f} \prod_{i=1}^k \hat{c}_{\lambda_i}^{p_i})$ using the cutting operation.

Let $C_{\hat{f}}(e_{2i-1})=V_{\mu(i)}$, $C_{\hat{f}}(e_{2i})=V_{\nu(i)}$ and let $C_{\hat{f}}(m_j)=V_{\zeta(j)}$, $j=1, \ldots, k-1$. The element $\hat{f}$ determines an invariant $\phi_{\hat{f}} \in \operatorname{Hom}(\mathscr{V}, D(q))$, where $\mathscr{V}=\bigotimes_{j=k-1}^1 \mathscr{V}'_j \otimes \bigotimes_{i=g}^1 \mathscr{V}_i$, with $\mathscr{V}_i=V^*_{\nu(i)} \otimes V^*_{\mu(i)} \otimes V_{\nu(i)} \otimes V_{\mu(i)}$, and $\mathscr{V}'_i=V^*_{\zeta(i)} \otimes V_{\zeta(i)}$.

Let V be a finite dimensional $\mathfrak{U}$-module and let $\eta_V(\lambda, \mu)$ be the elements of $\operatorname{Hom}(V_\lambda, V_\lambda \otimes V^* \otimes V)$ defined by $\eta_V(\lambda, \mu)=\sum_j \alpha^j \circ \alpha_j$ where $\sum_j \alpha^j \otimes \alpha_j = \delta(V; \lambda, \mu)$.

We can therefore rewrite the formula for $H(\hat{f} \prod_{i=1}^k \hat{c}_{\lambda_i}^{p_i})$ as

$$\begin{aligned}
&(q\dim V_\lambda)^{2g-1} \prod_{i=1}^{k-1} (q\dim V_{\lambda_i})^1 \, H\Big(\hat{f} \prod_{i=1}^k \hat{c}_{\lambda_i}^{p_i}\Big) \\
&\quad =(\mathrm{id}_\lambda \otimes \phi_{\hat{f}})\, \eta_{\zeta(k-1)}(\lambda, \lambda_{k-1}) \circ \cdots \circ \eta_{\zeta(1)}(\lambda, \lambda_1) \circ \xi_{\nu(g)\,\mu(g)}(\lambda) \circ \cdots \\
&\qquad \circ \xi_{\nu(1)\,\mu(1)}(\lambda),
\end{aligned}$$

where we have denoted $\lambda=\lambda_k$.

The analog of Proposition 13 holds. Indeed if λ is sufficiently far from the walls of the dominant chamber, we have

$$\eta_V(\lambda, \lambda-\mu)=\sum_i (\phi_{\lambda-\mu}^{v^i} \otimes \mathrm{id})\, \phi_\lambda^{Q(\lambda)^{-1} v_i},$$

where $\sum v^i \otimes v_i=\delta(V[\mu])$. Denote by $W(i)$, $i=1, \ldots, 4g+2k-2$, the appropriate weight spaces

$$V^*_{\zeta(k-1)}[\mu_{k-1}], V_{\zeta(k-1)}[\mu_{k-1}], \ldots, V^*_{\zeta(1)}[\mu_1], V_{\zeta(1)}[\mu_1],$$
$$V^*_{\nu(g)}[0], V^*_{\mu(g)}[0], V_{\nu(g)}[0], V_{\mu(g)}[0], \ldots, V^*_{\nu(1)}[0], V^*_{\mu(1)}[0], V_{\nu(1)}[0], V_{\mu(1)}[0].$$

If λ is sufficiently far from the walls of the Weyl chamber of dominant weights, we have

$$(q\dim V_\lambda)^{2g-1} \prod_{i=1}^{k-1} q\dim V_{\lambda_i} H\left(\hat{f} \prod_{i=1}^{k-1} \hat{c}_{\lambda-\mu_i}^{p_i} \hat{c}_\lambda^{p_k}\right)$$
$$= \phi_{\hat{f}} J_{1,\ldots,4g+2k-2}(\lambda) \prod_{j=1}^{k-1} Q_{W(2j)}^{-1}(\lambda)$$
$$\times \prod_{i=1}^{g} Q_{W(4i-1+2k-2)}^{-1}(\lambda)\, Q_{W(4i+2k-2)}^{-1}(\lambda)\, \omega_{g,k}, \tag{6.48}$$

where $\omega_{g,k}$ is the diagonal element of $\bigotimes_{i=1}^{4g+2k-2} W(i)$.

Again, there are appropriate rational functions $\tilde{R}_{\hat{f}}$ and $R_{\hat{f}}$, depending on the additional parameters $\mu_1, \ldots, \mu_{k-1}$, which recover (6.48) for sufficiently generic λ. The proof of Proposition 15 carries over to the multipuncture case:

PROPOSITION 16. *The functions $R_{\hat{f}}(\lambda; \mu_1, \ldots, \mu_{k-1})$ satisfy*

$$R_{\hat{f}}(\lambda; \mu_1, \ldots, \mu_{k-1}) = R_{\hat{f}}(w.\lambda; w(\mu_1), \ldots, w(\mu_{k-1})) \tag{6.49}$$

for any $w \in W_G$.

