Stavros Garoufalidis

Five selected publications (in one pdf file)

1. *On the Melvin-Morton-Rozansky conjecture* (with D.Bar-Natan), Inventiones Math. **125** (1996) 103-133.
2. *The colored Jones function is q-holonomic* (with T.T.Q. Lê), Geom. and Topology **9** (2005) 1253–1293.
3. *The quantum content of the gluing equations*, (with T. Dimofte), Geometry and Topology, **17** (2013) 1253–1315.
4. *The complex volume of* $\mathrm{SL}(n, \mathbb{C})$ *representations of 3-manifolds* (with D.P. Thurston and C. Zickert), Duke Math. J., **164** (2015) 2099–2160.
5. *Bloch groups, algebraic K-theory, units and Nahm's Conjecture*, (with F. Calegari and D. Zagier),
   preprint 2017, `arXiv:1712.04887`.

Invent. math. 125, 103–133 (1996)

Inventiones mathematicae

On the Melvin–Morton–Rozansky conjecture

Dror Bar-Natan[1,★], Stavros Garoufalidis[2]

[1] Department of Mathematics, Harvard University, Cambridge, MA 02138, USA;
e-mail: drorbn@math.huji.ac.il
[2] Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; e-mail: stavros@math.mit.edu

Oblatum 28-VII-1994 & 5-XI-1995 & 22-XI-1995

Abstract. We prove a conjecture stated by Melvin and Morton (and elucidated further by Rozansky) saying that the Alexander–Conway polynomial of a knot can be read from some of the coefficients of the Jones polynomials of cables of that knot (i.e., coefficients of the "colored" Jones polynomial). We first reduce the problem to the level of weight systems using a general principle, which may be of some independent interest, and which sometimes allows to deduce equality of Vassiliev invariants from the equality of their weight systems. We then prove the conjecture combinatorially on the level of weight systems. Finally, we prove a generalization of the Melvin–Morton–Rozansky (MMR) conjecture to knot invariants coming from arbitrary semi-simple Lie algebras. As side benefits we discuss a relation between the Conway polynomial and immanants and a curious formula for the weight system of the colored Jones polynomial.

Contents

★ *Current address*: Institute of Mathematics, The Hebrew University, Giv'at-Ram, Jerusalem 91904, Israel.
The first author was supported by NSF grant DMS-92-03382.
This paper is available electronically at http://www.ma.huji.ac.il//~drorbn and at ftp://ftp.ma.huji.ac.il/drorbn.

1. Introduction

1.1. The conjecture. In this paper, we will mostly be concerned with proving and explaining some of the motivation for the following conjecture, due to Melvin and Morton [MM, Mo]:

Conjecture 1. *Let $\hat{J}_{sl(2),\lambda}(K) \in \mathbf{Q}(q)$ be the "framing independent colored Jones polynomial" of the knot K, i.e., the framing independent Reshetikhin–Turaev invariant*[1] [RT] *of K colored by the $(d = \lambda+1)$-dimensional representation of $sl(2)$. Let $\hbar$ be a formal parameter, let $q = e^{\hbar}$, and let $[d]$ denote the "quantum integer d":*

$$[d] = \frac{q^{d/2} - q^{-d/2}}{q^{1/2} - q^{-1/2}} = \frac{e^{d\hbar/2} - e^{-d\hbar/2}}{e^{\hbar/2} - e^{-\hbar/2}} \, .$$

Then, expanding $\hat{J}/[d]$ in powers of d and $\hbar$ (this is possible by [MM]*),*

$$\frac{\hat{J}_{sl(2),\lambda}(K)(e^{\hbar})}{[d]} = \sum_{j,m \geqq 0} a_{jm}(K) d^j \hbar^m \, ,$$

we have:

(1) *"Above diagonal" coefficients vanish: $a_{jm}(K) = 0$ if $j > m$.*

(2) *"On diagonal" coefficients give the inverse of the Alexander–Conway polynomial:*

$$MM(K)(\hbar) \cdot A(K)(e^{\hbar}) = 1 \, , \tag{1}$$

where $A(q)$ is the Alexander–Conway polynomial (in its "Conway" normalization, as in example 2.8*) and MM is defined by*

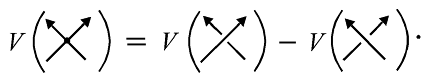
$$MM(K)(\hbar) = \sum_{m=0}^{\infty} a_{mm}(K) \hbar^m \, .$$

Notice that the colored Jones polynomial of a knot can be read from the Jones polynomials of cables of that knot (see, e.g. [MS]), and thus the above conjecture implies that the Alexander polynomial can be computed from the Jones polynomial and cabling operations.

Melvin and Morton arrived at (the rather unexpected) Conjecture 1 after noticing it in some special cases, and by noticing that the two sides of (1) seem to behave in the same way when acted on by the 'Adams operations' of [B-N2]. In his visit to Cambridge in November 1993, we informed L. Rozansky of the conjecture, and he was able [Ro1] to find a non-rigorous path integral "proof" of it, which easily leads to a generalization to other Lie algebras, as shown in Sect. 5. At the end of this introduction we will briefly review the main ideas of Rozansky's work on the MMR conjecture.

[1] I.e., $\hat{J}$ is obtained from the framing-dependent J either by multiplication of $q^{-C\cdot writhe}$ where C is the quadratic Casimir number of V_λ, or by evaluating J on K with its zero framing. We take the metric on $sl(2)$ to be the trace in the 2-dimensional representation.

1.2. Preliminaries. Before we can sketch our proof of the MMR conjecture, let us recall some facts about Vassiliev invariants and chord diagrams, which are the main tools used in the proof. We follow the notation of [B-N2]; see also [Va1, Va2, BL, Ko1]. A *Vassiliev invariant* of type m is a knot invariant V which vanishes whenever it is evaluated on a knot with more than m double points, where the definition of V is extended to knots with double points via the formula



The algebra $\mathscr{V}$ of all Vassiliev invariants (with values in some fixed ring) is filtered, with the type m subspace $\mathscr{F}_m\mathscr{V}$ containing all type m Vassiliev invariants. The associated graded space of $\mathscr{V}$ is isomorphic to the space $\mathscr{W}$ of all *weight systems*. A degree m weight system is a homogeneous linear functional of degree m on the graded vector space $\mathscr{A}^r$ of *chord diagrams* like in Fig. 1 divided by the $4T$ and *framing independence* relations explained in Figs. 2 and 3.

$\mathscr{A}^r$ is graded by the number of chords in a chord diagram. It is a commutative and co-commutative Hopf algebra with multiplication defined by juxtaposition, and with co-multiplication Δ defiend as the sum of all possible ways of 'splitting' a diagram. The co-algebra structure of $\mathscr{A}^r$ defines an algebra structure on $\mathscr{W}$. The Hopf algebra $\mathscr{A}$ is defined in the same way as $\mathscr{A}^r$, only without imposing the framing independence relation.

There are natural maps $W_m : \mathscr{F}_m\mathscr{V} \to \mathscr{G}_m\mathscr{W} = \mathscr{G}_m\mathscr{A}^{r*}$, where $\mathscr{G}_m$ obj denotes the degree m piece of a graded object obj. For a type m Vassiliev

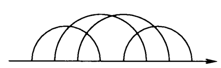


Fig. 1. A chord diagram.

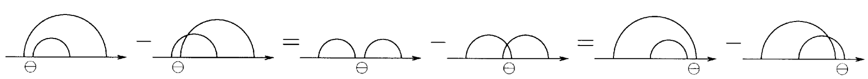


Fig. 2. To get the $4T$ relations, add an arbitrary number of chords in arbitrary positions (only avoiding the short intervals marked by a 'no-entry' sign $\ominus$) to all six diagrams in exactly the same way.

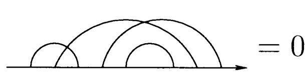


Fig. 3. The framing independence relation: any diagram containing a chord whose endpoints are not separated by the endpoints of other chords is equal to 0.

invariant V it is natural to think of $W_m(V)$ as "the m'th derivative of V". The maps W_m are compatible with the products of the spaces involved. Similar definitions can be made for framed knots, and the image of the corresponding map W_m will be $\mathcal{G}_m\mathcal{A}^*$.

1.3. Plan of the proof. It is well known [Gou, B-N1, B-N2, BL, Lin] that the coefficients of both the Conway and the Jones polynomials are Vassiliev invariants. Normally, Vassiliev invariants are not determined by their weight systems. However, in Sect. 2 we explain (following Kassel [Kas] and Le and Murakami [LM]) that when an invariant comes (in an appropriate sense) from a Lie algebra, it is in fact determined by its weight system. As this is the case for all the invariants appearing in Conjecture 1 (or rather, in the version of it that we actually prove Theorem 1), it is enough to prove Conjecture 1 (that is, Theorem 1) on the level of weight systems.

To do this, we analyze the weight systems of the Conway polynomial and of the invariant MM. In Sect. 3 we analyze the weight system W_C of the Conway polynomial. We find a simple characterization (Theorem 2) of it, and then we use this characterization to show that $W_C(D)$ is the determinant of the intersection matrix $\mathrm{IM}(D)$ (Definition 3.4) of the chord diagram D. In Sect. 4 we go through a rather complicated analysis of the weight system of MM, finding that it is given by the permanent of the intersection matrix. We then conclude the proof of the conjecture by showing that, in the sense of weight systems,

$$\log\det\,\mathrm{IM} + \log\,\mathrm{per}\,\mathrm{IM} = 0\,, \tag{2}$$

and thus the two weight systems are inverses of each other. Equation (2) is proven in the ends of Sects. 3 and 4, where the logarithm of the two weight systems involved are given in terms of explicit formulas.

In Sect. 5 we use similar techniques to generalize Conjecture 1 to arbitrary semi-simple Lie algebras. In Sect. 6.1 we discuss a curious relation between immanants and the algebra generated by the coefficients of the Conway polynomial, in Sect. 6.2 we sketch how the techniques of Sect. 4 can be used to get a formula for the weight system of the colored Jones polynomial, and in Sect. 6.3 we conjecture a generalization of Conjecture 1 beyond the realm of Lie algebras.

As noted before, we actually prove a variation of Conjecture 1 in which the normalizations are somewhat 'better' from the point of view of Sects. 2 and 5:

Theorem 1. *Expanding $\hat{J}/d$ in powers of $\lambda = d - 1$ and $\hbar$,*

$$\frac{\hat{J}_{sl(2),\lambda}(K)(e^{\hbar})}{d} = \sum_{j,m\geqq 0} b_{jm}(K)\lambda^j\hbar^m\,, \tag{3}$$

we have:

(1) "*Above diagonal*" *coefficients vanish*: $b_{jm}(K) = 0$ *if* $j > m$.

(2) *Up to a constant*, "*on diagonal*" *coefficients give the inverse of the Alexander–Conway polynomial*:

$$JJ(K)(\hbar)\cdot\frac{\hbar}{e^{\hbar/2}-e^{-\hbar/2}}A(K)(e^{\hbar})=1\ , \tag{4}$$

where JJ is defined by

$$JJ(K)(\hbar)=\sum_{m=0}^{\infty}b_{mm}(K)\hbar^m\ .$$

Claim 1.1. *Conjecture* 1 *and Theorem* 1 *are equivalent.*

Proof. Let b'_{jm} be the coefficients of the expansion of $\hat{J}/d$ in powers of d and $\hbar$. It is clear that Theorem 1 restated with b'_{jm} replacing b_{jm} is equivalent to the original Theorem 1. We have:

$$\sum a_{jm}d^j\hbar^m=\frac{\hat{J}}{[d]}=\frac{d}{[d]}\cdot\frac{\hat{J}}{d}=\frac{e^{\hbar/2}-e^{-\hbar/2}}{\hbar}\cdot\frac{d\hbar}{e^{d\hbar/2}-e^{-d\hbar/2}}\cdot\sum b'_{jm}\,d^j\hbar^m \tag{5}$$

The first factor in the right hand side of (5) is a power series in $\hbar$ alone in which the coefficient of $\hbar^0$ is 1, and thus it (or its inverse) cannot take below- or on-diagonal terms to go above the diagonal, and it does not change the coefficients on the diagonal. The second factor lives entirely *on* the diagonal and thus the first part of Conjecture 1 is equivalent to the first part of Theorem 1.

Restricted to the diagonal, (5) becomes

$$\sum a_{mm}d^m\hbar^m=\frac{d\hbar}{e^{d\hbar/2}-e^{-d\hbar/2}}\cdot\sum b'_{mm}d^m\hbar^m\ .$$

At $d=1$, we get

$$MM=\frac{\hbar}{e^{\hbar/2}-e^{-\hbar/2}}\cdot JJ\ ,$$

and it is clear that (1) and (4) are equivalent. □

1.4. Rozansky's work. Rozansky arrives at the MMR conjecture using the path integral interpretation of the Jones polynomial given in Witten's seminal paper [Wi]. Needless to say, path integrals have not yet been mathematically defined, but they can be used as a rich source of motivation. In our case they do in fact lead to the correct conjecture, though our proof of the conjecture is not a translation of the path integral argument to rigorous math, and we don't know how to translate the path integral argument into rigorous math. For the convenience of the reader we outline Rozansky's argument below. The reader may find our account somewhat more readable than Rozansky's [Ro1], as we have isolated the parts relevant to Conjecture 1 from his (much broader) paper, and skipped some of the details. We heartily recommend consulting with [Ro1] (as well as [Ro2, Ro3]) for the missing details and for many other related results.

Let us recall Witten's interpretation of the Jones polynomial. For a framed, oriented knot K in S^3, a choice V_λ of an irreducible $SU(2)$ representation of highest weight λ and an integer k, Witten introduces the following definition:

$$Z(K, V_\lambda; k) = \int_{\mathscr{A}} \mathscr{D}A e^{2\pi ik\, CS(A)} \mathscr{O}_{K, V_\lambda}(A)$$

where the (ill defined) path integral is over the space $\mathscr{A}$ of all $SU(2)$ connections on the trivial $SU(2)$ bundle over S^3, $CS : \mathscr{A} \to \mathbf{R}/\mathbf{Z}$ is the Chern–Simons action

$$CS(A) = \frac{1}{8\pi^2} \int_{S^3} \mathrm{tr}(A \wedge dA + \tfrac{2}{3} A \wedge A \wedge A)\,,$$

and $\mathscr{O}_{K, V_\lambda} : \mathscr{A} \to \mathbf{R}$ is the trace in the representation V_λ of the holonomy of the connection A along the knot K.

Using non-rigorous quantum field theory reasoning, Witten computed $Z(K, V_\lambda; k)$ and found that

$$Z(K, V_\lambda; k) = \sqrt{\frac{2}{k+2}} \sin\left(\frac{\pi}{k+2}\right) J_{sl(2), V_\lambda}(K) \left(\exp \frac{2\pi i}{k+2}\right),$$

where $J_{sl(2), V_\lambda}$ is the framing dependent colored Jones polynomial.

Now take a rational number $0 < a \ll 1$ (so that ka is a weight for many large integers k). Following Rozansky [Ro1], the path integral $Z(K, V_{ka}; k)$ (for such k) can be split into an integral over connections on a tubular neighborhood $\mathrm{Tub}(K)$ of the knot K and over connections on the complement $S^3 \backslash \mathrm{Tub}(K)$ with certain boundary conditions on the boundary $T^2 = \partial\, \mathrm{Tub}(K)$, followed by an integral over these boundary conditions. With the appropriate boundary conditions of [EMSS], the integral over the connections on $\mathrm{Tub}(K)$ can be restricted to an integral over flat connections, and on those it is proportional to $\delta(I_1 - e^{2\pi ia})$ *independently* of k, where I_1 is the holonomy along a meridian of K in $\partial\, \mathrm{Tub}(K)$ and $e^{2\pi ia}$ is considered in $SU(2)$ in the usual way. Therefore

$$Z(K, V_{ka}; k) = \int_{\mathscr{A}[S^3 \backslash \mathrm{Tub}(K)]_a} \mathscr{D}A e^{2\pi ik CS'(A)} \tag{6}$$

where the integral is over the connections on $S^3 \backslash \mathrm{Tub}(K)$ with holonomy $e^{2\pi ia}$ along any meridian of K. Here CS' is a modified Chern–Simons action dictated by the boundary conditions.

Rozansky now applies stationary phase approximation to calculate the large k limit of $Z(K, V_{ka}; k)$. The critical points of CS' are the flat $SU(2)$ connections on the knot complement with holonomy $e^{2\pi ia}$ around a meridian. Modulo gauge equivalence, the moduli space of such connections consists of only one connection A_a, for sufficiently small values of a.

By the stationary phase approximation, the leading order term of the path integral is proportional to

$$\frac{1}{\sqrt{8\pi}} \left(\frac{4\pi^2}{k}\right)^{\frac{1}{2}\left(h^0(A_a) - h^1(A_a)\right)} \sqrt{\tau_{RS}(A_a)} \cdot e^{2\pi ik CS'(A_a)}$$

where $h^j(A_a)$ is the dimension of the j'th cohomology of $S^3\backslash\mathrm{Tub}(K)$ with coefficients twisted by A_a, and $\tau_{RS}(A_a)$ is the $SU(2)$ Ray–Singer torsion of $S^3\backslash\mathrm{Tub}(K)$ twisted by A_a. Furthermore one can check that $h^1(A_a) = 0$, $h^0(A_a) = 1$, and $CS'(A_a) = 0$. The Ray–Singer torsion splits into three factors, one for each algebra component of $SU(2)$. The torsion in the Cartan direction is 1, and in the remaining two directions the torsions are equal, and each contributes the square root of the $U(1) \subset SU(2)$ torsion using the representation of $\pi_1(S^3\backslash\mathrm{Tub}(K))$ sending the meridian to $e^{2\pi ia} \in U(1)$. Summarizing, we get

$$\sqrt{\frac{2}{k+2}}\sin\left(\frac{\pi}{k+2}\right)J_{sl(2),V_{ka}}(K)\left(\exp\frac{2\pi i}{k+2}\right)\underset{k\to\infty}{\sim}\frac{1}{\sqrt{2k}}\,\tau_{RS}(S^3\backslash\mathrm{Tub}(K),e^{2\pi ia})\,.$$

Cheeger [Ch] and Müller [Mü] proved that the Ray–Singer torsion is equal to the Reidemeister torsion, which by Milnor [Mi] and Turaev [Tu] was shown to be proportional to the *inverse* of the Alexander polynomial $A(K)$ of K, evaluated at $e^{2\pi ia}$. With the correct constant of proportionality $(2\sin\pi a)$ in place and ignoring factors that converge to 1 as $k\to\infty$, we get

$$\frac{\pi}{k}J_{sl(2),V_{ka}}(K)\left(\exp\frac{2\pi i}{k}\right)\underset{k\to\infty}{\longrightarrow}\frac{\sin\pi a}{A(K)(e^{2\pi ia})}\,.$$

See [Ro1, (2.8) and following paragraph] for an explanation why the J computed here is 'in zero framing'. Thus $J = \hat{J}$ and

$$\pi a\sum_{j,m\geqq 0} b_{jm}(K)(2\pi i)^m a^j k^{j-m}\underset{k\to\infty}{\longrightarrow}\frac{\sin\pi a}{A(K)(e^{2\pi ia})}\,.$$

This proves (on the level of rigor of path integrals) that $b_{jm} = 0$ if $j - m > 0$, and, taking $a = \hbar/2\pi i$ and disregarding all strictly positive powers of k, it also proves Theorem 1 (on the same level of rigor).

2. A reduction to weight systems

Let us start with some generalities that (sometimes) allow us to deduce equality of invariants from the equality of their weight systems. In this section, we mostly interpret and adapt to our needs the deep results of Kassel [Kas] and Le and Murakami [LM], who followed Kohno [Koh] and Drinfel'd [Dr1, Dr2].

2.1. Canonical Vassiliev invariants. A fundamental (and not too surprising) result in the theory of Vassiliev invariants is that every degree m weight system comes from a type m Vassiliev invariant, and that the resulting Vassiliev invariant is well-defined up to Vassiliev invariants of lower types (see e.g. [Ko1] and [B-N2]); in other words, the sequence

$$0\to\mathcal{F}_{m-1}\mathcal{V}\to\mathcal{F}_m\mathcal{V}\to\mathcal{G}_m\mathcal{A}^{r\star}\to 0\,,\tag{7}$$

is exact. The standard way of proving this fact is to construct a splitting $V_m : \mathcal{G}_m\mathcal{A}^{r\star}\to\mathcal{F}_m\mathcal{V}$ for each m. These splittings can be assembled together in a unique way to form a *universal Vassiliev invariant* Z with values in the graded completion of $\mathscr{A}^r$, satisfying

$$V_m(W) = W \circ Z \tag{8}$$

for each degree m weight system W. In fact, usually Z is first constructed, and only then the splittings V_m are defined from it via (8).