Assume that $G = SU(2)$ in the rest of the section; ν is the fundamental weight as before. Denote the set of weights of $V_{\zeta(i)}$ by $Z_i \nu$, where Z_i is a finite subset of $\mathbb{Z}$ invariant under $n \to -n$. The key point is that if we fix $\lambda = n\nu$, then only finitely many possible sets of μ's can make a nonzero contribution to (6.47). More concretely, we need to show that

$$\sum_{n_j \in Z_j, n \geq N} \frac{R_{\hat{f}}(n\nu; n_1\nu, \ldots, n_{k-1}\nu)}{[n+1]_q^{2g-1} \prod_{j=1}^k [n+1-n_j]_q} \prod_{j=1}^k \frac{\sin(2\pi(n+1-n_j))\, x_j}{\sin(2\pi x_j)} \tag{6.50}$$

for N large enough admits an asymptotic expansion of the form stated in Theorem 3. This series can also be written:

$$\sum_{\epsilon_j = \pm 1, n_j \in Z_j, n \geq N} \frac{R_{\hat{f}}((n-1)\,\nu; n_1\nu, \ldots, n_{k-1}\nu)}{[n]_q^{2g-1} \prod_{j=1}^k [n-n_j]_q} \prod_{j=1}^k \frac{\epsilon_j e^{\epsilon_j(n-n_j)\,\underline{x_j}}}{2i\sin(2\pi x_j)}. \tag{6.51}$$

Using the symmetry property of $R_{\hat{f}}((n-1)\,\nu; n_1\nu, \ldots, n_{k-1}\nu)$, the series has an asymptotic expansion, if all the series

$$\sum_{n \in \mathbb{Z}}^{*} \frac{R_{\hat{f}}((n-1)\,\nu; n_1\nu, \cdots, n_{k-1}\nu)}{[n]_q^{2g-1} \prod_{j=1}^k [n-n_j]_q} \prod_{j=1}^k \frac{e^{\sum_{j=1}^k \epsilon_j(n-n_j)\underline{x_j}}}{2i\sin(2\pi x_j)} \tag{6.52}$$

have asymptotic expansions. These series are exactly of the type already studied in the one puncture case and they satisfy the asymptotic property when $\sum_j \epsilon_j x_j \notin \mathbb{Z}$. This last condition is exactly the condition of being "nonspecial." ∎

7. CONCLUSION

The purpose of this section is to organize our program, review what we have done so far and to formulate the conjecture which served as the original motivation for starting this work. For simplicity, we consider the one-puncture case only. We state the cornerstones of our program in the following theorem-conjecture. We describe what has been done in this paper below.

THEOREM 4. *Let G be a simple compact connected and simply connected Lie group and Σ be a compact genus g Riemann surface with a puncture at a point p. Fix a regular nonspecial conjugacy class σ of G.*

1. *Then there is a canonical local q-deformation $A^q_\sigma = F^q(\mathfrak{M}_\sigma[\Sigma_p])$ defined over $D(q)$ with a structure map* red: $A^q_\sigma \to F(\mathfrak{M}_\sigma[\Sigma_p])$ *covering the evaluation map $D(q) \to \mathbb{C}$.*

2. *There is a cyclic functional*

$$\mathrm{Tr}^q_\sigma \colon A^q_\sigma \to \mathrm{Mer}^{\mathrm{a.e.}}_1(q)$$

with values in meromorphic functions in the variable q on the unit disc which have an asymptotic expansion in the variable $\hbar$ as $\hbar \to i0^+$, where $q = e^{\pi i \hbar}$.

3. *Define* $\mathrm{Tr}^\hbar_\sigma = \mathrm{aexp} \circ \mathrm{Tr}^q_\sigma$, *where* aexp: $\mathrm{Mer}^{\mathrm{a.e.}}_1 \to \mathbb{C}[\![\hbar]\!]$ *is the asymptotic expansion. Then the trace functional $s^* \mathrm{Tr}^\hbar_\sigma$ on $F(\mathfrak{M}^G_\sigma)[\![\hbar]\!]$, which is the pull-back of $\mathrm{Tr}^\hbar_\sigma$ via a local section, extends to all smooth functions $C^\infty(\mathfrak{M}^G_\sigma)[\![\hbar]\!]$.*

4. *The equality*

$$c(g, G, q)\, \delta(\tilde{x})\, \mathrm{Tr}^\hbar_\sigma(1)) = \hbar^{\mathrm{rank}\, G(g-1)} \int_{\mathfrak{M}^G_{\sigma(\tilde{x})}} e^{\omega_x/\hbar} \hat{A}(\mathfrak{M}^G_{\sigma(\tilde{x})}) \tag{7.53}$$

holds, where ω_x is the standard symplectic form (cf. Section 2.2) and $c(g, G, q)$ is the normalization constant defined in Section 1.4.

Results of the present paper.

1. We give a precise and efficient construction of the quantum moduli algebra A^q_σ. The proof that this algebra is a local q-deformation will be given in a followup publication.

2. We construct a formal series in the general case, but the statement is proved for $G = SU(2)$ only. While the convergence of the series is not hard to show for a general G, the proof of the existence of the asymptotic expansion requires a significant improvement of the analysis in Section 6.

3. This is a technical statement that is true for an arbitrary q-deformation. The proof will be given elsewhere.

4. We show this equality only for $G = SU(2)$ in the present paper.

Assuming for now all of the statements listed above, we may finally formulate our main conjecture:

CONJECTURE 1. *Let $x \in \mathfrak{t}$ be nonspecial. Then the characteristic class of the deformation $A^q_\sigma(\tilde{x})$ is $\omega_x/\hbar$, i.e., A^q_σ is basic, and up to an appropriate power of $\hbar$ the functional* $\mathrm{Tr}^{\hbar}_{\sigma}$ *is the canonical trace.*

We expect this conjecture to hold in complete generality, for all groups and arbitrary number of punctures. We do not know how to approach this conjecture at the moment. Clearly, the results of [3] are relevant, but it is not clear how to make the connection rigorous. What we have shown in the case of $G = SU(2)$ in this paper is that the two statements in the conjecture are consistent:

If A^q_σ is basic then the asymptotic trace is canonical.

Finally, note that for groups other than $SU(n)$, the moduli space of flat connections with fixed holonomies is an orbifold even for generic inserted conjugacy classes. Our results could shed some light on the correct form of the orbifold version of the index theorem of Fedosov and Nest and Tsygan.

APPENDIX A: BACKGROUND ON QUANTUM GROUPS

In this Appendix we collected facts about quantum groups of which we make use in the main text. We recall the basic definitions in Section A.1 and write down the explicit formulas in the $\mathfrak{sl}_2$ case in Section A.2. The monographs [12, 25] are the basic references for this part. In Section A.3 we recall the definitions of dynamical quantum groups and dynamical Weyl groups defined in [16–18]. We also present some explicit computations in the $\mathfrak{sl}_2$ case. These are based on a brief remark in [37], but are not available in this form in the literature.

A.1. *Quantum Universal Algebra*

For an indeterminate q, define the following elements of $\mathbb{Z}[q, q^{-1}]$: $[m]_q = (q^m - q^{-m})/(q - q^{-1})$ for $m \in \mathbb{Z}$, $[n]_q! = [n]_q \cdots [1]_q$, for $n \in \mathbb{N}$, and

$$\begin{bmatrix} n \\ m \end{bmatrix}_q = \frac{[n]_q!}{[m]_q!\,[n-m]_q!}, \qquad 0 \leqslant m \leqslant n.$$

Let $\mathfrak{g}$ be a finite dimensional complex simple Lie algebra of rank r with Cartan matrix (a_{ij}), and let d_i be the coprime positive integers such that the matrix $d_i a_{ij}$ is symmetric. Introduce the notation $q_i = q^{d_i}$.

$\mathfrak{U}_q(\mathfrak{g})$ is the $\mathbb{C}(q)$ Hopf algebra generated by K_i, K_i^{-1}, e_i, f_i, $i = 1, \ldots, r$ satisfying the defining relations:

$$K_i K_i^{-1} = K_i^{-1} K_i = 1, \qquad K_i K_j = K_j K_i,$$

$$K_i e_j K_i^{-1} = q_i^{a_{ij}} e_j, \qquad K_i f_j K_i^{-1} = q_i^{-a_{ij}} f_j,$$

$$e_i f_j - f_j e_i = \delta_{ij} \frac{K_i - K_i^{-1}}{q_i - q_i^{-1}},$$

$$\sum_{r=0}^{1-a_{ij}} \begin{bmatrix} 1-a_{ij} \\ r \end{bmatrix}_{q_i} (-1)^r e_i^{1-a_{ij}-r} e_j e_i^r = 0,$$

$$\sum_{r=0}^{1-a_{ij}} \begin{bmatrix} 1-a_{ij} \\ r \end{bmatrix}_{q_i} (-1)^r f_i^{1-a_{ij}-r} f_j f_i^r = 0.$$

The coproduct is defined by:

$$\Delta(e_i) = e_i \otimes 1 + K_i^{-1} \otimes e_i, \qquad \Delta(f_i) = f_i \otimes K_i + 1 \otimes f_i, \qquad \Delta(K_i) = K_i \otimes K_i.$$

We denote the counit by ϵ and the antipode by S. The sum of positive roots may be expressed as $2\rho = \sum_{i=1}^r m_i \alpha_i$ with $m_i \in \mathbb{N}$. Define $K_{2\rho} = \prod_{i=1}^r K_i^{m_i}$. Then for every $a \in \mathfrak{U}_q(\mathfrak{g})$ we have $S^2(a) = K_{2\rho} a K_{2\rho}^{-1}$.

As the above relations are defined over the ring $\mathscr{A} = \mathbb{Z}[q, q^{-1}]$, by adding to the list of generators the divided powers $(K_i - K_i^{-1})/(q_i - q_i^{-1})$, one can define the "nonrestricted" integral form of $\mathfrak{U}_q(\mathfrak{g})$, an $\mathscr{A}$-subalgebra $\mathfrak{U}_{\mathscr{A}}(\mathfrak{g})$ of $\mathfrak{U}_q(\mathfrak{g})$ such that the natural map $\mathfrak{U}_{\mathscr{A}}(\mathfrak{g}) \otimes_{\mathscr{A}} \mathbb{C}(q) \to \mathfrak{U}_q(\mathfrak{g})$ is an isomorphism of $\mathbb{C}(q)$-algebras. Then one can specialize q to a nonzero complex number $q_0 \in \mathbb{C}$ by $U_{q_0}(\mathfrak{g}) = \mathfrak{U}_{\mathscr{A}}(\mathfrak{g}) \otimes_{\mathscr{A}} \mathbb{C}$, using the homomorphism $\mathrm{ev}_{q_0} : \mathscr{A} \to \mathbb{C}$, which sends q to q_0. In particular, one can set $q = 1$ and obtain that $U_1(\mathfrak{g})$ is essentially isomorphic to $U(\mathfrak{g})$. (One needs to set $K_i = 1$ as well; for the details see [12, Section 9.2]). When we speak of a representation V of the quantum group, we assume that the action of the operators K_i is diagonalizable and has eigenvalues of the form q^n, $n \in \mathbb{C}$. This ensures a "good limit" as $q \to 1$, i.e., an appropriate action of $U(\mathfrak{g})$ on $V/(q-1)V$.

In the text, we enlarged the ring $\mathscr{A}$ to the ring $D(q)$, which consists of those rational functions in q which have no poles at $q=1$ or inside the unit disc, except possibly at 0. To simplify the notation, we denoted this algebra by $\mathfrak{U} = \mathfrak{U}_{\mathscr{A}}(\mathfrak{g}) \otimes_{\mathscr{A}} D(q)$.