A-priori, there appears to be no knot theoretic reason to expect that there would be a preferred choice for the splittings V_m, or, equivalently, for Z. However, rather surprisingly, it seems that such a preferred choice for Z does exist. Indeed, for reasons discovered by Drinfel'd [Dr1, Dr2] and elucidated further by Kassel [Kas] and Le and Murakami [LM], many of the known constructions [B-N3, Ca, Kas, Ko1, LM] of a universal Vassiliev invariant give the same (hard to compute but rather well behaved) answer.[2] Let us call this preferred universal Vassiliev invariant $\mathbf{Z}^{\mathbf{K}}$.

Definition 2.1. *A* **Canonical type *m* Vassiliev invariant** *V is a type m Vassiliev invariant lying in the image of the splitting of* (7) *defined by $\mathbf{Z}^{\mathbf{K}}$. In a simpler language, let $\mathbf{Z}^{\mathbf{K}}_m$ be the projection of $\mathbf{Z}^{\mathbf{K}}$ into $\mathscr{G}_m\mathscr{A}^r$. V is a canonical type m Vassiliev invariant iff*

$$V = W_m(V) \circ \mathbf{Z}^{\mathbf{K}}_m .$$

Definition 2.2. *Let $\hbar$ be a formal parameter. A* **Vassiliev power series** *is an element*

$$V \in \sum_{m=0}^{\infty} (\mathscr{F}_m\mathscr{V})\hbar^m .$$

That is to say, it is a power series $V = V_0 + V_1\hbar + \dots$ in which the coefficient V_m of $\hbar^m$ is a Vassiliev invariant of type m. The **weight system** *$W(V)$ of V will be the sum of the weight systems of the coefficients of V* (*which makes sense in the graded completion $\bar{\mathscr{W}}$ of $\mathscr{W}$*):

$$W(V) = \sum_{m=0}^{\infty} W_m(V_m) \in \bar{\mathscr{W}} .$$

Definition 2.3. *A Vassiliev power series $V = \sum V_m\hbar^m$ is called* **canonical** *if each of its coefficients V_m is canonical. Equivalently, if $\hbar^{\deg}$ is the operator that multiplies every degree m diagram by $\hbar^m$ and $\mathbf{Z}^{\mathbf{K}}_\hbar \stackrel{\text{def}}{=} \hbar^{\deg} \circ \mathbf{Z}^{\mathbf{K}}$, then V is canonical iff*

$$V = W(V) \circ \mathbf{Z}^{\mathbf{K}}_\hbar .$$

Obviously, two canonical Vassiliev power series (*or canonical Vassiliev invariants*) *are equal iff their weight systems are equal. Sometimes, as is the case in this paper, it is easier to verify equality of weight systems and then*

[2] [B-N2, Pi2] differ only by a normalization, and the incomplete perturbative Chern–Simons constructions [AS1, AS2, B-N1, Ko2] are conjectured to also give the same answer.

use it to deduce the equality of the corresponding canonical invariants rather than proving the equality of the invariants directly.

2.2. Examples of canonical Vassiliev power series. In this section we will establish, through a sequence of examples, that the invariants appearing in Theorem 1 are canonical.

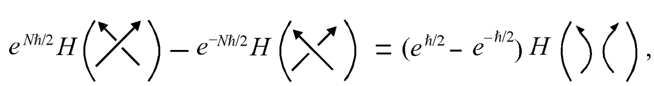
Example 2.4. The type 0 invariant 1, whose value on all knots (having no double points) is 1, is both a canonical type 0 Vassiliev invariant and a canonical Vassiliev power series. Its weight system ε is defined by

$$\varepsilon(D) = \begin{cases} 1 & \text{if } \deg D = 0 \text{ (namely, if } D = \longrightarrow \text{ is the empty diagram)} \ , \\ 0 & \text{otherwise} \ . \end{cases}$$

Kassel [Kas, Theorem 8.3, Chapter XX] and Le and Murakami [LM, Theorem 10], using the techniques of Kohno [Koh] and Drinfel'd [Dr1, Dr2], have shown that the Reshetikhin–Turaev [RT] invariant associated with a semisimple Lie algebra $\mathfrak{g}$ and a representation V (and a metric t on $\mathfrak{g}$) is a canonical Vassiliev power series when evaluated at $q = e^{\hbar}$ and expanded in powers of $\hbar$.[3] (Both the framed version $J_{\mathfrak{g},V}$ and unframed version $\hat{J}_{\mathfrak{g},V}$ are canonical; for the framed version, $\mathscr{A}$ has to replace $\mathscr{A}^r$ in the definitions of this section. For the unframed version (at least when V is irreducible), simply notice that it can always be obtained from the framed version by multiplying the Lie algebra by an Abelian Lie algebra). We will use this crucial result twice, in Example 2.5 and in Example 2.6.

Example 2.5. By [Kas, LM], the invariant $\hat{J}_{sl(2),\lambda}$ of Conjecture 1 is a canonical Vassiliev power series, and hence the invariants b_{jm} of Theorem 1 are canonical of type m, and JJ is a canonical Vassiliev power series. The invariants a_{jm} and MM are not canonical as $[d]$ depends on $\hbar$.

Example 2.6. The HOMFLY polynomial, defined by the relations



$$H\ (c\text{-component unlink}) = \left(\frac{e^{N\hbar/2} - e^{-N\hbar/2}}{e^{\hbar/2} - e^{-\hbar/2}}\right)^c \ ,$$

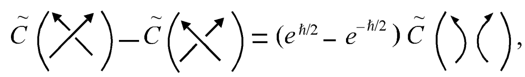
is a canonical Vassiliev power series, as it is the Reshetikhin–Turaev invariant associated with the Lie algebra $sl(N)$ in its defining representation.

Example 2.7. Divide the HOMFLY polynomial by N and take the limit $N \to 0$. The limit exists because the limit

$$\lim_{N\to 0} \frac{e^{N\hbar/2} - e^{-N\hbar/2}}{N} = \hbar$$

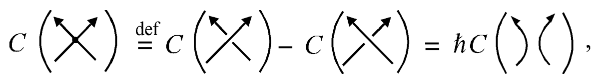
[3] Thus they gave an affirmative answer to problem 4.9 of [B-N2].

exists. The result is a canonical Vassiliev power series $\tilde{C}$ satisfying



$$\tilde{C}\ (c\text{-component unlink}) = \begin{cases} \frac{\hbar}{e^{\hbar/2} - e^{-\hbar/2}} & \text{if } c = 1 \\ 0 & \text{otherwise}. \end{cases} \tag{9}$$

Recall that the Conway polynomial C [Co, Kau] (considered as a polynomial in $\hbar$) is defined by the relations:



$$C\ (c\text{-component unlink}) = \begin{cases} 1 & \text{if } c = 1 \\ 0 & \text{otherwise}. \end{cases} \tag{10}$$

Comparing (9) and (10), we see that the Conway polynomial itself is not a canonical Vassiliev power series, but its renormalized reparametrized version

$$\tilde{C}(\hbar) = \frac{\hbar}{e^{\hbar/2} - e^{-\hbar/2}}\, C(e^{\hbar/2} - e^{-\hbar/2})$$

is a canonical Vassiliev power series.

Example 2.8. The Alexander polynomial, defined by $A(z) = C(z^{1/2} - z^{-1/2})$, is *not* a canonical Vassiliev power series, but it becomes canonical when multiplied by $\frac{\hbar}{e^{\hbar/2} - e^{-\hbar/2}}$ and evaluated at $z = e^{\hbar}$ (as this product is $\tilde{C}$).

2.3. Products. The product (in the natural sense) of two Vassiliev power series is a Vassiliev power series, and the weight system of such a product is the product of the weight systems of the factors.

Proposition 2.9. *The product of any two canonical Vassiliev power series is a canonical Vassiliev power series.*

Proof. It can be shown that the universal Vassiliev invariant $\mathbf{Z}^{\mathbf{K}}$ is 'group-like'; it satisfies $\Delta\mathbf{Z}^{\mathbf{K}}(K) = \mathbf{Z}^{\mathbf{K}}(K) \otimes \mathbf{Z}^{\mathbf{K}}(K)$ for any knot K. This property is an immediate consequence of the Kontsevich integral formula for $\mathbf{Z}^{\mathbf{K}}$ described in [Ko1, B-N2][4]. Now, if $V_{1,2}$ are canonical, then

$$\begin{aligned}
&(W(V_1 V_2) \circ \mathbf{Z}^{\mathbf{K}}_{\hbar})(K) \\
&\quad = (W(V_1)W(V_2))(\mathbf{Z}^{\mathbf{K}}_{\hbar}(K)) && \text{[B-N2, exercise 3.10]} \\
&\quad = (W(V_1) \otimes W(V_2))(\Delta\mathbf{Z}^{\mathbf{K}}_{\hbar}(K)) && \mathscr{A}^r \text{ is a Hopf algebra} \\
&\quad = (W(V_1) \otimes W(V_2))(\mathbf{Z}^{\mathbf{K}}_{\hbar}(K) \otimes \mathbf{Z}^{\mathbf{K}}_{\hbar}(K)) && \mathbf{Z}^{\mathbf{K}} \text{ is group-like} \\
&\quad = (W(V_1) \circ \mathbf{Z}^{\mathbf{K}}_{\hbar})(K)(W(V_2) \circ \mathbf{Z}^{\mathbf{K}}_{\hbar})(K) \\
&\quad = V_1(K)V_2(K)\,,
\end{aligned}$$

and thus $V_1 \cdot V_2$ is also canonical. □

[4] A similar but different statement is [LM, Theorem 4].

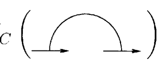
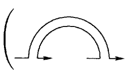
It follows from Examples 2.4, 2.5, and 2.8 and from Proposition 2.9 that both sides of equation (4) are canonical Vassiliev power series, and thus it is enough to prove (4) (as well as the vanishing of b_{jm} for $j > m$) on the level of weight systems. That is, we need to show that

$$W_{JJ} \cdot W_C = \varepsilon\,, \tag{11}$$

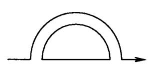
where W_{JJ} is the weight system of JJ, W_C is the weight system of $\tilde{C}$ (which is equal to the weight system of C), and ε is as in Example 2.4.

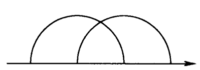
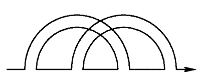
3. The Conway polynomial

3.1. The Conway weight system. The defining relations (10) of C, become the following relations on the level of W_C:

$W_C\Big($  $\Big) = W_C\Big($  $\Big)$ and $W_C\,(c \text{ cycles}) = \begin{cases} 1 & \text{if } c = 0 \\ 0 & \text{otherwise.} \end{cases}$

In other words, to compute W_C of a given chord diagram D, "thicken" all chords in D into bands, and count the number of cycles in the resulting diagram; if it is greater than $0, W_C(D)$ is 0, and otherwise it is 1. For example,

$\theta \stackrel{\text{def}}{=}$  $\longrightarrow$  $\longrightarrow 1 \text{ cycle} \longrightarrow 0,$

$X \stackrel{\text{def}}{=}$  $\longrightarrow$  $\longrightarrow 0 \text{ cycles} \longrightarrow 1.$

These two examples can be combined as in the following definition:

Definition 3.1. *An* (m_1, m_2)**-caravan** *or simply a* **caravan** *is the chord diagram* $\theta^{m_1}X^{m_2}$ *made of* m_1 *single-hump-camels and* m_2 *double-hump-camels, as in Fig.* 4. *It is a chord diagram of degree* $m = m_1 + 2m_2$.

Proposition 3.2.

$$W_C(an(m_1, m_2)\text{-}caravan) = \begin{cases} 1 & \text{if } m_1 = 0 \\ 0 & \text{otherwise}\,. \end{cases} \qquad \square$$

3.2. The 2T relation. It is clear that W_C is invariant under the "$2T$" or "slide" relations shown in Fig. 5. Indeed, after thickening the chords l and r, it is clear

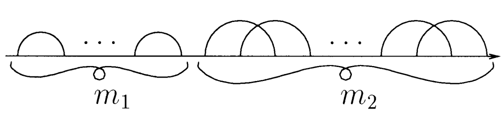


Fig. 4. An (m_1, m_2)-caravan.

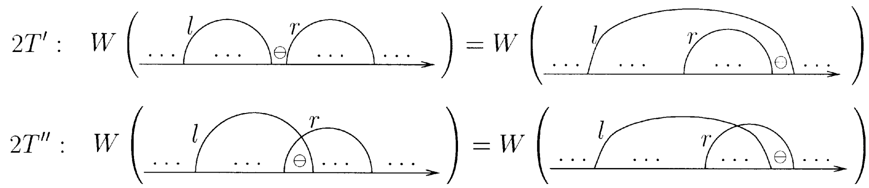


Fig. 5. The $2T$ relations. In these figures, ellipsis denote possible other chords, while a 'no-entry' sign ($\ominus$) means that no chords can end in the corresponding interval. For definiteness, we drew the 'far' end of the chord l left of the chord r, but it can be anywhere else in the diagram.

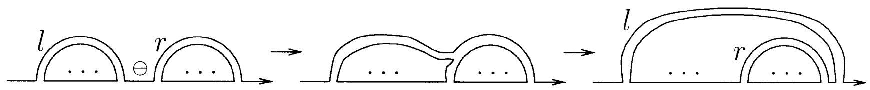


Fig. 6. Deriving the relation $2T'$ by sliding l over r.

that it is possible to 'slide' l over r as in Fig. 6 without changing the topology of the resulting diagram.

Let $\mathscr{G}_m\mathscr{D}$ be the set of all chord diagrams of degree m. The following theorem[5] is a characterization of the Conway weight system:

Theorem 2. *If a map* $W : \mathscr{G}_m\mathscr{D} \to \mathbf{Z}$ *satisfies the* $2T$ *relations and the same* '*initial condition*' *as in proposition* 3.2, *then it is the Conway weight system* W_C.

Proof. It is enough to show that modulo $2T$ relations, every chord diagram D is equivalent to a caravan. If D has a pair of intersecting chords r_1 and r_2, thicken both of them and slide all other chords out and to the left as in Fig. 7. The result is that a double-hump-camel (an X diagram) is factored out. Use induction to simplify the rest. If D has no pairs of intersecting chords, than it must have a 'small' chord r, a chord whose endpoints are not separated by the endpoints of any other chords. Thicken r, and slide all other chords over it and to the left. The result is that a single-hump-camel (a θ diagram) is factored out. Again, use induction to simplify the rest. □

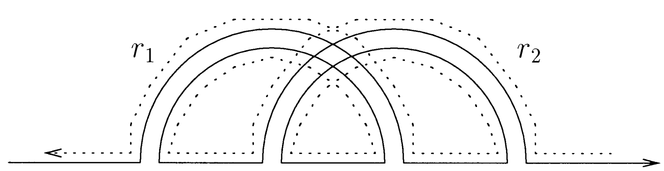


Fig. 7. Factoring out a double-hump-camel. Slide all other chords out following the path marked by a dotted line.

---

[5] P.M. Melvin commented that this is simply the classification theorem for surfaces presented as 'a box with handles'.

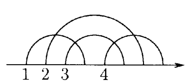
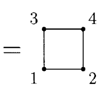
Exercise 3.3. Show that the space of maps $W : \mathcal{G}_m\mathcal{D} \to \mathbf{Z}$ satisfying the $2T$ relations is spanned by the coefficients of various powers of N in $D \mapsto W_{gl(N),V_N}(D)$, where $W_{gl(N),V_N}(D)$ is the weight assigned to D using the Lie algebra $gl(N)$ in its defining representation V_N as in Sect. 4.1 below. Show that such a map that also satisfies the framing independence relations has to be proportional to W_C.

3.3. The intersection graph and the intersection matrix. In this section, we will use Theorem 2 to find a determinant formula for W_C.

Definition 3.4 (See also [CDL1, CDL2, CDL3]). *Let D be a degree m chord diagram. The* **labeled intersection graph** LIG(D) *of D will be the graph whose vertices are the chords of D, numbered from* 1 *to m by the order in which they appear along the 'base line' of D from left to right, and in which two vertices are connected by an edge iff the corresponding two chords in D intersect. The* **intersection matrix** IM(D) *of D is the anti-symmetric variant of the $m \times m$ adjacency matrix of* LIG(D) *defined by*

$$\mathrm{IM}(D)_{ij} = \begin{cases} \operatorname{sign}(i-j) & \textit{if chords } i \textit{ and } j \textit{ of } D \textit{ intersect (where chords} \\ & \quad \textit{of } D \textit{ are numbered from left to right),} \\ 0 & \textit{otherwise.} \end{cases}$$

Example 3.5.

$D =$  , $\mathrm{LIG}(D) =$  , $\mathrm{IM}(D) = \begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$

Example 3.6. The labeled intersection graph of an (m_1, m_2)-caravan is the disconnected union of m_1 single vertices and m_2 graphs like •—•. Its intersection matrix is block diagonal, with the blocks on the diagonal being m_1 copies of the 1×1 zero matrix and m_2 copies of the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.

Exercise 3.7. Show that if the labeled intersection graph of a chord diagram is connected, then the diagram is determined by its intersection matrix. Deduce that in general the intersection matrix determines the class of diagram modulo $4T$ relations.

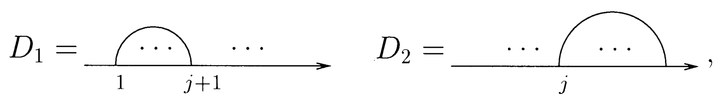
Hint 3.8. Start from a connected labeled intersection graph of a chord diagram, remove one vertex so that the resulting graph is still connected (this is possible!), use induction, and show that there is a unique way to re-install the missing chord.

In the light of the above exercise, it is not surprising that one can find a formula for the weight system of the Conway polynomial in terms of the intersection matrix, as found in the theorem below. A mild generalization of this theorem is in Sect. 6.1. Even though the exercise suggests it should be possible, we have not been able to find nice formulae for other weight systems (beyond those of Sect. 6.1) in terms of the intersection matrix.

Theorem 3. *For any chord diagram D,*

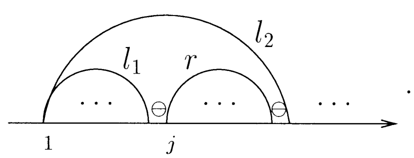
$$W_C(D) = \det \mathrm{IM}(D)$$

Proof. Let $W : \mathscr{G}_m\mathscr{D} \to \mathbf{Z}$ be defined by $W(D) = \det \mathrm{IM}(D)$. By Theorem 2, it is enough to prove that W satisfies the $2T$ relations and the initial conditions of Proposition 3.2. The latter fact is trivial; simply compute the determinant of the block diagonal matrix in Example 3.6. Let us now prove that W satisfies the $2T$ relations. First, notice that W is 'independent of the basepoint of D'. That is, if the diagram D_2 is obtained from the diagram D_1 by moving the left-most vertex of D_1 to the right end,



then $W(D_1) = W(D_2)$. Indeed, except for the labeling the intersection graphs of D_1 and D_2 are the same, and so $\mathrm{IM}(D_2)$ is obtained from $\mathrm{IM}(D_1)$ by reversing all the signs in the first row of $\mathrm{IM}(D_1)$, re-installing it as row number j for some j, and then doing exactly the same to the first column of $\mathrm{IM}(D_1)$. The effect of the row operations is to multiply $\det \mathrm{IM}(D_1)$ by some sign, and then the column operations multiply by the same sign once again. The end result is that $\det \mathrm{IM}(D_1) = \det \mathrm{IM}(D_2)$, as required.

By repeating the above process a few times, we may assume that the chord l in the $2T'$ relation is chord number 1, and so we need to prove that $W(D_1) = W(D_2)$ where $D_1(D_2)$ is the diagram obtained by ignoring $l_2(l_1)$ in the figure



In this figure, it is clear that any other chord can intersect either none of the chords l_1, l_1 and r, or exactly two of them. Using this and some case-checking, it is clear that $\mathrm{IM}(D_2)$ is obtained from $\mathrm{IM}(D_1)$ by adding its jth rows to its first row, and then doing the same column operation. Therefore $\det \mathrm{IM}(D_1) = \det \mathrm{IM}(D_2)$, as required. The same argument also proves the $2T''$ relation. □

In the following two exercises, we outline two alternative proofs of Theorem 3:

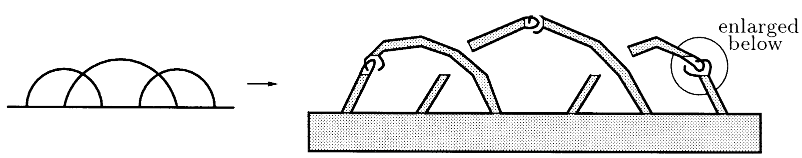
Exercise 3.9. (Melvin) Let F be the surface obtained by thickening a chord diagram D (that is, thicken all chords *and* the base line), and let ∂F be its boundary. $W_C(D) = 1$ if $H_0(\partial F) = \mathbf{Z}$, and otherwise, $W_C(D) = 0$. Now consider the following long exact sequence:

$$\begin{tikzcd} H_1(F) \arrow[r, "p_\star"] & H_1(F,\partial F) \arrow[d, "\gamma"] \arrow[r, "\delta"] & H_0(\partial F) \arrow[r, "i_\star"] & H_0(F) = \mathbf{Z} \arrow[r] & 0 \\ & H^1(F) \arrow[ur, phantom, "\begin{pmatrix}\text{Poincaré}\\ \text{duality}\end{pmatrix}"] & & & \end{tikzcd}$$

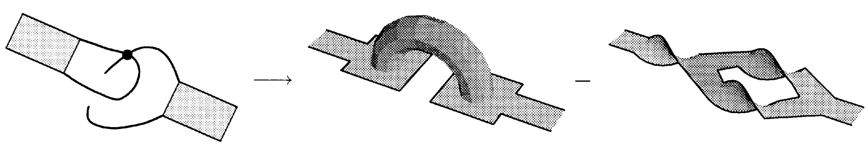
We are interested in knowing when $H_0(\partial F) = \mathbf{Z}$, which is when $p_\star$ is an epimorphism, which is when $\gamma \circ p_\star$ is an epimorphism. Show that in the basis suggested by the chords of D, $\gamma \circ p_\star$ is given by the matrix $\mathrm{IM}(D)$, and use this to deduce Theorem 3. (We wish to thank C. Kassel for reminding us that the determinant of an anti-symmetric matrix is always non-negative).

Exercise 3.10. Deduce theorem 3 from the fact (see e.g. [Kau, Chapter 7]) that the Alexander polynomial of a knot K is given by $\det(z^{-1}\theta - z\theta^T)$, where θ is Seifert pairing matrix for some Seifert surface for K, and θ^T is its transpose.

Hint 3.11. First, take the 'pre-Seifert surface' of a specific singular embedding of a chord diagram as in:



Then resolve all the double points to overcrossings and undercrossings, while extending the 'pre-Seifert surface' to a Seifert surface as in:



It is now easy to compute the $2m \times 2m$ Seifert pairing matrices of the resulting surfaces in terms of the $m \times m$ intersection matrix of the original chord diagram and the over/under choices at the double points.

3.4. The logarithm of the Conway weight system. Expanding $\det \mathrm{IM}(D)$ as a sum over permutations, we only need to consider those permutations of chords(D) which map any chord to a different chord intersecting it. Such permutations can be considered as 'walks' on $\mathrm{LIG}(D)$. Let us introduce the relevant terminology:

Definition 3.12. *A* **Hamilton cycle** *in* $\mathrm{LIG}(D)$ *is a directed cycle H of length* > 1 *in* $\mathrm{LIG}(D)$ *containing no repeated vertices. For example, the graph in example* 3.5 *has two Hamilton cycles of length* 4, *four of length* 2, *and none of any other length. The* **descent** $d(H)$ *of a Hamilton cycle H is the number* *of label-decreases along the cycle. For example, the cycle* $1 \to 2 \to 4 \xrightarrow{\star} 3 \xrightarrow{\star} 1$ *in Example* 3.5 *has descent* 2, *corresponding to the two stared label-decreases. A* **cycle decompositions** $H = \uplus H_\alpha$ *is a cover of the vertex set of* LIG(D) *by a collection of unordered disjoint Hamilton cycle, and the* **descent** $d(H)$ *of* H *is defined by* $d(H) = \sum d(H_\alpha)$.

Expanding $\det \mathrm{IM}(D)$, and taking account of signs, we find that

$$W_C(D) = \sum_{H=\uplus_\alpha H_\alpha} (-1)^{\sigma_H}(-1)^{d(H)} , \tag{12}$$

where σ_H is the permutation of the vertices of LIG(D) underlying H. Notice that if H contains a cycle of odd length, then $(-1)^{d(H)}$ is odd under reversing the orientation of that cycle, while $(-1)^{\sigma_H}$ does not change under that operation. Therefore, summation can be restricted to cycle decompositions containing no odd cycles. For such cycle decompositions, $(-1)^{\sigma_H} = (-1)^{|H|}$, where $|H|$ is the number of cycles in H, and thus

$$W_C(D) = \sum_{H=\uplus_\alpha H_\alpha} (-1)^{|H|}(-1)^{d(H)} . \tag{13}$$

Recall (see. e.g. [B-N2]) that the algebra structure on weight systems is defined by

$$(W_1 \cdot W_2)(D) = \sum_{\substack{\text{splittings} \\ D=D_1 \uplus D_2}} W_1(D_1) \cdot W_2(D_2) . \tag{14}$$

Using the power series expansion of the exponential function, we find that

$$(\exp W)(D) = \sum_{\substack{\text{unordered splittings} \\ D=\uplus D_\alpha}} \prod_\alpha W(D_\alpha) ,$$

and if W depends on D only through LIG(D), we find

$$(\exp W)(D) = \sum_{\substack{\text{unordered splittings} \\ \text{LIG } D=\uplus G_\alpha}} \prod_\alpha W(G_\alpha) ,$$

using the obvious definition for a splitting of a labeled graph.

Proposition 3.13.

$$(\log W_C)(D) = -\sum_H (-1)^{d(H)} ,$$

where the sum extends over all Hamilton cycles H *covering all the vertices of* LIG(D) (*i.e.*, *all cycle decompositions into a single cycle*).

Proof. Simply exponentiate both sides of this equation and use the discussion in the preceeding paragraph to recover (13). □

4. Understanding W_{JJ}

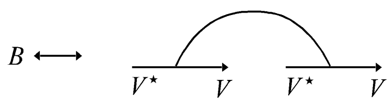
The purpose of this section is to understand W_{JJ}, the weight system underlying the invariant JJ. The invariant JJ, as defined in the statement of Theorem 1, has to do with the Lie algebra $sl(2)$. So let us start by recalling the relation between Lie algebras and weight systems.

4.1. Lie algebras and weight systems. Let $\mathfrak{g}$ be a Lie algebra over some ground field $\mathbf{F}$, let t be a metric (ad-invariant symmetric non-degenerate quadratic form) on $\mathfrak{g}$, and let V be a representation of $\mathfrak{g}$. Given this information, one can construct a weight system [B-N1, B-N2]. Let us recall how this is done.

Choose some basis $\{\mathfrak{g}_a\}_{a=1}^{\dim \mathfrak{g}}$ of $\mathfrak{g}$. Let (t_{ab}) be the matrix corresponding to the metric t in the basis $\{\mathfrak{g}_a\}$; that is, $t_{\mathrm{ab}} = t(\mathfrak{g}_a, \mathfrak{g}_b)$. Let the matrix (t^{ab}) be the inverse of the matrix (t_{ab}), and let $B \in (V^\star \otimes V) \otimes (V^\star \otimes V) = \mathrm{End}(V \otimes V)$ be given by

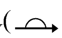
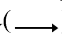
$$B = \sum_{a,b=1}^{\dim \mathfrak{g}} t^{ab} \mathfrak{g}_a \otimes \mathfrak{g}_b \ .$$

We will represent B symbolically by the diagram



(15)

With this notation for B, one can view a chord diagram of degree m as a recipe for how to contract m copies of B and get a tensor $\mathcal{T}(D) \in \mathrm{End}\, V$. This is best explained by an example; see Fig. 8.

One can show (see [B-N1, B-N2]) that the resulting tensor $\mathcal{T}(D)$ is independent of the choice of the basis of $\mathfrak{g}$ (indeed, already B is independent of that choice), is an intertwinner, and that the map $D \mapsto \mathrm{tr}\, \mathcal{T}(D)$ satisfies the $4T$ relation, and hence it descends to a map $W_{\mathfrak{g},V} : \mathcal{A} \to \mathbf{F}$ (the metric t is usually suppressed from the notation). If V is an irreducible representation and C is its quadratic Casimir number (the ratio $W_{\mathfrak{g},V}($$)/W_{\mathfrak{g},V}($$)$), one can define

$$\hat{W}_{\mathfrak{g},V} = W_{\mathfrak{g}\oplus u(1), \hat{V}} \ ,$$

where $\hat{V} = V \otimes \sqrt{-C}$ and $\sqrt{-C}$ denotes the 1-dimensional representation of the 1-dimensional Lie algebra $u(1)$, in which the unit norm generator acts by multiplication by $\sqrt{-C}$. Notice that the representations V and $\hat{V}$ are in the same vector space, and that $\hat{W}_{\mathfrak{g},V}(D)$ can be computed using the same procedure as in Fig. 8, only everywhere replacing B by $\hat{B}$, where $\hat{B} = B - C \cdot I$.

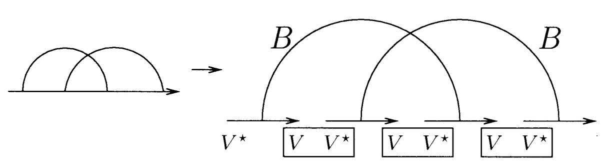


Fig. 8. The construction of $\mathcal{T}(D)$. The B components are as in (15), and pairs of spaces surrounded by a box should be contracted. The two un-boxed spaces are $V^\star$ and V, and thus the result is a tensor in $V^\star \otimes V = \mathrm{End}\, V$.

Recall from Sect. 2.2 that $J_{\mathfrak{g},V}(q)$ is the (framing dependent) Reshetikhin–Turaev knot invariant associated with the algebra $\mathfrak{g}$ and the representation V (and the metric t), and that (when V is irreducible) $\hat{J}_{\mathfrak{g},V}(q) = q^{-C \cdot writhe} \cdot J_{\mathfrak{g},V}(q)$ is its framing independent version. Consider both invariants as Vassiliev power series in the formal parameter $\hbar$ by substituting $q = e^{\hbar}$.

Proposition 4.1. *The weight system (in the sense of Definition 2.2) of $J_{\mathfrak{g},V}$ is $W_{\mathfrak{g},V}$ and (when V is irreducible) the weight system of $\hat{J}_{\mathfrak{g},V}$ is $\hat{W}_{\mathfrak{g},V}$.*

Proof. The framing dependent part is in [Pi1]; it follows easily from the relation $R - (R^{21})^{-1} = \hbar B + o(\hbar)$ satisfied by the quantum Yang–Baxter matrix R. The framing independent part follows from the fact [B-N2, Exercise 6.33] that the weight system corresponding to a direct sum of Lie algebras (and tensor products of representations) is the product of the weight systems of the algebras (and representations) involved, and from a direct (and very simple) analysis of the weight system of $\exp(-\hbar C \cdot writhe)$ and of the weight system $W_{u(1),\sqrt{-C}}$ (see [B-N2, Exercise 6.34]). □

Let us now switch from general consideration to the particular case of $\mathfrak{g} = sl(2)$ and $V = V_\lambda$.

4.2. Understanding $\hat{B}$. In one of the standard models[6] of the representation V_λ, it is spanned by vectors $v_0, \ldots, v_\lambda$, satisfying

$$hv_k = (\lambda - 2k)v_k \ ,$$

$$yv_k = (k+1)v_{k+1}, \quad \text{and} \quad xv_k = (\lambda - k + 1)v_{k-1} \ ,$$

where

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{16}$$

is the standard basis of $sl(2)$. Using the standard scalar product on $sl(2)$ ($\langle M_1, M_2 \rangle = \mathrm{tr}(M_1 M_2)$), we have $\frac{1}{2}\langle h,h\rangle = \langle x,y\rangle = \langle y,x\rangle = 1$, with all other scalar products between h, x, and y vanishing.

Therefore,

$$\hat{B} = y \otimes x + x \otimes y + \tfrac{1}{2} h \otimes h - C \cdot I \ ,$$

where C, the quadratic Casimir number of V_λ, is given by $C = \lambda(\lambda+2)/2$ (see e.g. [Hu, Exercise 4 in Sect. 23]).

[6]Here and later in this paper, we follow the notation of [Hu] for Lie algebras and their representations.

By an explicit computation, we find that

$$
\begin{aligned}
\hat{B}(v_k \otimes v_{k'}) &= (k+1)(\lambda - k' + 1) v_{k+1} \otimes v_{k'-1} \\
&\quad + (\lambda - k + 1)(k' + 1) v_{k-1} \otimes v_{k'+1} \\
&\quad + \tfrac{1}{2}((\lambda - 2k)(\lambda - 2k') - \lambda(\lambda + 2)) v_k \otimes v_{k'} \qquad (17) \\
&= \lambda (B^+ + B^- + I)(v_k \otimes v_{k'}) + (\text{terms of degree 0 in } \lambda), \qquad (18)
\end{aligned}
$$

where

$$B^+ = \sum_{\varepsilon=0,1} (-1)^\varepsilon B^+_\varepsilon; \qquad B^+_\varepsilon (v_k \otimes v_{k'}) = -(k+1) v_{k+\varepsilon} \otimes v_{k'-\varepsilon},$$

and

$$B^- = \sum_{\varepsilon=0,1} (-1)^\varepsilon B^-_\varepsilon; \qquad B^-_\varepsilon (v_k \otimes v_{k'}) = -(k'+1) v_{k-\varepsilon} \otimes v_{k'+\varepsilon}.$$

Proof of part 1 of Theorem 1. Recall that $\hat{B} = \hat{B}(\lambda)$ depends on λ. We wish to study this λ dependence. The different $\hat{B}(\lambda)$'s lie in different spaces, but this is not a serious problem: Let $\hat{V}_\infty$ be the vector space spanned by infinitely many basis vectors $\{v_k\}_{k=0}^\infty$, and extend $\hat{B}(\lambda)$ for all λ to be elements of $\mathrm{End}(\hat{V}_\infty \otimes \hat{V}_\infty)$ using the explicit formula (17). For a chord diagram D, $\mathscr{T}(D) \in \mathrm{End}(\hat{V}_\infty)$ can be constructed as before as in Fig. 8 (no infinite sums occur!), and when restricted to $\hat{V}_\lambda$, the new definition generalizes the old one.

Now that the different $\hat{B}(\lambda)$'s can be compared, equation (18) shows that $\hat{B}(\lambda)$ is at most linear in λ and thus $\mathscr{T}(D)$ is at most of degree m in λ, where $m = \deg D$. Taking the trace of an intertwinner (back again in $\hat{V}_\lambda$!) multiplies by $\lambda + 1$, the dimension of $\hat{V}_\lambda$, and that factor is canceled by the denominator in (3). Finally, by the general considerations of Sect. 2, the result on the level of knot invariants follows from the level of weight systems. □

4.3. Understanding W_{JJ}. Clearly, in computing $W_{JJ}(D)$ for some degree m chord diagram D, it is enough to consider $B^+ + B^- + I$, the coefficient of λ in $\hat{B}$. So let $T(D)$ be the operator constructed as in Fig. 8, only with $B^+ + B^- + I$ replacing B. As $\mathscr{T}(D)$ is an intertwinner, $\mathscr{T}(D) = W_{JJ}(D)I$. Similarly, let $T'(D)$ be the same, only with $B^+ + B^-$ replacing B, and let $W'_{JJ}(D)I$ satisfy $T'(D) = W'_{JJ}(D)I$. It is easy to verify that $W_{JJ} = W'_{JJ} \cdot W_1$, where the product is taken using the coproduct on $\mathscr{A}$ (the space spanned by chord diagrams), and $W_1 \in \mathscr{A}^\star$ satisfies $W_1(D) = 1$ for *any* chord diagram D.

Let D be a degree m chord diagram, and let $(C_\gamma)_{\gamma=1}^m$ be the chords of D, numbered from left to right as in definition 3.4. We are interested in computing $T(D)v_{k(1)}$, or, almost equivalently, $T'(D)v_{k(1)}$, for some non-negative integer $k(1)$. Looking again at Fig. 8 and at (18), we see that $T'(D)v_{k(1)}$ can be computed as follows:

- Sum over the 4^m possible ways of marking the chords $(C_\gamma)_{\gamma=1}^m$ of D by signs $s(\gamma) \in \{+, -\}$ and numbers $\varepsilon(\gamma) \in \{0, 1\}$, corresponding to the choice between $\{B_0^+, B_1^+, B_0^-, B_1^-\}$. Take the term marked by (s, ε) with a sign $\prod_\gamma(-1)^{\varepsilon(\gamma)}$.

- For each fixed choice of (s, ε), add a term determined as follows: Set $k = k(1)$. 'Feed' the marked diagram $D^{(s,\varepsilon)}$ with the vector v_k on the left, and push it right passing it through the vertices of D. Each vertex corresponds to some simple operation, dictated by the marking on the chord C_γ connected to it. The operation is to add or subtract $\varepsilon(\gamma)$ to k, and to multiply by either 1 or $-(k+1)$, using the current value of k for the multiplication. The end result, as read at the right end of $D^{(s,\varepsilon)}$, is proportional to the original $v_{k(1)}$; our term is the corresponding constant of proportionality.

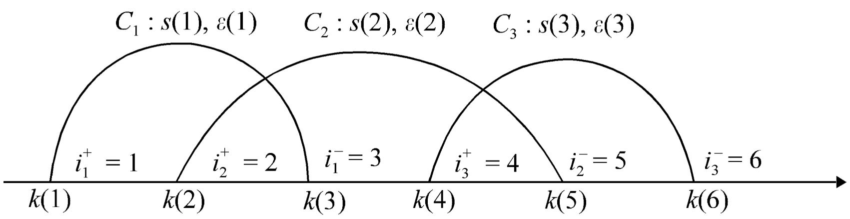
To make the above algorithm more precise and write the result in a closed form, we need to make some definitions. First, number the vertices of D from left to right, beginning with 1 and ending with $2m$. Let i_γ^+ (i_γ^-) be the number of the left (right) end of the chord C_γ, and let the *domain* of C_γ be

$$\operatorname{dom} C_\gamma = (i_\gamma^+, i_\gamma^-] = \{i \in \mathbf{N} : i_\gamma^+ \leqq i < i_\gamma^-\} \, .$$

Let $k(i)$ be the value of k before passing the i'th vertex. It is easy to check that

$$k(i) = k(1) + \sum_{\{\gamma : i \in \operatorname{dom} C_\gamma\}} s(\gamma)\varepsilon(\gamma) \, .$$

Our notation is summarized by the following example:



(19)

Using this notation, the algorithm becomes the following formula:

$$W'_{JJ}(D) = (-1)^m \sum_{\substack{s \in \{+,-\}^m \\ \varepsilon \in \{0,1\}^m}} \prod_{\gamma=1}^m (-1)^{\varepsilon(\gamma)}(1 + k(i_\gamma^{s(\gamma)}))$$

Define the 'difference' operators $\delta/\delta\varepsilon(\gamma)$ on polynomials P in the variables $\varepsilon(\gamma)$, $\gamma = 1, \ldots, m$ by

$$\frac{\delta P}{\delta\varepsilon(\gamma)} = P|_{\varepsilon(\gamma)=1} - P|_{\varepsilon(\gamma)=0} \, . \tag{20}$$

With this definition,

$$W'_{JJ}(D) = (-1)^m \sum_{s \in \{+,-\}^m} \left(\prod_{\gamma=1}^m \frac{\delta}{\delta\varepsilon(\gamma)}\right) \times \left(\prod_{\gamma=1}^m \left(1 + k(1) + \sum_{\{\delta : i_\gamma^{s(\gamma)} \in \operatorname{dom} C_\delta\}} s(\delta)\varepsilon(\delta)\right)\right)$$

Notice that in the above formula we take the m'th partial difference (with respect to $\varepsilon(1), \dots, \varepsilon(m)$) of a polynomial of degree at most m in these variables. By an easy to prove partial difference analog of Taylor's theorem, the result is the coefficient of $\varepsilon(1)\cdots\varepsilon(m)$ in

$$(-1)^m \sum_{s\in\{+,-\}^m} \prod_{\gamma=1}^{m} \left(1 + k(1) + \sum_{\{\delta : i_\gamma^{s(\gamma)} \in \operatorname{dom} C_\delta\}} s(\delta)\varepsilon(\delta) \right) .$$

As only one $\varepsilon(\delta)$ can be picked up from any factor in the product over $\gamma = 1, \dots, m$, this coefficient is the (properly signed) number of choices of an $\varepsilon(\delta)$ for each of these γ's, or, in other words,

$$W'_{JJ}(D) = (-1)^m \sum_{s\in\{+,-\}^m} \sum_{\{\Delta\in S_m : \forall\gamma\, i_\gamma^{s(\gamma)} \in \operatorname{dom} C_{\Delta(\gamma)}\}} \prod_{\gamma=1}^{m} s(\Delta(\gamma)) .$$

The condition in the summation over the permutation Δ can be made a little stronger. Notice that if for a given γ both $i_\gamma^+ \in \operatorname{dom} C_{\Delta(\gamma)}$ and $i_\gamma^- \in \operatorname{dom} C_{\Delta(\gamma)}$ (that is, both ends of the chord C_γ are within the domain of the chord $\Delta(\gamma)$), then the terms with $s(\gamma) = (+)$ cancel the terms with $s(\gamma) = (-)$ in the above sum, and thus summation can be restricted to the cases where this does not happen. In these cases, for each Δ there is a unique choice for the $s(\gamma)$'s for which $\forall\gamma\, i_\gamma^{s(\gamma)} \in \operatorname{dom} C_{\Delta(\gamma)}$. Denote this choice by $s(\Delta, \gamma)$ and get

$$W'_{JJ}(D) = \sum_{\{\Delta\in S_m : \forall\gamma\, C_\gamma \text{ intersects or equals } C_{\Delta(\gamma)}\}} \prod_{\gamma=1}^{m} (-s(\Delta,\gamma)) .$$

Finally, if $\gamma = \Delta(\gamma)$, then necessarily $s(\gamma) = (+)$ and thus $s(\Delta, \gamma) = (+)$. This means that the possibility 'C_γ equals $C_{\Delta(\gamma)}$' can be removed from the above equation by multiplying it by W_1. Thus,

$$W_{JJ}(D) = \sum_{\{\Delta\in S_m : \forall\gamma\, C_\gamma \text{ intersects } C_{\Delta(\gamma)}\}} \prod_{\gamma=1}^{m} (-s(\Delta,\gamma)) .$$

A moment's reflection shows that this formula proves the following proposition:

Proposition 4.2. *$W_{JJ}(D)$ is the permanent* $\operatorname{per} \mathrm{IM}(D)$ *of the intersection matrix* $\mathrm{IM}(D)$ *of D. (Recall that the permanent of a matrix is defined as a sum over permutations in exactly the same way as the determinant, only without the signs).*

4.4. The logarithm of the JJ weight system

Proposition 4.3.

$$(\log W_{JJ})(D) = \sum_H (-1)^{d(H)} ,$$

where the sum extends over all cycle decompositions of $\mathrm{LIG}(D)$ *into a single cycle.*

Proof. Expand $\operatorname{per} \mathrm{IM}(D)$ as a sum over permutations just as in (12), and get

$$W_{JJ}(D) = \sum_{H=\cup_\alpha H_\alpha} (-1)^{d(H)} .$$

Now take the logarithm as in Proposition 3.13. □

Comparing this with proposition 3.13, we find that $\log W_C + \log W_{JJ} = 0$, proving equation (11) and concluding the proof of the Melvin–Morton–Rozansky conjecture.