For every complex weight $\lambda \in \mathfrak{h}^*$, define the Verma module M_λ as the universal $\mathfrak{U}$-module generated by a vector v_λ and relations

$$K_i v_\lambda = q^{(\lambda, \alpha_i)} v_\lambda, \qquad e_i v_\lambda = 0, \qquad i = 1, ..., r.$$

Remark 14. Note that we defined (., .) to be the basic inner product on $\mathfrak{t}^*$, normalized by the condition, that the long roots have square length 2 (cf. Section 1.1). The formulas above then work for simply laced Lie algebras only, and we assume in what follows that $\mathfrak{g}$ is such a Lie algebra. In the non-simply laced case one needs to normalize the inner product in such a way that the short roots have square length 2.

If $\lambda \in \Omega^+$ is a dominant integral weight, then M_λ has a unique finite dimensional quotient V_λ which is irreducible.

Define the q-dimension of V_λ by $q\dim V_\lambda = \mathrm{Tr}_{V_\lambda}(K_{2\rho}) \in \mathbb{N}[q, q^{-1}]$. From the classical Weyl formula we have $q\dim V_\lambda = \prod_{\alpha \in \Delta^+} [(\lambda+\rho, \alpha)]_q / [(\rho, \alpha)]_q$.

The Hopf algebra $\mathfrak{U}_q(\mathfrak{g})$ has some additional special properties:

- *quasitriangularity*: there is an operator R in a completion of $\mathfrak{U}_q(\mathfrak{g})^{\otimes 2}$ such that for every $a \in \mathfrak{U}_q(\mathfrak{g})$, $\Delta^{op}(a) = R\Delta(a) R^{-1}$, and $(\Delta \otimes \mathrm{id})(R) = R_{13} R_{23}$, $(\mathrm{id} \otimes \Delta)(R) = R_{13} R_{12}$.
- *ribbon algebra*: it can be completed with the element $u = \sum_i S(b_i)\, a_i$, and with a central element v (the ribbon element), defined by $v^2 = uS(u)$, and $\Delta(v) = (R_{21} R_{12})^{-1} (v \otimes v)$. The action of v on irreducible representations V_λ is the constant $v(\lambda) = q^{-C(\lambda)}$ where $C(\lambda)$ is the value of the classical quadratic Casimir in the classical representation associated to the dominant weight λ. We use the symbol μ for the grouplike element $uv^{-1} = K_{2\rho}$. It plays an important role in the "attaching the candy cane" transformation in Section 5.2. While the natural pairing $V^* \otimes V \to \mathbb{C}(q)$ is invariant, the natural pairing $V \otimes V^* \to \mathbb{C}(q)$ is not. It needs to be twisted with μ^{-1}: $v, w^* \mapsto w^*(\mu^{-1} v)$.

A.2. *The Example* $\mathfrak{U}_q(sl(2))$

In this paragraph we write down some explicit formulas for the objects defined above in the simplest case of $\mathfrak{g} = \mathfrak{sl}_2$.

The R-matrix belongs to a completion of $\mathfrak{U}_q(\mathfrak{sl}_2)^{\otimes 2}$ and is given by $R = R^0 q^{h \otimes h/2}$, where $R^0 = \exp_{q^{-2}}((q-q^{-1})\, e \otimes f)$ and $K = q^h$. The q-exponential is defined by $\exp_\alpha(z) = \sum_{n=0}^{+\infty} (1-\alpha)^n\, z^n/(\alpha; \alpha)_n$, where $(a; b)_n = \prod_{k=0}^{n-1} (1 - ab^k)$.

The irreducible representations V_m of $\mathfrak{U}_q(\mathfrak{sl}_2)$ are classified by a positive integer m. We can choose a basis v_k^m, $0 \leqslant k \leqslant m$ of V_m, on which the action of the generators of $\mathfrak{U}_q(\mathfrak{sl}_2)$ is

$$Kv_k^m = q^{(m-2k)} v_k^m, \qquad ev_k^m = [m-k+1]_q\, v_{k-1}^m, \qquad fv_k^m = [k+1]_q\, v_{k+1}^m.$$

A basis of the Verma module M_λ is denoted u_k^λ, $k \in \mathbb{N}$, on which the action of the generators of $\mathfrak{U}_q(\mathfrak{sl}_2)$ is

$$Ku_k^\lambda = q^{(\lambda-2k)} u_k^m, \qquad eu_k^\lambda = [k]_q\, [\lambda-k+1]_q\, u_{k-1}^\lambda, \qquad fu_k^\lambda = u_{k+1}^\lambda.$$

The action of the ribbon element v on V_m is well defined and is simply multiplication by the constant $q^{-m(m+2)/2}$.

One can define a quantum version of the Weyl reflection as follows. Define an algebra automorphism T of $\mathfrak{U}_q(\mathfrak{sl}_2)$ by

$$T(e) = -qf, \qquad T(f) = -q^{-1}e, \qquad T(q^h) = q^{-h}.$$

The operator T intertwines Δ, the standard coalgebra structure of $\mathfrak{U}_q(\mathfrak{sl}_2)$, with the opposite coalgebra structure Δ^{op}; i.e., $\Delta^{\mathrm{op}} T = (T \otimes T)\, \Delta$. Then one can define an element w such that $T(a) = waw^{-1}$ and $\Delta(w) = R^{-1}(w \otimes w)$. Its action on V_m can easily be computed: $wv_k^m = (-1)^k\, q^{-m^2/4-k} v_{m-k}^m$. We have $w^2 = (-1)^m$ id on V_m.

A.3. *Dynamical Quantum Groups*

Here we give a short survey of the formalism of fusion matrices introduced in [16, 17] and dynamical quantum Weyl groups from [18, 37] in a form necessary for our applications. The original papers mainly deal with intertwiners between Verma modules instead of irreducible finite dimensional modules. We have had lots of help from Pavel Etingof here.