5. The MMR conjecture for general semi-simple Lie algebras

Let $\hat{J} = \hat{J}_{\mathfrak{g}, V_\lambda}(K) \in \mathbf{Q}(q)$ be the framing-independent Reshetikhin–Turaev invariant of the knot K for the semi-simple Lie algebra $\mathfrak{g}$ and the irreducible representation V_λ of $\mathfrak{g}$ of highest weight λ. (The metric on $\mathfrak{g}$ will be the Killing form $\langle \cdot, \cdot \rangle$). In this section we will prove an analog of Theorem 1 (and thus of Conjecture 1) for $\hat{J}$.

Choose a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, denote by Φ the set of all roots of $\mathfrak{g}$ in $\mathfrak{h}^\star$, and by Φ^+ the set of all positive roots. Let $\langle \cdot, \cdot \rangle$ also denote the scalar product on $\mathfrak{h}^\star$ induced by the Killing form.

The following theorem is suggested by the same reasoning as in Sect. 1.4, only replacing $SU(2)$ by $\mathfrak{g}$. The main difference is that $\tau_{RS}(A_a)$ splits into a product of $\dim \mathfrak{g}$ Abelian torsions, rather than just 3. The torsions along the Cartan directions are still 1, while those along the negative roots pair with those along the positive roots to give a product of Alexander polynomials (appearing under the alias $\tilde{C}$, discussed in Examples 2.7 and 2.8):

Theorem 4. (*Proven in Sects.* 5.1–5.4). *Regarding $\hat{J}(K)(e^\hbar)/\dim V_\lambda$ as a power series in $\hbar$ whose coefficients are polynomials in λ, we have*:

(1) *The coefficient $\hat{J}_m$ of $\hbar^m$ is of degree at most m in λ.*

(2) *If $JJ_\mathfrak{g}$ is the power series in $\hbar$ whose degree m coefficient is the homogeneous degree m piece of $\hat{J}_m$, then*

$$JJ_\mathfrak{g}(K)(\hbar) \cdot \prod_{\alpha\in\Phi^+} \tilde{C}(K)(\langle\lambda,\alpha\rangle\hbar) = 1 . \tag{21}$$

(Since on a simple Lie algebra every invariant scalar product is a multiple of the Killing form and the left-hand-side of (21) is clearly multiplicative under taking the direct sum of Lie algebras, it follows that (21) still holds when $\langle \cdot, \cdot \rangle$ is replaced by an arbitrary invariant scalar product on $\mathfrak{g}$, in both the $\tilde{C}$ part of the equation and in the definition of $\hat{J}$.)

As in Sect. 2, it is enough to prove Theorem 4 on the level of weight systems. Furthermore, in the light of Theorem 1, in order to prove (21) it is enough to prove that

$$W_{JJ,\mathfrak{g}} = \prod_{\alpha\in\Phi^+} W_{JJ} \circ \langle\lambda,\alpha\rangle^{\deg} , \tag{22}$$

where $\langle\lambda,\alpha\rangle^{\deg}$ is defined as in Definition 2.3.

5.1. Lie-algebraic preliminaries. Let $\mathfrak{g} = \mathfrak{h} \oplus (\oplus_{\alpha\in\Phi} L_\alpha)$ be the root space decomposition of $\mathfrak{g}$. Recall (e.g. [Hu]) that $\mathfrak{h}$ is orthogonal to all the L_α's, that L_α is orthogonal to L_β unless $\alpha+\beta = 0$ and that one can find $x_\alpha \in L_\alpha,\ y_\alpha \in L_{-\alpha}$, for all $\alpha \in \Phi$ so that

Setting $h_\alpha = [x_\alpha, y_\alpha]$, the triple $\{x_\alpha, y_\alpha, h_\alpha\}$ spans a subalgebra of $\mathfrak{g}$ isomorphic to $sl(2)$ via the map $(x_\alpha, y_\alpha, h_\alpha) \mapsto (x, y, h)$, where $\{x, y, h\}$ are as in (16). (23)

$$\langle x_\alpha, y_\alpha\rangle = 2/\langle\alpha,\alpha\rangle\ . \tag{24}$$

For any $\lambda \in \mathfrak{h}^\star$ and $\alpha \in \Phi \subset \mathfrak{h}^\star$, one has $\lambda(h_\alpha) = 2\langle\lambda,\alpha\rangle/\langle\alpha,\alpha\rangle$. (25)

An additional property worth recalling is

$$\text{For any } \alpha,\beta \in \Phi, [L_\alpha, L_\beta] \subset L_{\alpha+\beta}\ . \tag{26}$$

Choose a total ordering $<$ of Φ^+ for which $\alpha,\beta < \alpha+\beta$ for any $\alpha,\beta \in \Phi^+$. (For example, you can order the roots by the lengths of their projections on some generic vector in the fundamental Weyl chamber). Let $v_0 \in V_\lambda$ be a highest weight vector; that is, a vector satisfying $hv_0 = \lambda(h)v_0$ for all $h \in \mathfrak{h}$ and $x_\alpha v_0 = 0$ for all $\alpha \in \Phi^+$. Let $\mathbf{Z}_+\Phi^+ = \{\sum_{\alpha\in\Phi^+} k_\alpha\bar\alpha : \forall\alpha\ k_\alpha \in \mathbf{Z}_+\}$ be the semi-group of formal linear combinations of symbols $\bar\alpha$, one for each $\alpha \in \Phi^+$, with non-negative integer coefficients. Define a map $\{\cdot\} : \mathbf{Z}_+\Phi^+ \to \mathfrak{h}^\star$ by $\{\sum k_\alpha\bar\alpha\} = \sum k_\alpha\alpha$. Order $\mathbf{Z}_+\Phi^+$ lexicographically, that is, declare that $\sum k_\alpha\bar\alpha < \sum k'_\alpha\bar\alpha$ iff for some $\beta, k_\beta < k'_\beta$ and $k_\alpha = k'_\alpha$ for all $\alpha < \beta$. For any $k \in \mathbf{Z}_+\Phi^+$, set

$$v_k = \left(\prod_{\alpha\in\Phi^+} \frac{y_\alpha^{k_\alpha}}{k_\alpha!}\right) v_0\ , \tag{27}$$

where the k_α's are the coefficients of k and the product is taken using a decreasing order for the y_α's, so that, for example, if $\alpha > \beta$, then

$$v_k = \left(\cdots\frac{y_\alpha^{k_\alpha}}{k_\alpha!}\cdots\frac{y_\beta^{k_\beta}}{k_\beta!}\cdots\right) v_0\ . \tag{28}$$

The action of $\mathfrak{g}$ on V_λ is given by the following

Lemma 5.1. *With the notation as above we have that*

$$hv_k = (\lambda - \{k\})(h)v_k\ , \tag{29}$$

$$y_\alpha v_k = (k_\alpha + 1)v_{k+\bar\alpha} + \sum_{\substack{j\in\mathbf{Z}_+\Phi^+\\ j>k+\bar\alpha}} c_1(\alpha,k,j)v_j \tag{30}$$

$$x_\alpha v_k = \frac{2}{\langle\alpha,\alpha\rangle}\langle\lambda,\alpha\rangle v_{k-\bar\alpha} + \sum_{\substack{j\in\mathbf{Z}_+\Phi^+\\ j>k-\bar\alpha}} c_2(\lambda,\alpha,k,j)v_j + O(1)\ , \tag{31}$$

where c_1 does not depend on λ, c_2 is linear in λ, and here and in the next few paragraphs $O(1)$ means terms independent of λ.

The importance of the precise form of the 'remainder terms' in the above lemma will be better understood after reading the proof of Lemma 5.2. We therefore postpone the proof of Lemma 5.1 to Sect. 5.4.

5.2. Understanding $\hat{B}$. As in Sect. 4, the key to understanding $W_{JJ,\mathfrak{g}}$ is to first understand $\hat{B} \in \mathrm{End}(\hat{V}_\lambda \otimes \hat{V}_\lambda)$, where $\hat{V}_\lambda = V_\lambda \otimes \sqrt{-C}$ and $\sqrt{-C}$ denotes the 1-dimensional representation of the 1-dimensional Lie algebra $u(1)$, in which the unit norm generator acts by multiplication by $\sqrt{-C}$, and C is the quadratic Casimir number of V_λ.

Let $\{h_i\}_{i=1}^r$ be the arbitrary $\langle \cdot, \cdot \rangle$-orthonormal basis of $\mathfrak{h}$. Using (24), we find that

$$\hat{B} = \sum_{\alpha\in\Phi^+} \frac{\langle\alpha,\alpha\rangle}{2}(x_\alpha \otimes y_\alpha + y_\alpha \otimes x_\alpha) + \sum_{i=1}^r h_i \otimes h_i - C \cdot I \, .$$

Since the quadratic Casimir number C of the representation V_λ is $\langle \lambda + 2\rho, \lambda\rangle$, where $\rho = 1/2 \sum_{\alpha\in\Phi^+} \alpha$ is half the sum of the positive roots [Hu, Exercise 4 in Sect. 23], we also have that

$$\begin{aligned}
&\left(\sum_{i=1}^4 h_i \otimes h_i - C\right) v_k \otimes v_{k'} \\
&\quad = (((\lambda - \{k\}) \otimes (\lambda - \{k'\}))(\textstyle\sum h_i \otimes h_i) - C)v_k \otimes v_{k'} \quad \text{by Lemma 5.1} \\
&\quad = (\langle \lambda - \{k\}, \lambda - \{k'\}\rangle - \langle \lambda, \lambda + 2\rho\rangle)v_k \otimes v_{k'} \qquad \text{by Pythagoras' Theorem} \\
&\quad = -\langle \lambda, 2\rho + \{k\} + \{k'\}\rangle v_k \otimes v_{k'} + O(1) \\
&\quad = -\sum_{\alpha\in\Phi^+} \langle\lambda,\alpha\rangle(1 + k_\alpha + k'_\alpha)v_k \otimes v_{k'} + O(1) \qquad \text{expanding } \rho, \{k\}, \{k'\} \, .
\end{aligned}$$

Using the above formula and Lemma 5.1 we get that

$$\hat{B} = \sum_{\alpha\in\Phi^+} \langle\lambda,\alpha\rangle(B_\alpha^+ + B_\alpha^- + I) + B_{\text{rest}} + O(1) \tag{32}$$

where

$$\begin{aligned}
B_\alpha^+(v_k \otimes v_{k'}) &= -(k_\alpha + 1) \sum_{\varepsilon=0,1} (-1)^\varepsilon v_{k+\varepsilon\bar{\alpha}} \otimes v_{k'-\varepsilon\bar{\alpha}} \\
B_\alpha^-(v_k \otimes v_{k'}) &= -(k'_\alpha + 1) \sum_{\varepsilon=0,1} (-1)^\varepsilon v_{k-\varepsilon\bar{\alpha}} \otimes v_{k'+\varepsilon\bar{\alpha}} \\
B_{\text{rest}}(v_k \otimes v_{k'}) &= \sum_{\substack{j,j'\in\mathbf{Z}_+\Phi^+ \\ j+j'>k+k'}} c_3(\lambda,k,k',j,j')v_j \otimes v_{j'} \, ,
\end{aligned}$$

and where c_3 (which is a simple combination of $c_{1,2}$) is linear in λ.

Since $\hat{B}$ is at most linear in λ we conclude the first part of Theorem 4 as in Sect. 4.2.

5.3. Understanding $W_{JJ,\mathfrak{g}}$. Reading Sect. 4.3 once again and looking at Fig. 8, we see that $W_{JJ,\mathfrak{g}}(D)$ is a certain summation over all the possible ways of labeling the chords of D by I, B_α^+, B_α^-, or B_{rest}.

Lemma 5.2. *In the summation making $W_{JJ,\mathfrak{g}}(D)$, terms containing a chord labeled by B_{rest} can be ignored.*

Proof. This statement is best proven by an example. Let $k(i)$ be the value of k before passing the i's vertex of D, as in (19) (but notice that now $k(i)$ is in $\mathbf{Z}_+\Phi^+$ rather than in $\mathbf{Z}_+$). Similarly, let $k(7)$ be the value of k after passing the sixth vertex (assuming, for the sake for this example, that D is the diagram in (19)). As $\mathscr{T}(D)$ is an intertwinner, it has to be a multiple of the identity and thus $k(7) = k(1)$. On the other hand, by (32) (and remembering that in as much as $W_{JJ,\mathfrak{g}}$ is concerned, we need not care about the $O(1)$ term), we find that

$$\begin{aligned} k(1) + k(3) &\geqq k(2) + k(4)\,, \\ k(2) + k(5) &\geqq k(3) + k(6)\,, \\ k(4) + k(6) &\geqq k(5) + k(7)\,. \end{aligned}$$

Adding these inequalities, we get $k(1) \geqq k(7)$, and this inequality becomes strict if any of the previous ones is strict. As we know that $k(1) \geqq k(7)$ cannot be strict, we learn that none of the previous ones is, and thus we can ignore B_{rest} (as it would correspond to a strict inequality). □

Therefore, in computing $W_{JJ,\mathfrak{g}}(D)$, it is enough to consider

$$\sum_{\alpha\in\Phi^+} \langle\lambda,\alpha\rangle(B_\alpha^+ + B_\alpha^- + I)\,. \tag{33}$$

Nicely enough, the different summands in (33) are 'decoupled'. For each $\alpha, B_\alpha^\pm$ cares only about the α components of the $k(i)$'s, and changes only these components. This amounts to saying that $W_{JJ,\mathfrak{g}}$ is the product of the weight systems corresponding to the different summands. Comparing the definition of $B_\alpha^\pm$ with the definition of $B^\pm$ in Sect. 4, we find that we've proven (22) and hence we've proven Theorem 4. □

5.4. Proof of Lemma 5.1. (29) is just the well known statement that the y_α's act as 'lowering operators'. To prove (30), let us compute $y_\alpha\prod_\beta(y_\beta^{k_\beta}/k_\beta!)$ (using the same convention as in (28) for the ordering of products). To bring this expression to the form of (27), we need to commute y_α to its place, next to the term $y_\alpha^{k_\alpha}/k_\alpha!$. This done, the result is

$$\cdots y_\alpha \frac{y_\alpha^{k_\alpha}}{k_\alpha!} \cdots = \cdots (k_\alpha + 1)\frac{y_\alpha^{k_{\alpha+1}}}{(k_\alpha+1)!} \cdots\,,$$

explaining the first term in (30). However, en route to its place, we needed to commute y_α with various y_β's for which $\beta > \alpha$. By (26) and the choice of the order $<$, such commutators are proportional to y_γ's with even bigger γ's, explaining the remainder term in (30). To be fair, the resulting y_γ's also need to be taken to their respective places, at the cost of some more commutators proportional to even bigger y_δ's, but that doesn't disturb (30). A complete argument can be given using the PBW theorem for the subalgebra of $\mathfrak{g}$ generated by $\{y_\beta : \beta > \alpha\}$, but we don't feel this is necessary.

The proof of (31) is a little harder, but goes along similar lines. Consider an expression like $x_\alpha \prod_\beta (y_\beta^{k_\beta}/k_\beta!)$. Commuting x_α all the way to the right, we get a product that kills the highest weight vector v_0. Along the way, we pick up three kinds of commutators:

(1) First, we pick some $[x_\alpha, y_\beta]$'s, with $\beta > \alpha$. By (26), if $\beta > \alpha$, $[x_\alpha, y_\beta]$ is proportional to some y_γ, resulting in terms which are products of y's, and thus they fall into the third summand of (31), $O(1)$.

(2) We then pick the term containing $[x_\alpha, y_\alpha^{k_\alpha}]$, which, using (23), gives

$$\prod_{\beta>\alpha} \frac{y_\beta^{k_\beta}}{k_\beta!} \cdot \frac{1}{k_\alpha!} \left(\sum_{i=1}^{k_\alpha} y_\alpha^{i-1} h_\alpha y_\alpha^{k_\alpha - i} \right) \cdot \prod_{\beta<\alpha} \frac{y_\beta^{k_\beta}}{k_\beta!} \ .$$

By (29), applied to v_0 this is $\lambda(h_\alpha) v_{k-\bar{\alpha}} + O(1)$, and by (25), this is

$$\frac{2}{\langle \alpha, \alpha \rangle} \langle \lambda, \alpha \rangle v_{k-\bar{\alpha}} + O(1) \ ,$$

explaining the first term in (31).

(3) Finally, we get terms containing $[x_\alpha, y_\beta]$'s, with $\beta < \alpha$. By (26), if $\beta < \alpha$, $[x_\alpha, y_\beta]$ is proportional to some x_γ with $\gamma < \alpha$. Such x_γ are pushed to the right recursively using the same procedure we've used so far, at the cost of (at most) terms independent of λ and terms linear in λ, as in case (2) above, but with $v_{k-\bar{\gamma}}$ (or $v_{k-\bar{\delta}}$ for even smaller δ) replacing $v_{k-\bar{\alpha}}$. Such terms fall into the middle term of (31). □

6. Odds and ends

6.1. Immanants and the Conway polynomial. Theorem 3 and Proposition 4.2 show (in particular) that both the map $D \mapsto \det \mathrm{IM}(D)$ and the map $D \mapsto \operatorname{per} \mathrm{IM}(D)$ are weight systems. It is tempting to look for common generalizations of these two weight systems. In this section, which may be of some independent interest, we sketch just such a generalization. The basic idea is that just where the character of the alternating representation of the symmetric group S_m is used in the definition of det and the character of the trivial representation is used in the definition of per, one can put the character of an arbitrary representation of S_m:

Definition 6.1. Let $[\sigma]$ denote the conjugacy class of permutation σ. Let ZS_m be the free $\mathbf{Z}$-module generated by the conjugacy classes of S_m. Let $ZS_\star$ be the graded $\mathbf{Z}$-module whose degree m piece is ZS_m. The natural embedding $\iota : S_m \times S_n \to S_{m+n}$ makes $ZS_\star$ an algebra by setting $[\sigma][\tau] = [\iota(\sigma, \tau)]$. Identifying $ZS_\star$ with its dual by declaring each individual conjugacy class $[\sigma]$ to be of unit norm, the product on $ZX_\star$ becomes a co-product on $ZS_\star^\star = ZS_\star$.

Exercise 6.2. Verify that with the above product and co-product $ZS_\star$ becomes a graded commutative and co-commutative Hopf algebra, and that the primitive elements of $ZS_\star$ are exactly the classes of cyclic permutations (and thus $ZS_\star$ has exactly one generator in each degree).

Definition 6.3. (Compare with [Lit]) Let M be an $m \times m$ matrix. The **universal immanant** imm M of M is defined by

$$\operatorname{imm} M = \sum_{\sigma \in S_m} [\sigma] \prod_{i=1}^{m} M_{i\sigma i} \in ZS_m \ .$$

(Exactly the same as the definition of det M, only with $[\sigma]$ replacing $(-1)^\sigma$).

Composing the universal immanant with characters of arbitrary representations of S_m, one gets specific complex valued "immanants". Taking the representation to be the alternating representations, one gets det M. Taking it to be the trivial representation, one gets per M. Much is known about many other immanants; see e.g. [GJ, St1, St2].

In our context, we will be interested in the universal immanant of the intersection matrix of a chord diagram. By abuse of notation, we will write imm D for imm IM(D).

Theorem 5. (1) *The map* imm: $\{chord\ diagrams\} \to ZS_\star$ *descends to a well defined map* imm: $\mathscr{A}^r \to ZS_\star$.

(2) *The thus defined* imm: $\mathscr{A}^r \to ZS_\star$ *is a morphism of Hopf algebras.*

(3) *The image of adjoint map* $\mathrm{imm}^\star : ZS_\star^\star = ZS^\star \to \mathscr{A}^{r\star} = \mathscr{W}$ *is the subalgebra of* $\mathscr{W}$ *generated by the weight systems of the coefficients of the Conway polynomial.*

Proof. (sketch) Let L_m be the degree of m piece of $\log W_C$, and let $C_m \in S_m$ be a cyclic permuation. Re-interpreted in our new language, Proposition 3.13 is simply the statement $\mathrm{imm}^\star[C_m] = -L_m$ and equation (14) becomes the multiplicativity of $\mathrm{imm}^\star$. It follows that the image of $\mathrm{imm}^\star$ is equal to the subalgebra of the algebra of functionals on chord diagrams generated by the L_m's. As L_m is known to be a weight system and the product of two weight systems is again a weight system, it follows that the image of $\mathrm{imm}^\star$ is in $\mathscr{W}$ and thus imm descends to $\mathscr{A}^r$. Finally notice that the algebra generated by the L_m's is equal to the algebra generated by the weight systems of the coefficients of the Conway polynomial. □

It is easy to check (or deduce from Theorem 5) the $\mathrm{imm}^\star[\sigma] = 0$ if σ has a cycle of an odd length. By evaluating $\mathrm{imm}^\star[\sigma]$ on chord diagrams whose intersection graph is a union of polygons of an even number of sides, one can see that $\mathrm{imm}^\star$ restricted to permutations with no cycles or odd length is injective.

Exercise 6.4. Check that if IM(D) is replaced by IM$(D) + \lambda I$ for any non-zero constant λ and $\mathscr{A}^r$ and $\mathscr{W}$ are replaced by $\mathscr{A}$ and $\mathscr{A}^\star$ in the statement of theorem 5, the theorem remains valid, with the unique element of $\mathscr{G}_1\mathscr{A}^\star$ adjoined to the generators of the image of $\mathrm{imm}^\star$.

6.2. A curious formula for the weight system of the colored Jones polynomial. (A sketch). The key to the understanding of W_{JJ} in Sect. 4.3 was to rewrite (17) in a nicer form, equation (18). There is an even nicer form, however, that also includes the terms independent of λ: (suppressing '$\otimes$' symbols)

$$\begin{aligned}\hat{B}(v_k v_{k'}) = \lambda\Bigg((k+1)\underbrace{(v_{k+1}v_{k'-1} - v_k v_{k'})}_{\text{part 1}} \\ -(k'+1)\times\underbrace{(v_k v_{k'} - v_{k-1}v_{k'+1})}_{\text{part 2}} + v_k v_{k'}\Bigg) \\ +(k-k')\underbrace{(v_{k+1}v_{k'-1} - v_{k-1}v_{k'+1})}_{\text{part 3}} + v_{k+1}v_{k'-1} + v_{k-1}v_{k'+1} \\ - kk'\underbrace{(v_{k+1}v_{k'-1} - 2v_k v_{k'} + v_{k-1}v_{k'+1})}_{\text{part 4}}\,.\end{aligned}$$

Following roughly the same steps as in Sect. 4.3, parts 1 and 2 of the above equation become 'derivatives' like in (20). Part 3 also becomes a derivative, but with an additional factor of 2 as in it '$\Delta k = 2$'. Part 4 becomes a 'second derivative', and all other parts remain '0th order'. These 'differentiations' mean that we want to look at the coefficients of certain monomials in the ε's of Sect. 4.3, and when all the dust settles we remain with the following (completely self-contained) formula:

Theorem 6. *Let D be a chord diagram of degree m, and let $i_\gamma^\pm$ and* $\operatorname{dom} C_\gamma$ *be as in Sect.* 4.3. *Let $\varepsilon(\gamma)$ be commuting indeterminates, and let*

$$k(i) = \sum_{\{\gamma: i\in \operatorname{dom} C_\gamma\}} \varepsilon(\gamma)\,.$$

Then $W_{\hat{J}}(D)$ (the weight of D in the weight system of the framing-independent Reshetikhin–Turaev invariant of $sl(2)$ in the $(\lambda+1)$st dimensional representation) is the term independent of all the $\varepsilon(\gamma)$'s in

$$(\lambda+1)\prod_{\gamma=1}^{m}\left((\lambda+2)\left(1+\frac{k(i_\gamma^+)-k(i_\gamma^-)}{\varepsilon(\gamma)}\right) - 2\,\frac{k(i_\gamma^+)k(i_\gamma^-)}{\varepsilon(\gamma)^2}\right). \qquad \square$$

Exercise 6.5. Deduce the equality $W_{JJ}(D) = \operatorname{per}\operatorname{IM}(D)$ from the above theorem.

Arguing similarly but starting from the 'framed' $B = x\otimes y + y\otimes x + h\otimes h/2$, one finds that the weight of D in the weight system of the framing-*dependent* Reshetikhin–Turaev invariant of $sl(2)$ in the $(\lambda+1)$st dimensional representation is the term independent of all the $\varepsilon(\gamma)$'s in

$$(\lambda+1)\prod_{\gamma=1}^{m}\left((\lambda+2)\left(1+\frac{\lambda}{2}+\frac{k(i_\gamma^+)-k(i_\gamma^-)}{\varepsilon(\gamma)}\right)-2\,\frac{k(i_\gamma^+)k(i_\gamma^-)}{\varepsilon(\gamma)^2}\right).$$

Remark 6.6. Experimentally (on a computer) we found that the above formulas appear to be (by far) the best method for computing the corresponding weight systems. But, in some sense, we do not understand them very well:

(1) Our only proof that the above formulas satisfy the $4T$ relation is by tracing them back to $sl(2)$. It would be interesting to find a direct proof.

(2) We do not know how to generalize these formulas to other Lie algebras.

(3) We do not know how to view these formulas in the context of Rozansky's work. More specifically, it should be possible to push exercise 6.5 a little further and get formulas for the 'sub-diagonal' invariants $JJ_n = \Sigma_m b_{m-n,m}\hbar^m$ (for small n), and it should be possible to expand (6) in powers of $1/k$ using Feynman diagrams. The $1/k^n$ term in (6) should equal JJ_n. In this paper we dealt with the case $n = 0$ but we don't know how to deal with higher values of n.

6.3. A further generalization. If, as conjectured in [B-N2], all weight systems come from Lie algebras, then there should be a way of stating and proving Theorem 4 without any reference to Lie algebras. We do not have a precise analog of the statement; without a Lie algebra, it is not clear what λ is and in which space it should be. However, on the level of group representations, $\psi^n V_\lambda = V_{n\lambda} +$ (representations of a smaller highest weight), and thus the Adams operations ψ^n, which have a generalization to arbitrary weight systems [B-N2], can play a role similar to 'scaling λ'. We thus arrive at the following conjecture[7]:

Conjecture 2. *Let W be an arbitrary weight system, let n be an integer, and let $\hat{W}^n = \widehat{\psi^{n\star}W}$ be the deframed version (as in* [B-N2, *Exercise* 3.16]) *of $W \circ \psi^n$, where ψ^n is the nth Adams operation on chord diagrams. Then*

(1) *For any fixed chord diagram D of degree m, $\hat{W}^n(D)$ is a polynomial in n of degree at most m.*

(2) *Let $\hat{W}^{n,m}(D)$ be the degree m piece of $\hat{W}^n(D)$. Then the weight system $\hat{W}^{n,m}$ is in the algebra generated by the coefficients of the Conway polynomial.*

A similar statement should hold on the level of knot invariants, using the '0-framing' of a knot for the Adams operations.

Acknowledgement. We wish to thank N. Bergeron, P. Diaconis, C. Kassel, R.C. Kirby, H.R. Morton, L. Rozansky, S. Sawin, C.H. Taubes and V.G. Turaev for their many useful comments. Especially we wish to thank D. Kazhdan for his critical reading and P.M. Melvin for suggesting Exercise 3.9 and the use of permanents.

[7] Added in proof: This conjecture was proven in November 1995 by A. Kricher, B. Spence, and I. Aitchison. See their Melbourne University and Queen Marry and Westfield College preprint, *Cabling the Vassiliev Invariants.*

References

[AS1] S. Axelrod, I.M. Singer: Chern–Simons perturbation theory. Proc. XXth DGM Conference (New York, 1991) (S. Catto and A. Rocha, eds) World Scientific, 3–45 1992

[AS2] S. Axelrod, I.M. Singer: Chern–Simons perturbation theory II. J. Differ. Geom. **39**, 173–213 (1994)

[BL] J.S. Birman, X.-S. Lin: Knot polynomials and Vassiliev's invariants. Invent. Math. **111**, 225–270 (1993)

[B-N1] D. Bar-Natan: Perturbative aspects of the Chern–Simons topological quantum field theory. Ph.D thesis, Princeton Univ., June 1991

[B-N2] D. Bar-Natan: On the Vassiliev knot invariants, Topology **34**, 423–472 (1995)

[B-N3] D. Bar-Natan: Non-associative tangles. Georgia Inter. Topology Conference proceedings (to appear)

[Ca] P. Cartier: Construction combinatoire des invariants de Vassiliev–Kontsevich des nœuds, C.R. Acad. Sci. Paris **316**, Série I, 1205–1210 (1993)

[Ch] J. Cheeger: Analytic torsion and the heat equation. Ann. Math. **109**, 259–322 (1979)

[Co] J.H. Conway: An enumeration of knots and links and some of their algebraic properties, in Computational Problems in Abstract Algebra. 329–358, Pergamon, New York, 1970

[CDL1] S.V. Chmutov, S.V. Duzhin, S.K. Lando: Vassiliev knot invariants I. Introduction. Adv. Sov. Math. **21**, 117–126 (1994)

[CDL2] S.V. Chmutov, S.V. Duzhin: Vassiliev knot invariants II. Intersection graph conjecture for trees, Adv. Sov. Math. **21**, 127–134 (1994)

[CDL3] S.V. Chmutov, S.K. Lando: Vassiliev knot invariants III. Forest algebra and weighted graphs. Adv. Sov. Math. **21**, 135–145 (1994)

[Dr1] V.G. Drinfel'd: Quasi-Hopf algebras. Leningrad Math. J. **1**, 1419–1457 (1990)

[Dr2] V.G. Drinfel'd: On quasitriangular Quasi-Hopf algebras and a group closely connected with Gal$(\bar{\mathbb{Q}}/\mathbb{Q})$, Leningrad Math. J. **2**, 829–860 (1991)

[EMSS] S. Elitzur, G. Moore, A. Schwimmer, N. Seiberg: Remarks on the canonical quantization of the Chern–Simons–Witten theory. Nucl. Phys. **B326**, 108–134 (1989)

[GJ] I.P. Goulden, D.M. Jackson: Immanants of combinatorial matrices. J. Algebra **148**, 305–324 (1992)

[Gou] M. Goussarov: A new form of the Conway–Jones polynomial of oriented links, in Topology of manifolds and varieties (O. Viro, editor), Amer. Math. Soc., 167–172, Providence, 1994

[Hu] J.E. Humphreys, Introduction to Lie algebras and representation theory, Springer-Verlag GTM **9**, New York, 1972

[Kas] C. Kassel: Quantum groups. Springer-Verlag GTM **155**, Heidelberg, 1994

[Kau] L.H. Kauffman: On knots, Princeton Univ. Press, Princeton, 1987

[Koh] T. Kohno: Monodromy representaions of braid groups and Yang–Baxter equations. Ann. Inst. Fourier **37**, 139–160 (1987)

[Ko1] M. Konstevich: Vassiliev's knot invariants. Adv. in Sov. Math. **16(2)**, 137–150 (1993)

[Ko2] M. Konstevich: (unpublished)

[LM] T.Q.T. Le, J. Murakami: The universal Vassiliev–Kontsevich invariant for framed oriented links. hep-th/9401016. Compos. Math. (to appear)

[Lin] X.-S. Lin: Vertex models, quantum groups and Vassiliev's knot invariants. Columbia Univ preprint, 1991

[Lit] D.E. Littlewood: The theory of group characters, Clarendon, Oxford 1950

[MM] P.M. Melvin, H.R. Morton: The coloured Jones function. Commun. Math. Phys. **169**, 501–520 (1995)

[Mi] J. Milnor: A duality theorem for Reidemeister torsion. Ann. Math. **76**, 137–147 (1962)

[Mo] H.R. Morton: The coloured Jones function and Alexander polynomial for torus knots. Math. Proc. Camb. Philos. Soc. **117**, 129–136 (1995)

[MS] H.R. Morton, P. Strickland: Jones polynomial invariants of knots and satellites. Math. Proc. Camb. Phil. Soc. **109**, 83–103 (1991)

[Mü] W. Müller: Analytic torsion and the *R*-torsion of Riemannian manifolds. Adv. Math. **28**, 233–305 (1978)

[Pi1] S. Piunikhin: Weights of Feynman diagrams, link polynomials and Vassiliev knot invariants. Moscow State Univ. (preprint, 1992)

[Pi2] S. Piunikhin: Combinatorial expression for universal Vassiliev link invariant. Commun. Math. Phys. **168**, 1–22 (1995)

[RT] N.Yu. Reshetikhin, V.G. Turaev: Ribbon graphs and their invariants derived from quantum groups. Commun. Math. Phys. **127**, 1–26 (1990)

[Ro1] L. Rozansky: A contribution of the trivial flat connection to the Jones polynomial and Witten's invariant of 3d Manifolds. hep-th/9401061 (preprint, January 1994)

[Ro2] L. Rozansky: A contribution of the trivial flat connection to the Jones polynomial and Witten's invariant of 3d Manifolds II. hep-th/9403021 (preprint, March 1994)

[Ro3] L. Rozansky: Reshetikhin's formula for the Jones polynomial of a link: Feynman diagrams and Milnor's linking numbers, hep-th/9403020 (preprint, March 1994)

[St1] J. Stembridge: Some conjectures for immanants. Can. J. Math. **44**, 1079–1099 (1992)

[St2] J. Stembridge: Immanants of totally positive matrices are nonnegative. Bull. London Math. Soc. **23**, 422–428 (1991)

[Tu] V. Turaev: Reidemeister torsion in knot theory. Russ. Math. Surveys **41**, 97–147 (1986)

[Va1] V.A. Vassiliev: Cohomology of knot spaces, Theory of Singularities and its Applications (Providence) (V.I. Arnold, ed.), Amer. Math. Soc., Providence, 1990

[Va2] V.A. Vassiliev: Complements of discriminants of smooth maps: topology and applications, Trans. of Math. Mono. **98**, Amer. Math. Soc., Providence, 1992

[Wi] E. Witten: Quantum field theory and the Jones polynomial. Commun. Math. Phys. **121**, 360–376 (1989)

ISSN 1364-0380 (on line) 1465-3060 (printed)

Geometry & Topology
Volume 9 (2005) 1253–1293
Published: 24 July 2005

The colored Jones function is q–holonomic

Stavros Garoufalidis
Thang T Q Lê

School of Mathematics, Georgia Institute of Technology
Atlanta, GA 30332-0160, USA

Email: `stavros@math.gatech.edu`, `letu@math.gatech.edu`

URL: `http://www.math.gatech.edu/~stavros`

Abstract

A function of several variables is called holonomic if, roughly speaking, it is determined from finitely many of its values via finitely many linear recursion relations with polynomial coefficients. Zeilberger was the first to notice that the abstract notion of holonomicity can be applied to verify, in a systematic and computerized way, combinatorial identities among special functions. Using a general state sum definition of the colored Jones function of a link in 3–space, we prove from first principles that the colored Jones function is a multisum of a q–proper-hypergeometric function, and thus it is q–holonomic. We demonstrate our results by computer calculations.

AMS Classification numbers Primary: 57N10

Secondary: 57M25

Keywords: Holonomic functions, Jones polynomial, Knots, WZ algorithm, quantum invariants, D–modules, multisums, hypergeometric functions

Proposed: Walter Neumann
Seconded: Joan Birman, Vaughan Jones

Received: 28 October 2004
Revised: 20 July 2005

1 Introduction

1.1 Zeilberger meets Jones

The colored Jones function of a framed knot $\mathcal{K}$ in 3–space

$$J_{\mathcal{K}}\colon\ \mathbb{N} \longrightarrow \mathbb{Z}[q^{\pm 1/4}]$$

is a sequence of Laurent polynomials that essentially measures the Jones polynomial of a knot and its cables. This is a powerful but not well understood invariant of knots. As an example, the colored Jones function of the 0–framed right-hand trefoil is given by

$$J_{\mathcal{K}}(n) = \frac{q^{1/2-n/2}}{1-q^{-1}} \sum_{k=0}^{n-1} q^{-kn}(1-q^{-n})(1-q^{1-n})\dots(1-q^{k-n}).$$

Here $J_{\mathcal{K}}(n)$ denotes the Jones polynomial of the 0–framed knot $\mathcal{K}$ colored by the n–dimensional irreducible representation of $\mathfrak{sl}_2$, and normalized by $J_{\mathrm{unknot}}(n) = (q^{n/2} - q^{-n/2})/(q^{1/2} - q^{-1/2})$.

Only a handful of knots have such a simple formula. However, as we shall see all knots have a *multisum* formula. Another way to look at the colored Jones function of the trefoil is via the following 3–term recursion formula:

$$J_{\mathcal{K}}(n) = \frac{q^{n-1} + q^{4-4n} - q^{-n} - q^{1-2n}}{q^{1/2}(q^{n-1} - q^{2-n})} J_{\mathcal{K}}(n-1) + \frac{q^{4-4n} - q^{3-2n}}{q^{2-n} - q^{n-1}} J_{\mathcal{K}}(n-2)$$

with initial conditions: $J_{\mathcal{K}}(0) = 0, \quad J_{\mathcal{K}}(1) = 1$.

In this paper we prove that the colored Jones function of any knot satisfies a linear recursion relation, similar to the above one. For a few knots this was obtained by Gelca and his colleagues [13, 14]. (In [13] a more complicated 5–term recursion formula for the trefoil was established).

Discrete functions that satisfy a nontrivial difference recursion relation are known by another name: they are q–holonomic.

Holonomic functions were introduced by IN Bernstein [2, 3] and M Saito. The latter coined the term holonomic, that is a function which is entirely determined by the law of its differential equation, together with finitely many initial conditions. Bernstein used holonomic functions to prove a conjecture of Gelfand on the analytic continuation of operators. Holonomicity and the related notion of D–modules are a tool in studying linear differential equations from the point of view of algebra (differential Galois theory), algebraic geometry, and category theory. For an excellent introduction on holonomic functions and their properties, see [5] and [7].

Our approach to the colored Jones function owes greatly to Zeilberger's work. Zeilberger noticed that the abstract notion of holonomicity can be applied to verify, in a systematic and computerized way, combinatorial identities among special functions, [35] and also [33, 28].

A starting point for Zeilberger, the so-called *operator approach*, is to replace functions by the recursion relations that they satisfy. This idea leads in a natural way to noncommutative algebras of operators that act on a function, together with left ideals of annihilating operators.

To explain this idea concretely, consider the operators E and Q which act on a *discrete function* (that is, a function of a discrete variable n) $f\colon \mathbb{N} \longrightarrow \mathbb{Z}[q^{\pm}]$ by:

$$(Qf)(n) = q^n f(n) \qquad (Ef)(n) = f(n+1).$$

It is easy to see that $EQ = qQE$, and that E, Q generate a noncommutative q–*Weyl algebra* generated by noncommutative polynomials in E and Q, modulo the relation $EQ = qQE$:

$$\mathcal{A} = \mathbb{Z}[q^{\pm}]\langle Q, E\rangle/(EQ = qQE)$$

Given a discrete function f as above, consider the *recursion ideal* $\mathcal{I}_f = \{P \in \mathcal{A} \,|\, Pf = 0\}$. It is easy to see that it is a left ideal of the q–Weyl algebra. We say that f is *q–holonomic* iff $\mathcal{I}_f \neq 0$.

In this paper we prove that:

Theorem 1 *The colored Jones function of every knot is q–holonomic.*

Theorem 1 and its companion Theorem 2 are effective, as their proof reveals.