Notation. We presume the notation of Section 1.1 and 6. For any $\mathfrak{h}$-diagonalizable representation V of the algebra $\mathfrak{U}$ and weight $\mu \in \mathfrak{h}^*$, denote by $V[\mu]$ the μ-weight space in V. Thus we have $V = \bigoplus_\mu V[\mu]$. Note that for a highest weight representation U_λ with highest weight λ, we have $U_\lambda[\mu] = 0$ unless $\mu \leqslant \lambda$ with respect to the standard partial order (cf. Section 1.1). Also, we will denote the weight of a vector v of pure weight in a representation V by $\mathrm{wt}(v)$.

DEFINITION 4. Let V be a finite dimensional representation of $\mathfrak{U}$ and U_λ, U_μ be highest weight representations generated by the vectors u_λ and u_μ, correspondingly. The *expectation value* $\langle\phi\rangle$ of an intertwiner ϕ from $\mathrm{Hom}(U_\lambda, U_\mu \otimes V)$ is defined by the equation $\phi(u_\lambda) = u_\mu \otimes \langle\phi\rangle + \sum_i u_i \otimes v_i$, where $\mathrm{wt}(u_i) < \mu$. This yields a map

$$\langle . \rangle \colon \mathrm{Hom}(U_\lambda, U_\mu \otimes V) \to V[\lambda - \mu].$$

We will be interested in two cases: when $U_\lambda = V_\lambda$ the irreducible highest weight representation for a dominant integral weight λ, and when $U_\lambda = M_\lambda$ the Verma module for arbitrary $\lambda \in \mathfrak{h}^*$. We denote the intertwiners in the second case by

$$\tilde{I}(V; \lambda, \mu) = \mathrm{Hom}(M_\lambda, M_\mu \otimes V) \quad \text{and} \quad \tilde{I}^*(V; \lambda, \mu) = \mathrm{Hom}(M_\mu, M_\lambda \otimes V^*).$$

DEFINITION 5. We will say that $\mu \in \mathfrak{t}^*$ is *not* generic with respect to the dominant weight $\nu \in \Omega^+$ if for some positive root $\alpha \in \Delta^+$ we have[3]

$$0 \leqslant (\alpha, \mu + \rho)\, \alpha \leqslant \nu.$$

LEMMA 10. *Let $\lambda, \mu \in \Omega^+$.*

1. *The natural map $\pi \colon \tilde{I}(V, \lambda, \mu) \to \mathrm{Hom}(M_\lambda, V_\mu \otimes V)$ factors through a map $\tilde{\pi} \colon \tilde{I}(V, \lambda, \mu) \to I(V, \lambda, \mu)$, which is compatible with the expectation value maps.*
2. *The expectation value map $\langle . \rangle \colon I(V; \lambda, \mu) \to V[\lambda - \mu]$ is injective.*
3. *Suppose that $V = V_\nu$ with $\nu \in \Omega^+$ and that μ is generic with respect to ν. Then the expectation value map $\langle . \rangle \colon \tilde{I}(V_\nu; \lambda, \mu) \to V_\nu[\lambda - \mu]$ has a canonical (right) inverse $v \mapsto \phi^v_\lambda$. In particular, the map $\langle\, , \rangle$ is surjective.*

The lemma implies the following important statement:

PROPOSITION 17. *If μ is generic with respect to ν, then the map*

$$\langle . \rangle \colon I(V_\nu; \lambda, \mu) \to V_\nu[\lambda - \mu]$$

is an isomorphism of vector spaces.

Remark 15. 1. One can extend the notion of genericity of μ with respect to V_ν to an arbitrary finite dimensional representation V by additivity and thus conclude that $\langle\, , \rangle \colon I(V; \lambda, \mu) \to V[\lambda - \mu]$ is an isomorphism.

[3] We remind the reader that we are assuming G is simply laced.

2. We will also denote by ϕ_λ^v the map $\tilde{\pi}\phi_\lambda^v \in I(V, \lambda, \lambda - \mathrm{wt}(v))$ if this causes no confusion.

3. Note that in the case of $\mathfrak{g} = \mathfrak{sl}_2$, there are only finitely many weights μ nongeneric with respect to a particular representation V.

Proof of the lemma. (1) Let $\phi \in \tilde{I}(V, \lambda, \mu)$ be a nonzero intertwiner. Then the image $\pi(\phi(M_\lambda))$ is a finite dimensional module of highest weight λ, which necessarily has to be isomorphic to V_λ. Clearly, then π factors through $\tilde{\pi}$.

(2) Let $\phi \in I(V, \lambda, \mu)$ and represent the image of the highest weight vector as $\phi(v_\lambda) = \sum_i x_i \otimes y_i$, where we assume that the x_is are vectors of pure weight in V_μ. Split the sum as

$$\sum_i x_i \otimes y_i = \sum_j x_j^{\max} \otimes y_i + \sum_l x_l \otimes y_l,$$

where the vectors $x_j^{\max}$ are the vectors of maximal weight among those which occur in the original sum. It follows from the intertwiner property that any such vector $x_j^{\max}$ has to be a singular vector, i.e., be killed by all of the e_i's. Then the statement follows since V_λ has only one such vector, the highest weight vector.