Theorem 2

(a) *The E–order of the colored Jones function of a knot is bounded above by an exponential function in the number of crossings.*

(b) *For every knot $\mathcal{K}$ there exist a natural number $n(\mathcal{K})$, such that $n(\mathcal{K})$ initial values of the colored Jones function determine the colored Jones function of $\mathcal{K}$. In other words, the colored Jones function is determined by a finite list. $n(\mathcal{K})$ is bounded above by an exponential function in the number of crossings.*

Computer calculations are given in Section 6. In relation to (b) above, notice that the q–Weyl algebra is *noetherian*; thus every left ideal is finitely generated. The theorem states more, namely that the we can compute (via elimination) a basis for the recursion ideal of the colored Jones function of a knot.

Let us end the introduction with some remarks.

Remark 1.1 The colored Jones function can be defined for every simple Lie algebra $\mathfrak{g}$. Our proof of Theorem 1 generalizes and proves that the $\mathfrak{g}$–colored Jones function of a knot is q–holonomic (except for G_2), see Theorem 6 below.

Remark 1.2 The colored Jones function can be defined for colored links in 3–space. Our proof of Theorem 1 proves that the colored Jones function of a link is q–holonomic in all variables, see Section 3.1.

Remark 1.3 It is well known that computing $J(n)$ for any fixed $n > 1$ is a *#P–complete* problem. Theorem 1 claims that this sequence of #P–complete problems is no worse than any of its terms.

Remark 1.4 The proof of Theorem 1 indicates that many statistical mechanics models, with complicated partition functions that depend on several variables, are holonomic, provided that their local weights are holonomic. This observation may be of interest to statistical mechanics.

1.2 Synonymous notions to holonomicity

We have chosen to phrase the results of our paper mostly using the high-school language of linear recursion relations. We could have used synonymous terms such as linear q–difference equations, or q–holonomic functions, or D–modules, or maximally overdetermined systems of linear PDEs which is more common in the area of algebraic analysis, see for example [24]. The geometric notion of D–modules gives rise to geometric invariants of knots, such as the characteristic variety introduced by the first author in [11]. The characteristic variety is determined by the colored Jones function of a knot and is conjectured to be isomorphic to the $\mathfrak{sl}_2(\mathbb{C})$–character variety of a knot, viewed from the boundary torus. This, so-called *AJ Conjecture*, formulated by the first author is known to hold for all torus knots (due to Hikami, [19]), and infinitely many 2–bridge knots (due to the second-author, [21]).

Thus, there is nontrivial geometry encoded in the linear recursion relations of the colored Jones function of a knot.

1.3 Plan of the proof

In Section 2, we discuss in detail the notion of a q–holonomic function. We give examples of q–holonomic functions (our building blocks), together with rules that create q–holonomic functions from known ones.

In Section 3, we discuss the colored Jones function of a link in 3–space, using state sums associated to a planar projection of the link. The colored Jones function is built out of local building blocks (namely, R–matrices) associated to the crossings, which are assembled together in a way dictated by the planar projection. The main observation is that the R–matrix is q–holonomic in all variables, and that the assembly preserves q–holonomicity. Theorem 1 follows. As a bonus, we present the colored Jones function as a multisum of a q–proper hypergeometric function.

In Section 4 we show that the cyclotomic function of a knot (a reparametrization of the colored Jones function, introduced by Habiro, with good integrality properties) is q–holonomic, too. We achieve this by studying explicitly a change of basis for representations of $\mathfrak{sl}_2$.

In Section 5 we give a theoretical review about complexity and computability of recursion relations of q–holonomic functions, following Zeilberger. These ideas solve the problem of finding recursion relations of q–holonomic functions which are given by multisums of q proper hypergeometric functions. It is a fortunate coincidence (?) that the colored Jones function can be presented by such a multisum, thus we can compute its recursion relations. Theorem 2 follows.

Section 6 is a computer implementation of the previous section, where we use `Mathematica` packages developed by A Riese.

In Section 7 we discuss the $\mathfrak{g}$–colored Jones function of a knot, associated to a simple Lie algebra $\mathfrak{g}$. Our goal is to prove that the $\mathfrak{g}$–colored Jones function is q–holonomic in all variables (see Theorem 6). In analogy with the $\mathfrak{g} = \mathfrak{sl}_2$ case, we need to show that the local building block, the R–matrix, is q–holonomic in all variables. This is a trip to the world of quantum groups, which takes up the rest of the section, and ends with an appendix which computes (by brute-force) structure constants of quantized enveloping Lie algebras in the rank 2 case.

1.4 Acknowledgements

The authors wish to express their gratitude to D Zeilberger who introduced them to the wonderful world of holonomic functions and its relation to discrete mathematics. Zeilberger is a philosophical co-author of the present paper.

In addition, the authors wish to thank A Bernstein, G Lusztig, and N Xie for help in understanding quantum group theory, G Masbaum and K Habiro for help in a formula of the cyclotomic expansion of the colored Jones polynomial, and D Bar-Natan and A Riese for computer implementations.

The authors were supported in part by National Science Foundation.

2 q–holonomic and q–hypergeometric functions

Theorem 1 follows from the fact that the colored Jones function can be built from elementary blocks that are q–holonomic, and the operations that patch the blocks together to give the colored Jones function preserve q–holonomicity.

IN Bernstein defined the notion of holonomic functions $f\colon \mathbb{R}^r \longrightarrow \mathbb{C}$, [2, 3]. For an excellent and complete account, see Bjork [4]. Zeilberger's brilliant idea was to link the abstract notion of holonomicity to the concrete problem of algorithmically proving combinatorial identities among hypergeometric functions, see [35, 33] and also [28]. This opened an entirely new view on combinatorial identities.

Sabbah extended Bernstein's approach to holonomic functions and defined the notion of a q–holonomic function, see [31] and also [6].

2.1 q–holonomicity in many variables

We briefly review here the definition of q–holonomicity. First of all, we need an r–dimensional version of the q–Weyl algebra. Consider the operators E_i and Q_j for $1 \leq i, j \leq r$ which act on discrete functions $f\colon \mathbb{N}^r \longrightarrow \mathbb{Z}[q^{\pm}]$ by:

$$\begin{aligned}(Q_i f)(n_1, \dots, n_r) &= q^{n_i} f(n_1, \dots, n_r)\\ (E_i f)(n_1, \dots, n_r) &= f(n_1, \dots, n_{i-1}, n_i + 1, n_{i+1}, \dots, n_r).\end{aligned}$$

It is easy to see that the following relations hold:

$$\begin{array}{ll} Q_i Q_j = Q_j Q_i & E_i E_j = E_j E_i \\ Q_i E_j = E_j Q_i \,\text{for}\, i \neq j & E_i Q_i = q Q_i E_i \end{array} \tag{Rel_q}$$

We define the *q–Weyl algebra* $\mathcal{A}_r$ to be a noncommutative algebra with presentation

$$\mathcal{A}_r = \frac{\mathbb{Z}[q^{\pm 1}]\langle Q_1, \dots, Q_r, E_1, \dots, E_r \rangle}{(\mathrm{Rel}_q)}.$$

Given a discrete function f with domain $\mathbb{N}^r$ or $\mathbb{Z}^r$ and target space a $\mathbb{Z}[q^{\pm 1}]$–module, one can define the left ideal $\mathcal{I}_f$ in $\mathcal{A}_r$ by

$$\mathcal{I}_f := \{P \in \mathcal{A}_r | Pf = 0\}.$$

If we want to determine a function f by a finite list of initial conditions, it does not suffice to ensure that f satisfies one nontrivial recursion relation if $r \geq 2$. The key notion that we need instead is q–holonomicity.

Intuitively, a discrete function $f\colon \mathbb{N}^r \longrightarrow \mathbb{Z}[q^{\pm}]$ is q–holonomic if it satisfies a *maximally overdetermined system* of linear difference equations with polynomial coefficients. The exact definition of holonomicity is through homological dimension, as follows.

Suppose $M = \mathcal{A}_r/I$, where I is a left $\mathcal{A}_r$–module. Let F_m be the sub-space of $\mathcal{A}_r$ spanned by polynomials in Q_i, E_i of total degree $\leq m$. Then the module $\mathcal{A}_r/I$ can be approximated by the sequence $F_m/(F_m \cap I), m = 1, 2, \ldots.$ It turns out that, for $m >> 1$, the dimension (over the fractional field $\mathbb{Q}(q)$) of $F_m/(F_m \cap I)$ is a polynomial in m whose degree $d(M)$ is called the *homological dimension* of M.

Bernstein's *famous inequality* (proved by Sabbah in the q–case, [31]) states that $d(M) \geq r$, if $M \neq 0$ and M has *no monomial torsions*, ie, any non-trivial element of M cannot be annihilated by a monomial in Q_i, E_i. Note that the left $\mathcal{A}_r$–module $M_f := \mathcal{A}_r \cdot f \cong \mathcal{A}_r/\mathcal{I}_f$ does not have monomial torsion.

Definition 2.1 We say that a discrete function f is q–holonomic if $d(M_f) \leq r$.

Note that if $d(M_f) \leq r$, then by Bernstein's inequality, either $M_f = 0$ or $d(M_f) = r$. The former can happen only if $f = 0$.

Although we will not use in this paper, let us point out an alternative cohomological definition of dimension for a finitely generated $\mathcal{A}_r$ module M. Let us define

$$c(M) := \min\{j \in \mathbb{N} \mid \mathrm{Ext}^j_{\mathcal{A}_r}(M, \mathcal{A}_r) \neq 0\}.$$

Then the *homological dimension* $d(M) := 2r - c(M)$ equals to the dimension $d(M)$ defined above.

Closely related to $\mathcal{A}_r$ is the *q–torus algebra* $\mathcal{T}_r$ with presentation

$$\mathcal{T}_r = \frac{\mathbb{Z}[q^{\pm 1}]\langle Q_1^{\pm 1}, \ldots, Q_r^{\pm 1}, E_1^{\pm 1}, \ldots, E_r^{\pm 1}\rangle}{(\mathrm{Rel}_q)}.$$

Elements of $\mathcal{T}_r$ acts on the set of functions with domain $\mathbb{Z}^r$, but not on the set of functions with domain $\mathbb{N}^r$. Note that $\mathcal{T}_r$ is simple, but $\mathcal{A}_r$ is not. If I is a left ideal of $\mathcal{T}_r$ then the dimension of $\mathcal{T}_r/I$ is equal to that of $\mathcal{A}_r/(I \cap A_r)$.

2.2 Assembling q–holonomic functions

Despite the unwelcoming definition of q–holonomic functions, in this paper we will use not the definition itself, but rather the *closure properties* of the set of q–holonomic functions under some natural operations.

Fact 0

- Sums and products of q–holonomic functions are q–holonomic.
- Specializations and extensions of q–holonomic functions are q–holonomic. In other words, if $f(n_1,\dots,n_m)$ is q–holonomic, the so are the functions
$$\begin{aligned} g(n_2,\dots,n_m) &:= f(a,n_2,\dots,n_m) \\ \text{and} \qquad h(n_1,\dots,n_m,n_{m+1}) &:= f(n_1,\dots,n_m). \end{aligned}$$
- Diagonals of q–holonomic functions are q–holonomic. In other words, if $f(n_1,\dots,n_m)$ is q–holonomic, then so is the function
$$g(n_2,\dots,n_m) := f(n_2,n_2,n_3,\dots,n_m).$$
- Linear substitution. If $f(n_1,\dots,n_m)$ is q–holonomic, then so is the function, $g(n'_1,\dots,n'_{m'})$, where each n'_j is a linear function of n_i.
- Multisums of q–holonomic functions are q–holonomic. In other words, if $f(n_1,\dots,n_m)$ is q–holonomic, the so are the functions g and h, defined by
$$\begin{aligned} g(a,b,n_2,\dots,n_m) &:= \sum_{n_1=a}^{b} f(n_1,n_2,\dots,n_m) \\ h(a,n_2,\dots,n_m) &:= \sum_{n_1=a}^{\infty} f(n_1,n_2,\dots,n_m) \end{aligned}$$
(assuming that the latter sum is finite for each a).

For a user-friendly explanation of these facts and for many examples, see [35, 33] and [28].

2.3 Examples of q–holonomic functions

Here are a few examples of q–holonomic functions. In fact, we will encounter only sums, products, extensions, specializations, diagonals, and multisums of these functions. In what follows we usually extend the ground ring $\mathbb{Z}[q^{\pm 1}]$ to the *fractional field* $\mathbb{Q}(q^{1/D})$, where D is a positive integer. We also use v to denote a root of q, $v^2 = q$.

For $n, k \in \mathbb{Z}$, let

$$\{n\} := v^n - v^{-n}, \qquad [n] := \frac{\{n\}}{\{1\}}, \qquad [n]! := \prod_{i=1}^{n} [i], \qquad \{n\}! := \prod_{i=1}^{n} \{i\}$$

$$\{n\}_k := \begin{cases} \prod_{i=1}^{k} \{n-i+1\}, & \text{if } k \geq 0 \\ 0 & \text{if } k < 0 \end{cases}$$

$$\left[\begin{array}{c} n \\ k \end{array} \right] := \begin{cases} \frac{\{n\}_k}{\{k\}_k} & \text{if } k \geq 0 \\ 0 & \text{if } k < 0 \end{cases}.$$

The first four functions are q–holonomic in n, and the last two, as well as the delta function $\delta_{n,k}$, are q–holonomic in both n and k.

2.4 q–hypergeometric functions

Definition 2.2 A discrete function $f\colon \mathbb{Z}^r \longrightarrow \mathbb{Q}(q)$ is *q–hypergeometric* iff $E_i f/f \in \mathbb{Q}(q, q^{n_1}, \dots, q^{n_r})$ for all $i = 1, \dots, r$.

In that case, we know generators for the annihilation ideal of f. Namely, let $E_i f/f = (R_i/S_i)|_{Q_i = q^{n_i}}$ for $R_i, S_i \in \mathbb{Z}[q, Q_1, \dots, Q_r]$. Then, the annihilation ideal of f is generated by $S_i E_i - R_i$.

All the functions in the previous subsections are q–hypergeometric.

Unfortunately, q–hypergeometric functions are not always q–holonomic. For example, $(n,k) \longrightarrow 1/[n^2 + k^2]!$ is q–hypergeometric but not q–holonomic. However, q–proper-hypergeometric functions are q–holonomic. The latter were defined by Wilf–Zeilberger as follows, [33, Sec.3.1]:

Definition 2.3 A *proper q–hypergeometric* discrete function is one of the form

$$F(n, \mathbf{k}) = \frac{\prod_s (A_s; q)_{a_s n + \mathbf{b}_s . \mathbf{k} + c_s}}{\prod_t (B_t; q)_{u_t n + \mathbf{v}_t . \mathbf{k} + w_t}} q^{A(n,\mathbf{k})} \xi^{\mathbf{k}} \tag{1}$$

where $A_s, B_t \in \mathbb{K} = \mathbb{Q}(q)$, a_s, u_t are integers, $\mathbf{b}_s, \mathbf{k}_s$ are vectors of r integers, $A(n, \mathbf{k})$ is a quadratic form, c_s, w_s are variables and ξ is an r vector of elements in $\mathbb{K}$. Here, as usual

$$(A; q)_n := \prod_{i=0}^{n-1} (1 - Aq^i).$$

3 The colored Jones function for $\mathfrak{sl}_2$

3.1 Proof of Theorem 1 for links

We will formulate and prove an analog of Theorem 1 (see Theorem 3 below) for colored links. Our proof will use a *state-sum* definition of the colored Jones function, coming from a representation of the quantum group $U_q(\mathfrak{sl}_2)$, as was discovered by Reshetikhin and Turaev in [29, 32].

Suppose L is a framed, oriented link of p components. Then the colored Jones function $J_L\colon \mathbb{N}^p \to \mathbb{Z}[q^{\pm 1/4}] = \mathbb{Z}[v^{\pm 1/2}]$ can be defined using the representations of braid groups coming from the quantum group $U_q(\mathfrak{sl}_2)$.

Theorem 3 *The colored Jones function J_L is q–holonomic.*

Proof We will present the definition of J_L in the form most suitable for us. Let $V(n)$ be the n–dimensional vector space over the field $\mathbb{Q}(v^{1/2})$ with basis $\{e_0, e_1, \dots, e_{n-1}\}$, with $V(0)$ the zero vector space.

Fix a positive integer m. A linear operator

$$A\colon\ V(n_1)\otimes\cdots\otimes V(n_m) \to V(n'_1)\otimes\cdots\otimes V(n'_m)$$

can be described by the collection

$$A^{b_1,\dots,b_m}_{a_1,\dots,a_m} \in \mathbb{Q}(v^{1/2}),$$

where

$$A(e_{a_1}\otimes\cdots\otimes e_{a_m}) = \sum_{b_1<n'_1,\dots,b_m<n'_m} A^{b_1,\dots,b_m}_{a_1,\dots,a_m}\, e_{b_1}\otimes\cdots\otimes e_{j_m}.$$

We will call $(a_1,\dots,a_m,b_1,\dots,b_m)$ *the coordinates* of the matrix entry $A^{b_1,\dots,b_m}_{a_1,\dots,a_m}$ of A, with respect to the given basis.

The *building block* of our construction is a pair of functions $f_\pm\colon \mathbb{Z}^5 \to \mathbb{Z}[v^{\pm 1/2}]$, given by

$$f_+(n_1,n_2;a,b,k)$$
$$:= (-1)^k v^{-((n_1-1-2a)(n_2-1-2b)+k(k-1))/2} \begin{bmatrix} b+k \\ k \end{bmatrix} \{n_1-1+k-a\}_k,$$

$$f_-(n_1,n_2;a,b,k)$$
$$:= v^{((n_1-1-2a-2k)(n_2-1-2b+2k)+k(k-1))/2} \begin{bmatrix} a+k \\ k \end{bmatrix} \{n_2-1+k-b\}_k.$$

The reader should not focus on the actual, cumbersome formulas. The main point is that:

Fact 1

- f_+ and f_- are q–proper hypergeometric and thus q–holonomic in all variables.

For each pair $(n_1, n_2) \in \mathbb{N}^2$ we define two operators

$$\mathcal{B}_+(n_1, n_2), \mathcal{B}_-(n_1, n_2)\colon\ V(n_1) \otimes V(n_2) \to V(n_2) \otimes V(n_1)$$

by

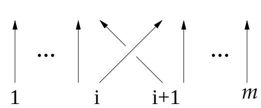
$$\begin{aligned}
(\mathcal{B}_+(n_1, n_2))_{a,b}^{c,d} &:= f_+(n_1, n_2; a, b, c-b)\, \delta_{c-b,a-d}, \\
(\mathcal{B}_-(n_1, n_2))_{a,b}^{c,d} &:= f_+(n_1, n_2; a, b, b-c)\, \delta_{c-b,a-d},
\end{aligned}$$

where $\delta_{x,y}$ is Kronecker's delta function. Although the coordinates (a, b, c, d) of the entry $\mathcal{B}_\pm(n_1, n_2))_{a,b}^{c,d}$ of the operators $\mathcal{B}_\pm(n_1, n_2)$ are defined for $0 \leq a, b \leq n_1$ and $0 \leq c, d \leq n_2$, the above formula makes sense for all non-negative integers a, b, c, d. This will be important for us. The following lemma is obvious.

Lemma 3.1 *The discrete functions $\mathcal{B}_\pm(n_1, n_2)_{a,b}^{c,d}$ are q–holonomic with respect to the variables (n_1, n_2, a, b, c, d).*

If we identify $V(n)$ with the simple n–dimensional $U_q(\mathfrak{sl}_2)$–module, with $e_i, i = 0, \dots, n-1$ being the *standard basis*, then $\mathcal{B}_+(n_1, n_2), \mathcal{B}_-(n_1, n_2)$ are respectively the braiding operator and its inverse acting on $V(n_1) \otimes V(n_2)$. This fact follows from the formula of the R–matrix, say, in [17, Chapter 3]. In particular, $\mathcal{B}_-(n_1, n_2)$ is the inverse of $\mathcal{B}_+(n_1, n_2)$. If one allows a, b, c, d in $\mathcal{B}_\pm(n_1, n_2)_{a,b}^{c,d}$ to run the set $\mathbb{N}$, then $\mathcal{B}_\pm(n_1, n_2)_{a,b}^{c,d}$ define the braid action on the Verma module corresponding to $V(n_1), V(n_2)$.

Let B_m be the braid group on m strands, with standard generators $\sigma_1, ..., \sigma_{m-1}$:

$\sigma_i =$ 

For each braid $\beta \in B_m$ and $(n_1, \dots, n_m) \in \mathbb{N}^m$, we will define an operator $\tau(\beta) = \tau(\beta)(n_1, \dots, n_m)$,

$$\tau(\beta)\colon\ V(n_1) \otimes \cdots \otimes V(n_m) \to V(n_{\bar{\beta}(1)}) \otimes \cdots \otimes V(n_{\bar{\beta}(m)}),$$

where $\bar{\beta}$ is the permutation of $\{1,\dots,m\}$ corresponding to β. The operator $\tau(\beta)$ is uniquely determined by the following properties: For an elementary braid σ_i, we have:

$$\tau(\sigma_i^{\pm 1}) = \mathrm{id}^{\otimes i-1} \otimes \mathcal{B}_{\pm}(n_i, n_{i+1})\,\mathrm{id}^{\otimes m-i-1}.$$

In addition, if $\beta = \beta'\beta''$, then $\tau(\beta) := \tau(\beta')\tau(\beta'')$. It is well-known that $\tau(\beta)$ is well-defined.

From Fact 0 and Lemma 3.1 it follows that

Lemma 3.2 *For any braid $\beta \in B_m$, the discrete function $\tau(\beta)(n_1,\dots,n_m)$, considered as a function with variables $n_1,\dots,n_m$ and all the coordinates of the matrix entry, is q–holonomic.*

Let K be the linear endomorphism of $V(n_1)\otimes\cdots\otimes V(n_m)$ defined by

$$K(e_{i_1}\otimes\cdots\otimes e_{i_m}) = v^{n_1+\cdots+n_m-2i_1-\cdots-2i_m-m}\, e_{i_1}\otimes\cdots\otimes e_{i_m}.$$

The inverse operator K^{-1} is well-defined.

Corollary 3.3 *For any braid $\beta \in B_m$, the discrete function*

$$\tilde{\tau}(\beta) := \tau(\beta)(n_1,\dots,n_m)\times K^{-1}$$

is q–holonomic in $n_1,\dots,n_m$ all all of the coordinates of the matrix entry.

In general, the trace of $\tilde{\tau}(\beta)$ is called the *quantum trace* of $\tau(\beta)$. Although the target space and source space maybe different, let us define the quantum trace of $\tau(\beta)(n_1,\dots,n_m))$ by

$$\mathrm{tr}_q(\beta)(n_1,\dots,n_m) := \sum_{1\le i\le m}\ \sum_{0\le a_i<n_i} \tilde{\tau}(\beta)(n_1,\dots,n_m)^{a_1,\dots,a_m}_{a_1,\dots,a_m}.$$

It follows from Fact 0 that $\mathrm{tr}_q(\beta)(n_1,\dots,n_m)$ is q–holonomic in $n_1,\dots,n_m$. Restricting this function on the diagonal defined by $n_i = n_{\bar{\beta}i}, i = 1,\dots,m$, we get a new function J_β of p variables, where p is the number of cycles of the permutation $\bar{\beta}$.

Suppose a framed link L can be obtained by closing the braid β. Then the colored Jones polynomial J_L is exactly J_β. Hence Theorem 1 follows. $\square$

Remark 3.4 In general, $J_{\mathcal{K}}(n)$ contains the fractional power $q^{1/4}$. If K has framing 0, then $J_{\mathcal{K}'}(n) := J_{\mathcal{K}}(n)/[n] \in \mathbb{Z}[q^{\pm 1}]$. See [20].

Remark 3.5 There is a variant of the colored Jones function $J_{L'}$ of a colored link L' where one of the components is broken. If β is a braid as above, let us define the *broken quantum trace* tr'_β by

$$\mathrm{tr}'_q(\beta)(n_1,\dots,n_m) := \sum_{2\le i\le m}\sum_{0\le a_i<n_i} \tilde\tau(\beta)(n_1,\dots,n_m)^{a_1,\dots,a_m}_{a_1,\dots,a_m}|_{a_1=0}.$$

Restricting this function on the diagonal defined by $n_i = n_{\bar\beta i}, i = 1,\dots,m$, we get a new function $J_{\beta'}$ of p variables, where p is the number of cycles of the permutation $\bar\beta$.

If L' denotes the broken link which is the closure of all but the first strand of β, then the colored Jones function $J_{L'}$ of L' satisfies $J_{L'} = J_{\beta'}$.

If L denotes the closure of the broken link L', then we have:

$$J_L = J_{L'} \times [\lambda]$$

where λ is the color of the broken component of L'.

3.2 A multisum formula for the colored Jones function of a knot

In this section we will give explicit multisum formulas for the $\mathfrak{sl}_2$–colored Jones function of a knot. The calculation here is computerized in Section 6.

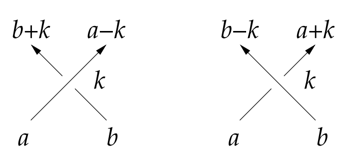
Consider a word $w = \sigma_{i_1}^{\epsilon_1}\dots\sigma_{i_c}^{\epsilon_c}$ of length m written in the standard generators $\sigma_1,\dots,\sigma_{s-1}$ of the braid group B_m with m strands, where $\epsilon_i = \pm 1$ for all i.

w gives rise to a braid $\beta\in B_m$, and we assume that the closure of β is a knot $\mathcal{K}$. Let $\mathcal{K}'$ denote *long knot* which is the closure of all but the first strand of β.

A *coloring* of $\mathcal{K}'$ is a tuple $\mathbf{k} = (k_1,\dots,k_c)$ of *angle variables* placed at the crossings of $\mathcal{K}'$.

Lemma 3.6 *There is a unique way to extend a coloring* $\mathbf{k}$ *of* $\mathcal{K}'$ *to a coloring of the crossings and part-arcs of* $\mathcal{K}'$ *such that*

- *around each crossing the following consistency relations are satisfied:*



- *The color of the lower-left incoming part-arc is* 0.

Moreover, the labels of the part-arcs are linear forms on $\mathbf{k}$*.*

Proof Start walking along the long knot starting at the incoming part-arc. At the first crossing, whether over or under, the label of the outgoing part-arc is determined by the label of the ending part-arc and the angle variable of the crossing. Thus, we know the label of the outgoing part-arc of the first crossing. Keep going. Since $\mathcal{K}'$ is topologically an interval, the result follows. □

For an example, see Figure 1.

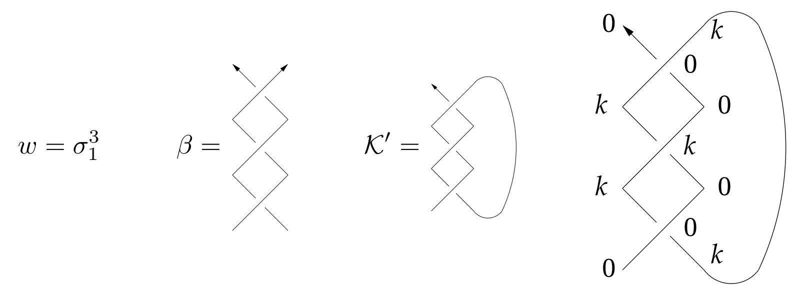


Figure 1: A word w, the corresponding braid β, its long closure $\mathcal{K}'$, and a coloring of $\mathcal{K}'$

Fix a coloring of $\mathcal{K}'$ determined by a vector $\mathbf{k}$. Let $b_i(\mathbf{k})$ for $i = 1, \ldots, m$ denote the labels of the top part-arcs of β. Let $x_j(\mathbf{k})$ and $y_j(\mathbf{k})$ denote the labeling of the left and right incoming part-arcs at the ith crossing of $\mathcal{K}'$ for $j = 1, \ldots, c$. According to Lemma 3.6, $b_i(\mathbf{k}), x_j(\mathbf{k})$ and $y_j(\mathbf{k})$ are linear forms on $\mathbf{k}$.

It is easy to see that

$$\mathrm{tr}'_q(\beta) = \sum_{\mathbf{k} \geq \mathbf{0}} F_w(n, \mathbf{k})$$

where

$$F_w(n, \mathbf{k}) := \prod_{i=2}^{m} v^{\frac{n}{2} - b_i(\mathbf{k})} \prod_{j=1}^{c} f_{\mathrm{sgn}(\epsilon_i)}(n, n; x_j(\mathbf{k}), y_j(\mathbf{k})).$$

is a q–proper hypergeometric function. Remark 3.5 then implies that

Proposition 3.7 *The colored Jones function of a long knot* $\mathcal{K}'$ *is a multisum of a* q*–proper hypergeometric function:*

$$J_{\mathcal{K}'}(n) = \sum_{\mathbf{k} \geq \mathbf{0}} F_w(n, \mathbf{k}).$$

Remark 3.8 If a long knot $\mathcal{K}'$ is presented by a planar projection D with c crossings (which is not necessarily the closure of a braid), then similar to the above there is a q–proper hypergeometric function $F_D(n,\mathbf{k})$ of $c+1$ variables such that $J_{\mathcal{K}'}(n) = \sum_{\mathbf{k}\geq\mathbf{0}} F_D(n,\mathbf{k})$. Of course, F_D depends on the planar projection. Occasionally, some of the summation variables can be ignored. This is the case for the right hand-trefoil (where the multisum reduces to a single sum) and the figure eight (where it reduces to a double sum).

D Bar-Natan has kindly provided us with a computerized version of Proposition 3.7, [1].

4 The cyclotomic function of a knot is q–holonomic

Habiro [15] proved that the colored Jones polynomial (of $\mathfrak{sl}_2$) can be rearranged in the following convenient form, known as the *cyclotomic expansion* of the colored Jones polynomial: For every 0–framed knot $\mathcal{K}$, there exists a function

$$C_{\mathcal{K}}\colon \mathbb{Z}_{>0} \to \mathbb{Z}[q^{\pm 1}]$$

such that
$$J_{\mathcal{K}}(n) = \sum_{k=1}^{\infty} C_{\mathcal{K}}(k) S(n,k),$$

where
$$S(n,k) := \{n+k-1\}_{2k-1}/(v-v^{-1}) = \frac{\prod_{n-k+1}^{n+k-1}(v^i - v^{-i})}{v-v^{-1}}.$$

Note that $S(n,k)$ does not depend on the knot $\mathcal{K}$. Note that J is determined from C and vice-versa by an upper diagonal matrix, thus C takes values in $\mathbb{Q}(q)$. The difficult part of Habiro's result is $C_{\mathcal{K}}$ takes values in $\mathbb{Z}[q^{\pm}]$. The integrality of the cyclotomic function is a crucial ingredient in the study of integrality properties of 3–manifold invariants, [15].

Theorem 4 *The cyclotomic function $C_{\mathcal{K}}\colon \mathbb{N} \to \mathbb{Z}[q^{\pm}]$ of every knot $\mathcal{K}$ is q–holonomic.*

Proof Habiro showed that $C_{\mathcal{K}}(n)$ is the quantum invariant of the knot $\mathcal{K}$ with color

$$P''(n) := \frac{\prod_{i=1}^{n-1}(V(2) - v^{2i-1} - v^{1-2i})}{\{2n-1\}_{2n-2}},$$

where $V(n)$ is the unique n–dimensional simple $\mathfrak{sl}_2$–module, and (retaining Habiro's notation with a shift $n \to n-1$) $P''(n)$ is considered as an element of the ring of $\mathfrak{sl}_2$–modules over $\mathbb{Q}(v)$.

Using induction one can easily prove that

$$P''(n) = \sum_{k=1}^{n} R(n,k)V(k),$$

where $R(k,n)$ is given by

$$R(n,k) = (-1)^{n-k} \frac{\{2k\}}{\{2n-1\}![2n]} \left[\begin{array}{c} 2n \\ n-k \end{array} \right].$$

We learned this formula from Habiro [15] and Masbaum [25]. Since

$$C_{\mathcal{K}}(n) = \sum_{k} R(n,k) J_{\mathcal{K}}(k)$$

and $R(n,k)$ is q–proper hypergeometric and thus q–holonomic in both variables n and k, it follows that $C_{\mathcal{K}}$ is q–holonomic. □

5 Complexity

In this section we show that Theorem 1 is effective. In other words, we give a priori bounds and computations that appear in Theorem 2.

5.1 Finding a recursion relation for multisums

Our starting point are multisums of q–proper hypergeometric functions. Recall the definition 2.3 of a q–proper hypergeometric function $F(n,\mathbf{k})$ from Section 2.4, and let G denote

$$G(n) := \sum_{\mathbf{k} \geq \mathbf{0}} F(n,\mathbf{k})$$

throughout this section.

With the notation of Equation (1), Wilf–Zeilberger show that:

Theorem 5 ([33, Sec.5.2])

(a) *$F(n,\mathbf{k})$ satisfies a k–free recurrence relation of order at most*

$$J^\star := \frac{(4STB^2)^r}{r!}$$

where $B = \max_{s,t}\{|\mathbf{b}_s|, |\mathbf{v}_t|, |a_s|, |u_t|\} + \max_{\mu,\nu} |a_{\mu,\nu}|$ where $a_{\mu,\nu}$ are the coefficients of the quadratic form A.

(b) *Moreover, $G(n)$ satisfies an inhomogeneous recursion relation of order at most $J^\star$.*

Let us briefly comment on the proof of this theorem. A *certificate* is an operator of the form

$$P(E,Q) + \sum_{i=1}^{r} (E_i - 1) R_i(E, E_1, \dots, E_r, Q, Q_1, \dots, Q_r)$$

that annihilates $F(n,\mathbf{k})$, where P and R_i are operators with P a polynomial in E, Q, with $P \neq 0$. Here E is the shift operator on n, E_i (for $i = 1, \dots, r$) are shift operators in k_i, and Q is the multiplication operator by q^k and Q_i (for $i = 1, \dots, r$) are the multiplication operator by q^{k_i}, where $\mathbf{k} = (k_1, \dots, k_r)$.

The important thing is that $P(E,Q)$ is an operator that does not depend on the summation variables $\mathbf{k}$. A certificate implies that for all $(n,\mathbf{k})$ we have:

$$P(E,Q)F(n,\mathbf{k}) + \sum_{i=1}^{r} (G_i(n, k_1, \dots, k_{i-1}, k_i + 1, k_{i+1}, \dots, k_r) - G_i(n, k_1, \dots, k_{i-1}, k_i, k_{i+1}, \dots, k_r)) = 0,$$

where $G_i(n,\mathbf{k}) = R_i F(n,\mathbf{k})$. Summing over $\mathbf{k} \geq \mathbf{0}$, it follows that $G(n)$ satisfies an inhomogeneous recursion relation $PG = \mathrm{error}(n)$. Here $\mathrm{error}(n)$ is a sum of multisums of q–proper hypergeometric functions of one variables less. Iterating the process, we finally arrive at a homogeneous recursion relation for G.

How can one find a certificate given $F(n,\mathbf{k})$? Suppose that F satisfies a $\mathbf{k}$–*free* recursion relation $AF = 0$, where $A = A(E, Q, E_1, \dots, E_r)$ is an operator that does not depend on the Q_i. Then, evaluating A at $E_1 = \dots E_r = 1$, we obtain that

$$A = A(E, Q, 1, \dots, 1) + \sum_{i=1}^{r} (E_i - 1) R_i(E, Q, E_1, \dots, E_r)$$

is a certificate.

How can we find a $\mathbf{k}$–free recursion relation for F? Let us write

$$A = \sum_{(i,\boldsymbol{j}) \in S} \sigma_{i,\boldsymbol{j}}(Q) E^i E^{\boldsymbol{j}}$$

where S is a finite set, $\boldsymbol{j} = (j_1, \dots, j_r)$, $E^{\boldsymbol{j}} = E_1^{j_1} \dots E_r^{j_r}$, and $\sigma_{i,\boldsymbol{j}}(Q)$ are polynomial functions in Q with coefficients in $\mathbb{Q}(q)$; see [30]. The condition

$AF = 0$ is equivalent to the equation $(AF)/F = 0$. Since F is q–proper hypergeometric, the latter equation is the vanishing of a rational function in $Q_1, \dots, Q_r$. By cleaning out denominators, this is equivalent to a system of *linear equations* (namely, the coefficients of monomials in Q_i are zero), with unknowns the polynomial functions $\sigma_{i,\boldsymbol{j}}$. For a careful discussion, see [30]. As long as there are more unknowns than equations, the system is guaranteed to have a solution. [33] estimate the number of equations and unknowns in terms of $F(n, \mathbf{k})$, and prove Theorem 5.

Wilf–Zeilberger programmed the above proof, see [28]. As time passes the algorithms get faster and more refined. For the state-of-the-art algorithms and implementations, see [26, 27] and [30], which we will use below.

Alternative algorithms of noncommutative elimination, using *noncommutative Gröbner basis*, have been developed by Chyzak and Salvy, [8]. In order for have Gröbner basis, one needs to use the following localization of the q–Weyl algebra

$$\mathcal{B}_r = \frac{Q(q, Q_1, \dots, Q_r)\langle E_1, \dots, E_r\rangle}{(\mathrm{Rel}_q)}.$$

and *Gröbner basis* [8].

In case $r = 1$, $\mathcal{B}_1$ is a *principal ideal domain* [7, Chapter 2, Exercise 4.5]. In that case one can associate an operator in $\mathcal{B}_1$ (unique up to units) that generates that annihilating ideal of $G(n)$. For a conjectural relation between this operator for the $\mathfrak{sl}_2$–colored Jones function of a knot and hyperbolic geometry, see [11].

Let us point out however that none of the above algorithms can find generators for the annihilating ideal of the multisum $G(n)$. In fact, it is an open problem how to find generators for the annihilating ideal of $G(n)$ in terms of generators for the annihilating ideal of $F(n, \mathbf{k})$, in theory or in practice. We thank M Kashiwara for pointing this out to us.

5.2 Upper bounds for initial conditions

In another direction, one may ask the following question: if a q–holonomic function satisfies a nontrivial recursion relation, it follows that it is uniquely determined by a finite number of initial conditions. How many? This was answered by Yen, [34]. If G is a discrete function which satisfies a recursion relation of order $J^\star$, consider its principal symbol $\sigma(q, Q)$, that is the coefficient of the leading E–term. The principal symbol lies in the commutative ring $\mathbb{Z}[q^\pm, Q^\pm]$ of Laurrent polynomials in two variables q and Q. For every n, consider the Laurrent polynomial $\sigma(q, q^n) \in \mathbb{Z}[q^\pm]$. If $\sigma(q, q^n) \neq 0$ for all n, then G is determined by $J^\star$ many initial values. Since $\sigma(q,Q) \neq 0$, it follows that $\sigma(q,q^n) \neq 0$ for large enough n. In fact, in [34, Prop.3.1] Yen proves that $\sigma(q,q^n) \neq 0$ if $n > \deg_q(\sigma)$, then $\sigma(q,q^n) \neq 0$, where the degree of a Laurrent polynomial in q is the difference between the largest and smallest exponent. Thus, G is determined by $J^{\star\star} := J^\star + \deg_q(\sigma)$ initial conditions.

Yen further gives upper bounds for $\deg_q(\sigma)$ in terms of the q–hypergeometric summand, see [34, Thm.2.9] for single sums. An extension of Yen's work to multisums, gives a priori upper bounds $J^{\star\star}$ in terms of the q–hypergeometric summand. These exponential bounds are of theoretical interest only, and in practice much smaller bounds are found by computer.

5.3 Proof of Theorem 2

Theorem 2 follows from Proposition 3.7 together with the discussion of Sections 5.1 and 5.2. □

Our luck with the colored Jones function is that we can identify it with a multisum of a q–proper hypergeometric function. Are we really lucky, or is there some deeper explanation? We believe that there is a underlying geometric reason for coincidence, which in a sense explains the underlying geometry of topological quantum field theory. We will postpone to a later publication applications of this principle to Hyperbolic Geometry; [11].

6 In computer talk

In this section we will show that Proposition 3.7 can be implemented by computer.

For every knot, one can write down a multisum formula for the colored Jones function, where the summand is q–hypergeometric. Occasionally, this multisum formula can be written as a single sum. There are various programs that can compute the recursion relations and their orders for multisums. In maple, one may use `qEKHAD` developed by Zeilberger [28]. In Mathematica, one may use the `qZeil.m` and `qMultiSum.m` packages of RISC developed by Paule and Riese [26, 27, 30].

6.1 Recursion relations for the cyclotomic function of twist knots

The twist knots Kp for integer p are shown in Figure 2. Their planar projections have $2|p|+2$ crossings, $2|p|$ of which come from the full twists, and 2 come from the negative clasp.

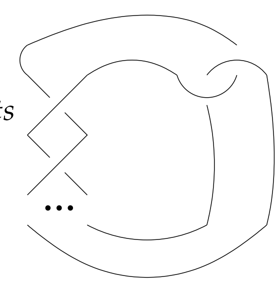


Figure 2: The twist knot K_p, for integers p. For $p=-1$, it is the Figure 8, for $p=0$ it is the unknot, for $p=1$ it is the left trefoil and for $p=2$ it is the Stevedore's ribbon knot.