(3) This statement is a slight generalization (λ, μ arbitrary) of Etingof and Styrkas [15, Proposition 2.1.]. ▌

The proposition allows us to introduce the basic objects of [16]: fusion matrices and the dynamical Weyl group operators. Below we will always assume that the necessary genericity conditions hold. In particular, we have an intertwiner $\phi_\lambda^v \in \tilde{I}(V, \lambda, \lambda - \mathrm{wt}(v))$ such that $\langle \phi \rangle = v$. If V, W are finite dimensional modules, define an endomorphism of $V \otimes W$, called *fusion matrix* and denoted $J_{VW}(\lambda)$, by the equation

$$\langle \phi^v_{\lambda - \mathrm{wt}(w)} \circ \phi^w_\lambda \rangle = J_{VW}(\lambda)(v \otimes w), \tag{A.54}$$

where v, w are pure weight vectors. In fact, there exists a universal element $J(\lambda)$ in a completion of $\mathfrak{U}^{\otimes 2}$ such that $J(\lambda)$ is represented by $J_{VW}(\lambda)$ on the module $V \otimes W$. This operator satisfies the so-called "dynamical cocycle equation":

$$(\mathrm{id} \otimes \Delta)\, J(\lambda)\, J_{23}(\lambda) = (\Delta \otimes \mathrm{id})\, J(\lambda)\, J_{12}(\lambda - h^{(3)}),$$

where the notation $h^{(3)}$ stands for the action of h on the 3rd tensor component.

One can generalize the definition of J_{VW} to tensor products with more components. These higher fusion matrices are also induced by universal elements, which may be written as

$$J_{1,2,\ldots,N}(\lambda) = J_{1,2\ldots N}(\lambda) \cdots J_{N-1,N}(\lambda), \tag{A.55}$$

where $J_{1,2\cdots N}(\lambda) = (id \otimes \Delta^{(N-1)})\, J(\lambda)$ and $\Delta^{(p)}\colon \mathfrak{U} \to \mathfrak{U}^{\otimes p}$ is the iterated coproduct.

If $v \in W[\lambda-\mu]^*$, $w \in W[\lambda-\mu]$, the linear map $\mathrm{Tr}_W(\phi^v_\lambda \otimes \mathrm{id}_W)\, \phi^w_\lambda$ is an intertwiner from V_λ to itself. Such an intertwiner is necessarily the identity times a constant, and we will use the same notation for this constant as for the linear map. This defines a nondegenerate pairing between $W[\lambda-\mu]^*$ and $W[\lambda-\mu]$, and we can define an invertible endomorphism $Q_W(\lambda)$ of $W[\lambda-\mu]$, such that

$$\mathrm{Tr}_W(\phi^v_\lambda \otimes \mathrm{id}_W)\, \phi^w_\lambda = \langle v, Q_W(\lambda)\, w\rangle; \tag{A.56}$$

i.e., $\langle v, Q_W(\lambda)\, w\rangle = \mathrm{Tr}_W(J_{W^*W}(\lambda)(v \otimes w))$. We can also define the universal element $Q(\lambda)$ in a completion of $\mathfrak{U}$ by $Q(\lambda) = \sum_i S(a_i)\, b_i$, where $J(\lambda) = \sum_i a_i \otimes b_i$. It is easy to check that this element is represented on W by $Q_W(\lambda)$.

Finally, we turn to the definition of the dynamical quantum Weyl group introduced in [17, 37]. For V finite dimensional, sufficiently generic $\lambda \in \Omega^+$ and an element $w \in W_G$, there is a canonical inclusion $M_{w.\lambda} \hookrightarrow M_\lambda$, which induces an isomorphism between $\tilde{I}(V; w.\lambda) \to \tilde{I}(V; \lambda)$. The expectation value map identifies these spaces with $V[v]$ and $V[w(v)]$, correspondingly, thus this isomorphism can be represented by an operator $A_{V,w}(\lambda)$: $V[v] \to V[w(v)]$. Again, this operator is induced by a universal element $A_w(\lambda)$ in a completion of $\mathfrak{U}$, such that on each finite dimensional $\mathfrak{U}_q(\mathfrak{g})$ module V, $A_w(\lambda)$ is represented by an endomorphism $A_{V,w}(\lambda)$. The operators $A_w(\lambda)$ satisfy the following two relations:

$$A_{ww'}(\lambda) = A_w(w'.\lambda)\, A_{w'}(\lambda),$$

$$\forall w, w' \in W_G,\ l(ww') = l(w) + l(w') \tag{A.57}$$

$$\Delta(A_w(\lambda))\, J(\lambda) = J(w.\lambda)\, A_w^{(2)}(\lambda)\, A_w^{(1)}(\lambda - h^{(2)}). \tag{A.58}$$

In particular, applying $(S \otimes id)$ followed by the algebra multiplication one obtains the relation between $Q(\lambda)$ and $Q(w.\lambda)$:

$$Q(\lambda) = S(A_w(\lambda - wh))\, Q(w.\lambda)\, A_w(\lambda). \tag{A.59}$$

A.4. *Explicit Computation in the* $\mathfrak{sl}_2$ *Case*

Here we present some explicit computations of the objects defined in the previous paragraph in the $\mathfrak{U}_q(\mathfrak{sl}_2)$ case. To simplify our notation, we will identify the weight $\lambda = lv$ with the integer l and the weight $\mathrm{wt}(v)$ with the symbol h.

Let $J(\lambda)$ be the fusion matrix of $\mathfrak{U}_q(\mathfrak{sl}_2)$. It can easily be computed using the ABRR linear equation [2], which reads

$$J(\lambda)(1 \otimes q^{(2\lambda+1)\,h-h^2/2}) = R^0_{21}(1 \otimes q^{(2\lambda+1)\,h-h^2/2})\,J(\lambda). \tag{A.60}$$

The computation results in the formula

$$J(\lambda) = \sum_{n=0}^{+\infty} \frac{q^n(1-q^{-2})^{2n}}{(q^{-2};q^{-2})_n}\,(f^n \otimes e^n j_n(\lambda)), \tag{A.61}$$

where $j_n(\lambda) = \prod_{k=0}^{n-1} (1-q^{2(\lambda-h-k)})^{-1}$.

Recall that $Q(\lambda) = \sum_i S(a_i)\,b_i$, where $J(\lambda) = \sum_i a_i \otimes b_i$. After a straightforward computation, we obtain from (A.61) that

$$Q(\lambda)\,v^m_k = {}_2\varphi_1(q^{-2(m-k+1)}, q^{2k}; q^{2(\lambda-m+2k)})(q^{-2})\,v^m_k,$$

where ${}_2\varphi_1$ is the basic hypergeometric function of base q^{-2} evaluated at q^{-2}. Using the Heine formula [22] we obtain

$$Q(\lambda)\,v^m_k = q^{-2(m-k+1)\,k}\,\frac{(q^{2(\lambda+k+1)};q^{-2})_k}{(q^{2(\lambda-m+2k)};q^{-2})_k}\,v^m_k. \tag{A.62}$$

From this explicit expression, an easy computation implies the relation:

$$Q(-\lambda-2)\,Q(\lambda+h) = q^{\frac{h^2}{2}+h}v. \tag{A.63}$$

If λ is a positive integer, then $\omega_\lambda = (f^{\lambda+1}/[\lambda+1]_q!)\,u^\lambda_0$ is a singular vector in M_λ and $\phi^{v^m_k}_\lambda(\omega_\lambda)$ is a singular vector in $M_{\lambda-(m-2k)} \otimes V_m$. The endomorphism $A_{V_m}(\lambda)$ is defined by $\phi^{v^m_k}_\lambda(\omega_\lambda) = \omega_{\lambda-(m-2k)} \otimes A_{V_m}(\lambda)(v^m_k)$ + terms of lower weight. It can be shown that $A_{V_m}(\lambda)$ is the value of the element $A(\lambda)$ acting on V_m, where

$$A(\lambda) = v^{-1}Kq^{-h^2/4}wQ(\lambda). \tag{A.64}$$

In order to show this equality, one can follow the proof of [18] or proceed along the lines of [37]: we can first compute $\phi^{v^m_k}_\lambda(u^\lambda_0)$ and then compute $\phi^{v^m_k}_\lambda(\omega^\lambda)$ by applying $f^{\lambda+1}/[\lambda+1]_q!$. This leads to

$$A_{V_m}(\lambda)(v^m_k) = \frac{[k+n]_q!}{[k]_q!}\,{}_3\varphi_2(\alpha_1, \alpha_2, \alpha_3; \beta_1, \beta_2)(q^{-2})\,v^m_{m-k}, \tag{A.65}$$

where $\alpha_1 = q^{2(\lambda-n+1)}$, $\alpha_2 = q^{2(-m+k-1)}$, $\alpha_3 = q^{2k}$, $\beta_1 = q^{-2n-2}$, $\beta_2 = q^{2(\lambda-n)}$, and $n = m-2k$.

By using the q-analog of Saalschütz formula [22], we arrive at

$$A_{V_m}(\lambda)(v_k^m) = (-1)^k \, q^{-k(m-k+1)} \frac{(q^{2(\lambda+k+1)}; q^{-2})_k}{(q^{2(\lambda-m+2k)}; q^{-2})_k} \, v_{m-k}^m \tag{A.66}$$

from which Eq. (A.64) follows.

ACKNOWLEDGMENTS

We are greatly indebted to Pavel Etingof for his help at all stages of this project and to Boris Tsygan for his insights and suggestions. Useful discussions with Jean-Michel Bismut, Victor Guillemin, and Michèle Vergne are also gratefully acknowledged.

REFERENCES

1. M. Atiyah and R. Bott, Yang–Mills equations over Riemann surfaces, *Philos. Trans. Roy. Soc. London Ser. A* **308** (1983), no. 1505, 523–615.
2. D. Arnaudon, E. Buffenoir, E. Ragoucy, and Ph. Roche, Universal solutions of quantum Yang-Baxter equations, *Lett. Math. Phys.* **44** (1998), 201–214.
3. A. Y. Alekseev, H. Grosse, and V. Schomerus, Combinatorial Quantization of the Hamiltonian Chern–Simons Theory I, *Comm. Math. Phys.* **172** (1995), 317–358. A. Y. Alekseev, H. Grosse, and V. Schomerus, Combinatorial quantization of the Hamiltonian Chern–Simons theory II, *Comm. Math. Phys.* **174** (1995), 561-604.
4. J. Andersen, J. Mattes, and N. Reshetikhin, The Poisson structure on the moduli space of flat connections and chord diagrams, *Topology* **35** (1996), 1069–1083.
5. J. Andersen, J. Mattes, and N. Reshetikhin, Quantization of the algebra of chord diagrams, *Math. Proc. Cambridge Philos. Soc.* **124** (1998), 451–467.
6. A. Beauville, Vector bundles on curves and generalized theta functions: recent results and open problems, *in* "Math. Sci. Res. Inst. Publ.," Vol. 28, pp. 17–33, Cambridge Univ. Press, Cambridge, UK, 1995. (Preprint: alg-geom/9404001)
7. H. Boden and Yi Hu, Variations of moduli of parabolic bundles, *Math. Ann.* **301** (1995), 539–559.
8. J.-M. Bismut and F. Labourie, Symplectic geometry and the Verlinde formulas, preprint, Université de Paris-Sud, 1998.
9. E. Bullock, C. Frohman, and J. Kania-Bartoszyńska, Topological interpretations of lattice gauge field theory, *Comm. Math. Phys.* **198** (1998), 47–81.
10. E. Bullock, C. Frohman, and J. Kania-Bartoszyńska, The Yang–Mills measure in the Kauffman bracket skein module, preprint, math.GT/0005218.
11. E. Buffenoir and Ph. Roche, Two dimensional lattice gauge theory based on a quantum group, *Comm. Math. Phys.* **170** (1995), 669–698. E. Buffenoir and Ph. Roche, Link invariants and combinatorial quantization of Hamiltonian Chern Simons theory, *Comm. Math. Phys.* **181** (1996), 331–365.
12. V. Chari and A. Pressley, "Quantum Groups," Cambridge Univ. Press, Cambridge, UK, 1994.

13. P. Deligne, Déformations de l'algèbre des fonctions d'une variété symplectique: comparaison entre Fedosov et De Wilde, Lecomte, *Selecta Math. (N.S.)* **1** (1995), 667–697.
14. M. De Wilde and P. Lecomte, Existence of star-Products and of formal deformations of the Poisson Lie algebra of arbitrary symplectic manifolds, *Lett. Math. Phys.* **7** (1983), 487–496.
15. P. Etingof and K. Styrkas, Algebraic integrability of Macdonald operators and representation theory of quantum groups, *Comp. Math.* **114** (1998), 125–152; preprint, math.QA/9603022
16. P. Etingof and A. Varchenko, Exchange dynamical quantum groups, *Comm. Math. Phys.* **205** (1999), 19–52; preprint, math.QA/9801135.
17. P. Etingof and A. Varchenko, Traces of intertwining operators for quantum groups and difference equations I, preprint, math.QA/9907181.
18. P. Etingof and A. Varchenko, Dynamical Weyl groups and applications, math.QA/0011001.
19. B. Fedosov, "Deformation Quantization and Index Theory," Mathematical Topics, Vol. 9, Akademie Verlag, Berlin, 1996.
20. V. V. Fock and A. A. Rosly, Poisson structures on moduli space of flat connections on Riemann surfaces and the r-matrices, *in* "Moscow Seminar in Math. Physics;" *Amer. Math. Soc. Transl. Ser. 2* **191**(1999), 67–86; q-alg/9802054
21. V. Guillemin, Star products on prequantizable symplectic manifolds, *Lett. Math. Phys.* **35** (1995), 85–89.
22. G. Gasper and R. Rahman, "Basic Hypergeometric Series," Encyclopedia of Mathematics and Its Applications, Vol. 35, Cambridge Univ. Press, Cambridge, UK, 1990.
23. L. C. Jeffrey and F. C. Kirwan, Intersection theory on moduli spaces of holomorphic bundles of arbitrary rank on a Riemann surface, *Ann. Math.* **148** (1998), 109–196.
24. A. V. Karabegov and M. Schlichenmaier, Identification of Berezin–Toeplitz deformation quantization, preprint, math.QA/00006063.
25. C. Kassel, "Quantum Groups," Springer-Verlag, Berlin/New York, 1995.
26. M. Kontsevich, Deformation quantization of Poisson manifolds, I, preprint, math.QA/9709040.
27. K. Liu, Heat kernel and moduli spaces I.-II., *Math. Res. Lett.* **3** (1996), 743–762; *Math. Res. Lett.* **4** (1997), 569-588.
28. E. Meinrenken and C. Woodward, Hamiltonian loop group actions and Verlinde factorization, *J. Differential Geom.* **50** (1998), 417–469.
29. R. Nest and B. Tsygan, Algebraic index theorem, *Comm. Math. Phys.* **172** (1995), 223–262.
30. M. Polyak and N. Reshetikhin, On 2D Yang–Mills theory and invariants of links, *in* "Deformation Theory and Symplectic Geometry Ascona, 1996," Math. Phys. Stud., Vol. 20, pp. 223–248, Kluwer Academic, Dordrecht.
31. N. Reshetikhin and V. Turaev, Ribbon Graphs and their invariants derived from quantum groups, *Comm. Math. Phys.* **127** (1990), 1–26.
32. A. Szenes, The combinatorics of the Verlinde formulas, *in* "Vector Bundles in Algebraic Geometry, Durham 1993," London Math. Soc. Lecture Notes Ser., Vol. 208, Cambridge Univ. Press, Cambridge, UK, 1995.
33. A. Szenes, Iterated residues and multiple Bernoulli polynomials, *Internat. Math. Res. Notices* **18** (1998).
34. A. Szenes, A residue theorem for trigonometric sums and Verlinde's formula, preprint, math.CO/0109038.
35. C. Teleman, The quantization conjecture revisited, preprint, math.AG/9808029.
36. M. Thaddeus, Geometric invariant theory and flips, *J. Amer. Math. Soc.* **9** (1996), 691–723.

37. V. Tarasov and A. Varchenko, Difference equations compatible with trigonometric KZ differential equations, preprint, math.QA/0002132.
38. V. Turaev, Algebras of loops on surfaces, algebras of knots, and quantization, *in* "Braid Group, Knot Theory and Statistical Mechanics, II," Adv. Ser. Math. Phys., Vol. 17, pp. 324–360, World Scientific, River Edge, NJ, 1994.
39. E. Verlinde, Fusion rules and moduli transformations in 2D conformal field theory, *Nuclear Phys. B* **300** (1988), 360–376.
40. A. Weinstein, Deformation quantization, *in* "Séminaire Bourbaki," Vol. 1993/1994, *Astérisque* **227** (1995).
41. A. Weinstein and Ping Xu, Hochschild cohomology and characteristic classes for star-products, Geometry of differential equations, *Amer. Math. Soc. Transl. Ser. 2* **186** (1998), 177–194.
42. E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis," Cambridge Mathematical Library, Cambridge Univ. Press, Cambridge, UK, 1996.
43. E. Witten, On quantum gauge theories in two dimensions, *Comm. Math. Phys.* **141** (1991), 153–209.