Masbaum, [25], following Habiro and Le gives the following formula for the cyclotomic function of a twist knot. Let $c(p,\cdot)$ denote the cyclotomic function of the twist knot Kp. Rearranging a bit Masbaum's formula [25, Eqn.(35)], we obtain that:

$$\begin{aligned} c(p,n) &= (-1)^{n+1} q^{n(n+3)/2} \\ &\sum_{k=0}^{\infty} (-1)^k q^{k(k+1)p+k(k-1)/2} (q^{2k+1}-1) \frac{(q;q)_n}{(q;q)_{n+k+1}(q;q)_{n-k}} \end{aligned} \tag{2}$$

The above sum has compact support for each n. Now, in computer talk, we have:

```
Mathematica 4.2 for Sun Solaris
Copyright 1988-2000 Wolfram Research, Inc.
 -- Motif graphics initialized --
In[1]:=<< qZeil.m

q-Zeilberger Package by Axel Riese -- ©RISC Linz -- V 2.35 (04/29/03)
```

For $p=-1$ (which corresponds to the Figure 8 knot) the program gives:

```
In[2]:= qZeil[q^(n(n + 3)/2) (-1)^(n + k + 1) q^(-k(k + 1))(q^(2k + 1)
        - 1)qfac[q, q, n]/(qfac[q, q, n + k + 1] qfac[q, q, n - k])
```

```
        q^(k(k - 1)/2), {k, 0, Infinity}, n, 1]

Out[2]= SUM[n] == SUM[-1 + n]
```

which means that $c(-1,n) = c(-1,n-1)$ in accordance to the discussion after [25, Thm.5.1] which states $c(-1,n) = 1$ for all n.

For $p = 1$ (which corresponds to the left hand trefoil) the program gives:

```
In[3]:= qZeil[q^(n(n + 3)/2) (-1)^(n + k + 1) q^(k(k + 1))(q^(2k + 1)
        - 1)qfac[q, q, n]/(qfac[q, q, n + k + 1] qfac[q, q, n - k])
        q^(k(k - 1)/2), {k, 0, Infinity}, n, 1]

                          1 + n
Out[3]= SUM[n] == -(q          SUM[-1 + n])
```

which means that $c(1,n) = -q^{n+1}c(1,n-1)$ in accordance to the discussion after [25, Thm.5.1] which states $c(1,n) = (-1)^n q^{n(n+3)/2}$ for all n.

Similarly, for $p = 2$ (which corresponds to Stevedore's ribbon knot) the program gives:

```
In[4]:= qZeil[q^(n(n + 3)/2) (-1)^(n + k + 1) q^(2k(k + 1))(q^(2k + 1)
        - 1)qfac[q, q, n]/(qfac[q, q, n + k + 1] qfac[q, q, n - k])
        q^(k(k - 1)/2), {k, 0, Infinity}, n, 1]
Out[4]:= No solution: Increase order by 1
```

which *proves* that $c(2,n)$ satisfies no first order recursion relation. It does satisfy a second order recursion relation, as we find by:

```
In[5]:= qZeil[q^(n(n + 3)/2) (-1)^(n + k + 1) q^(2k(k + 1))(q^(2k + 1)
        - 1) qfac[q, q, n]/(qfac[q, q, n + k + 1] qfac[q, q, n - k])
        q^(k(k - 1)/2), {k, 0, Infinity}, n, 2]
                          2 + 2 n       -1 + n
Out[4]= SUM[n] == -(q            (1 - q       ) SUM[-2 + n]) -

       1 + n            n    2 n
>     q       (1 + q - q  + q    ) SUM[-1 + n]
```

Thus, the program computes not only a recursion relation, but also the order of a minimal one. Experimentally, it follows that $c(p,n)$ satisfies a recursion relation of order $|p|$, for all p. Perhaps one can guess the form of a minimal order recursion relation for all twist knots.

Actually, more is true. Namely, the formula for $c(p,n)$ shows that it is a q–holonomic function in *both* variables (p,n). Thus, we are guaranteed to find recursion relations with respect to n and with respect to p. Usually, recursion relations with respect to p for fixed n are called *skein theory* for the nth colored Jones function, because the knot is changing, and the color is fixed.

Thus, q–holonomicity implies skein relations (with respect to the number of twists) for the nth colored Jones polynomial of twist knots, for every fixed n.

For computations of recursion relations of the cyclotomic function of twist knots, we refer the reader to [12].

6.2 Recursion relations for the colored Jones function of the figure 8 knot

The Mathematica package `qMultiSum.m` can compute recursion relations for q–multisums. Using this, we can compute equally easily the recursion relation for the colored Jones function. Due to the length of the output, we illustrate this by computing the recursion relation for the colored Jones function of the Figure 8 knot. Recall from Equation (2) for $p = -1$ and from the fact that $c(-1, n) = 1$ that the colored Jones function of the figure 8 knot is given by:

$$J_{K(-1)}(n) = \sum_{k=0}^{\infty} q^{nk} (q^{-n-1}; q^{-1})_k (q^{-n+1}; q)_k. \tag{3}$$

In computer talk,

```
In[6]:= qZeil[q^(n k) qfac[q^(-n-1),q^(-1),k] qfac[q^(-n+1),q,k],
        {k,0,Infinity},n,2]

                    -1 - n        n          2 n
                   q       (q + q ) (-q + q    )
Out[6]= SUM[n] == ---------------------------- -
                                  n
                              -1 + q

             -2 + n         -1 + 2 n
        (1 - q      ) (1 - q        ) SUM[-2 + n]
>       ----------------------------------------- +
                     n        -3 + 2 n
              (1 - q ) (1 - q         )

          -2 - 2 n        -1 + n 2        -1 + n
>       (q         (1 - q       )  (1 + q       )
```

```
      4    4 n    3 + n    1 + 2 n    3 + 2 n    1 + 3 n
>   (q  + q    - q      - q        - q        - q       ) SUM[-1 + n])/

              n        -3 + 2 n
>     ((1 - q ) (1 - q         ))
```

This is a second order *inhomogeneous* recursion relation for the colored Jones function. A third order homogeneous relation may be obtained by:

```
In[7]:= MakeHomRec[%, SUM[n]]

         2 + n    3    n
        q      (-q  + q ) SUM[-3 + n]
Out[7]= ---------------------------- -
             2    n     5    2 n
           (q  + q ) (-q  + q   )

      -2 - n   2    n    8    4 n      6 + n    7 + n    3 + 2 n
>    (q       (q  - q ) (q  + q    - 2 q      + q      - q        +

          4 + 2 n    5 + 2 n    1 + 3 n      2 + 3 n
>        q        - q        + q        - 2 q       ) SUM[-2 + n]) /

              n    5    2 n
>     ((q + q ) (q  - q   )) +

      -1 - n        n    4    4 n    2 + n      3 + n    1 + 2 n
>    (q       (-q + q ) (q  + q    + q      - 2 q      - q        +

          2 + 2 n    3 + 2 n      1 + 3 n    2 + 3 n
>        q        - q        - 2 q        + q       ) SUM[-1 + n]) /

                                       1 + n        n
          2    n         2 n          q      (-1 + q ) SUM[n]
>     ((q  + q ) (-q + q   )) + ---------------------- == 0
                                          n       2 n
                                    (q + q ) (q - q   )
```

Of course, we can clear denominators and write the above recursion relation using the q–Weyl algebra $\mathcal{A}$. Let us end with a matching the theoretical bound for the recursion relation from Section 5 with the computer calculated bound from this section. Using Theorem 5, it follows that the summand satisfies a recursion relation of order $J^\star = 1^2+1^2 = 2$. This implies that the colored Jones function of the Figure 8 knot satisfies an inhomogeneous relation of degree 2 as was found above. The program also confirms that the colored Jones function of the Figure 8 knot does not satisfy an inhomogeneous relation of order less than 2.

7 The colored Jones function for a simple Lie algebra

Fix a *simple complex Lie algebra* $\mathfrak{g}$ of *rank* ℓ. For every knot $\mathcal{K}$ and every finite-dimensional $\mathfrak{g}$–module V, called *the color* of the knot, one can define the *quantum invariant* $J_{\mathcal{K}}(V) \in \mathbb{Z}[q^{\pm 1/2D}]$, where D is the determinant of the Cartan matrix of $\mathfrak{g}$. Simple $\mathfrak{g}$–modules are parametrized by the set of *dominant weights*, which can be identified, after we choose fixed fundamental weights, with $\mathbb{N}^\ell$. Hence $J_{\mathcal{K}}$ can be considered as a function $J_{\mathcal{K}}\colon \mathbb{N}^\ell \to \mathbb{Z}[q^{\pm 1/2D}]$.

Theorem 6 *For every simple Lie algebra other than G_2, and a set of fixed fundamental weights, the colored Jones function $J_{\mathcal{K}}\colon \mathbb{N}^\ell \to \mathbb{Z}[q^{\pm 1/2D}]$ is q–holonomic.*

Hence the colored Jones function will satisfy some recursion relations, which, together with values at a finitely many initial colors, totally determine the colored Jones function $J_{\mathcal{K}}$.

Remark 7.1 The reason we exclude the G_2 Lie algebra is technical. Namely, at present we cannot prove that the structure constants of the multiplication of the quantized enveloping algebra of G_2 with respect to a standard PBW basis, are q–holonomic; see Remark A.3. We believe however, that the theorem also holds for G_2.

The proof occupies the rest of this section. We will define $J_{\mathcal{K}}$ using representation of the braid groups coming from the R–matrix acting on Verma modules (instead of finite-dimensional modules). We then show that the R–matrix is q–holonomic. The theorem follows from that fact that products and traces of q–holonomic matrices are q–holonomic.

7.1 Preliminaries

Fix a *Cartan subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$ and a basis $\{\alpha_1, \dots, \alpha_\ell\}$ of simple roots for the dual space $\mathfrak{h}^*$. Let $\mathfrak{h}^*_{\mathbb{R}}$ be the $\mathbb{R}$–vector space spanned by $\alpha_1, \dots, \alpha_\ell$. The *root lattice* Y is the $\mathbb{Z}$–lattice generated by $\{\alpha_1, \dots, \alpha_\ell\}$. Let X be the *weight lattice* that is spanned by the *fundamental weights* $\lambda_1, \dots, \lambda_\ell$. Normalize the *invariant scalar product* $(\cdot,\cdot)$ on $\mathfrak{h}^*_{\mathbb{R}}$ so that $(\alpha,\alpha)=2$ for every *short* root α. Let D be the determinant of the Cartan matrix, then $(x,y)\in\frac{1}{D}\mathbb{Z}$ for $x,y\in X$.

Let $s_i, i=1,\dots,\ell$, be the *reflection along the wall* $\alpha_i^{\perp}$. The *Weyl group* W is generated by $s_i, i=1\dots,\ell$, with the braid relations together with $s_i^2=1$. A word $w=s_{i_1}\dots s_{i_r}$ is *reduced* if w, considered as an element of W, can not be expressed by a shorter word. In this case the *length* $l(w)$ of the element $w\in W$ is r. The *longest element* ω_0 in W has length $t=(\dim(\mathfrak{g})-\ell)/2$, the number of positive roots of $\mathfrak{g}$.

7.1.1 The quantum group $\mathcal{U}$

The quantum group $\mathcal{U}=\mathcal{U}_q(\mathfrak{g})$ associated to $\mathfrak{g}$ is a *Hopf algebra* defined over $\mathbb{Q}(v)$, where v is the usual quantum parameter (see [17, 22]). Here our v is the same as v of Lusztig [22] and is equal to q of Jantzen [17], while our q is v^2. The standard generators of $\mathcal{U}$ are $E_\alpha, F_\alpha, K_\alpha$ for $\alpha\in\{\alpha_1,\dots,\alpha_\ell\}$. For a full set of relations, as well as a good introduction to quantum groups, see [17]. Note that all the K_α's commute with each other.

For an element $\gamma\in Y$, $\gamma=k_1\alpha_1+\cdots+k_\ell\alpha_\ell$, let $K_\gamma := K_{\alpha_1}^{k_1}\dots K_{\alpha_1}^{k_1}$.

There is a Y–*grading* on $\mathcal{U}$ defined by $|E_\alpha|=\alpha, |F_\alpha|=-\alpha$, and $|K_\alpha|=0$. If x is homogeneous, then

$$K_\gamma x = v^{(\alpha,|x|)} x K_\gamma.$$

Let $\mathcal{U}^+$ be the subalgebra of $\mathcal{U}$ generated by the E_α, $\mathcal{U}^-$ by the F_α, and $\mathcal{U}^0$ by the K_α. It is known that the map

$$\begin{aligned} \mathcal{U}^-\otimes\mathcal{U}^0\otimes\mathcal{U}^+ &\to \mathcal{U} \\ (x,x',x'') &\to xx'x'' \end{aligned}$$

is an isomorphism of *vector spaces*.

7.1.2 Verma modules and finite dimensional modules

Let $\lambda\in X$ be a weight. The *Verma module* $M(\lambda)$ is a $\mathcal{U}$–module with underlying vector space $\mathcal{U}^-$ and with the action of $\mathcal{U}$ that is uniquely determined by the following condition. Here η is the unit of the algebra $\mathcal{U}^-$:

$$\begin{aligned} E_\alpha\cdot\eta &= 0 \quad \text{for all} \quad \alpha \\ K_\alpha\cdot\eta &= v^{(\alpha,\lambda)}\eta \quad \text{for all} \quad \alpha \\ F_\alpha\cdot x &= F_\alpha x \quad \text{for all} \quad \alpha\in\{\alpha_1,\dots,\alpha_\ell\}, x\in\mathcal{U}^- \end{aligned}$$

If $(\lambda|\alpha_i) < 0$ for all $i = 1, ..., \ell$ then $M(\lambda)$ is irreducible. On the other hand if $(\lambda|\alpha_i) \geq 0$ for all $i = 1, ..., \ell$ (ie, λ is dominant), then $M(\lambda)$ has a unique proper maximal submodule, and the quotient $L(\lambda)$ of $M(\lambda)$ by the proper maximal submodule is a finite dimensional module (of type 1, see [17]). Every finite dimensional module of type 1 of $\mathcal{U}$ is a direct sum of several $L(\lambda)$.

7.1.3 Quantum braid group action

For each fundamental root $\alpha \in \{\alpha_1, \dots, \alpha_\ell\}$ there is an *algebra automorphism* $T_\alpha \colon \mathcal{U} \to \mathcal{U}$, as described in [17, Chapter 8]. These automorphisms satisfy the following relations, known as the *braid relations*, or *Coxeter moves*.

If $(\alpha, \beta) = 0$, then $T_\alpha T_\beta = T_\beta T_\alpha$.

If $(\alpha, \beta) = -1$, then $T_\alpha T_\beta T_\alpha = T_\beta T_\alpha T_\beta$.

If $(\alpha, \beta) = -2$, then $T_\alpha T_\beta T_\alpha T_\beta = T_\beta T_\alpha T_\beta T_\alpha$.

If $(\alpha, \beta) = -3$, then $T_\alpha T_\beta T_\alpha T_\beta T_\alpha T_\beta = T_\beta T_\alpha T_\beta T_\alpha T_\beta T_\alpha$.

Note that the Weyl group is generated by s_α with exactly the above relations, replacing T_α by s_α, and the extra relations $s_\alpha^2 = 1$.

Suppose $w = s_{i_1} \dots s_{i_r}$ is a reduced word, one can define

$$T_w := T_{\alpha_{i_1}} \dots T_{\alpha_{i_r}}.$$

Then T_w is well-defined: If w, w' are two reduced words of the same element in W, then $T_w = T_{w'}$. This follows from the fact that any two reduced presentations of an element of W are related by a sequence of Coxeter moves.

7.1.4 Ordering of the roots

Suppose $w = s_{i_1} s_{i_2} \dots s_{i_t}$ is a reduced word representing the longest element ω_0 of the Weyl group. For r between 1 and t let

$$\gamma_r(w) := s_{i_1} s_{i_2} \dots s_{i_{r-1}}(\alpha_{i_r}).$$

Then the set $\{\gamma_i, i = 1, ..., t\}$ is exactly the set of positive roots. We *totally order* the set of positive roots by $\gamma_1 < \gamma_2 < \dots < \gamma_t$. This order depends on the reduced word w, and has the following *convexity* property: If β_1, β_2 are two positive roots such that $\beta_1 + \beta_2$ is also a root, then $\beta_1 + \beta_2$ is between β_1 and β_2. In particular, the first and the last, γ_1 and γ_t, are always fundamental roots. Conversely, any convex total ordering of the set of positive roots comes from a reduced word representing the longest element of W.

7.1.5 PBW basis for $\mathcal{U}^-$,$\mathcal{U}^+$, and $\mathcal{U}$

Suppose $w = s_{i_1} \dots s_{i_t}$ is a reduced word representing the longest element of W. Let us define

$$\begin{aligned} e_r(w) &= T_{\alpha_{i_1}} T_{\alpha_{i_2}} \dots T_{\alpha_{i_{r-1}}}(E_{\alpha_{i_r}}), \\ f_r(w) &= T_{\alpha_{i_1}} T_{\alpha_{i_2}} \dots T_{\alpha_{i_{r-1}}}(F_{\alpha_{i_r}}). \end{aligned}$$

Then $|e_r| = \gamma_r = -|f_r|$. (We drop w if there is no confusion.)

If γ_r is one of the fundamental roots, $\gamma_r = \alpha \in \{\alpha_1, \dots, \alpha_\ell\}$, then $e_r(w) = E_\alpha$, $f_r(w) = F_\alpha$ (and do not depend on w).

For $t \geq j \geq i \geq 1$ let $\mathcal{U}^-[j,i]$ be the vector space spanned by $f_j^{n_j} f_{j-1}^{n_{j-1}} \dots f_i^{n_i}$, for all $n_j, n_{j-1}, \dots, n_i \in \mathbb{N}$ and let $\mathcal{U}^+[i,j]$ the vector space spanned by $e_i^{n_i} e_{i+1}^{n_{i+1}} \dots e_j^{n_j}$, for all $n_j, n_{j-1}, \dots, n_i \in \mathbb{N}$. It is known that $\mathcal{U}^- = \mathcal{U}^-[t,1]$ and $\mathcal{U}^+ = \mathcal{U}^+[1,t]$.

For $\mathbf{n} = (n_1, \dots, n_t) \in \mathbb{N}^t$, $\mathbf{j} = (j_1, \dots j_\ell) \in \mathbb{Z}^\ell$ and $\mathbf{m} = (m_1, \dots, m_t) \in \mathbb{N}^t$ let us define $\mathbf{f}^\mathbf{n}$, $\mathbf{K}^\mathbf{j}$ and $\mathbf{e}^\mathbf{m}$ by

$$\mathbf{f}^\mathbf{n}(w) := f_t^{n_t} \dots f_1^{n_1}, \qquad \mathbf{K}^\mathbf{j} := K_{j_1\alpha_1} \dots K_{j_\ell \alpha_\ell} \qquad \mathbf{e}^\mathbf{n}(w) := e_1^{n_1} \dots e_t^{n_t}.$$

Then as vector spaces over $\mathbb{Q}(v)\mathcal{U}^-$, $\mathcal{U}^+$ and $\mathcal{U}$ have *Poincare-Birkhoff-Witt* (in short, PBW) basis

$$\{\mathbf{f}^\mathbf{n} \,|\, \mathbf{n} \in \mathbb{N}^t\}, \qquad \{\mathbf{e}^\mathbf{m} \,|\, \mathbf{m} \in \mathbb{N}^t\}, \qquad \{\mathbf{f}^\mathbf{n}\mathbf{K}^\mathbf{j}\mathbf{e}^\mathbf{m} \mid \mathbf{n}, \mathbf{m} \in \mathbb{N}^t, \mathbf{j} \in \mathbb{Z}^\ell\}$$

respectively, associated with the reduced word w.

In order to simplify notation, we define $S := \mathbb{N}^t \times \mathbb{Z}^\ell \times \mathbb{N}^t$, and $x_\sigma := \mathbf{f}^n \, K^\mathbf{j} \, \mathbf{e}^m$. Thus,

$$\{x_\sigma \,|\, \sigma \in S\} \tag{4}$$

is a PBW basis of $\mathcal{U}$ with respect to the reduced word w.

7.1.6 A commutation rule

For $x, y \in \mathcal{U}$ homogeneous let us define

$$[x,y]_q := xy - v^{(|x|,|y|)} yx.$$

Note that, in general, $[y,x]_q$ is not proportional to $[x,y]_q$.

An important property of the PBW basis is the following commutation rule, see [18]. If $i < j$ then $[f_i, f_j]_q$ belongs to $\mathcal{U}^-[j-1, i+1]$ (which is 0 if $j = i+1$).

It follows that $\mathcal{U}^-[j,i]$ is an *algebra*. This allows us to sort algorithmically non-commutative monomials in the variables f_i. Also two consecutive variables always q–commute: $[f_i, f_{i+1}]_q = 0$.

Similarly, if $i < j$ then $[e_i, e_j]_q$ belongs to $\mathcal{U}^+[i+1, j-1]$ (which is 0 if $j = i+1$). It follows that $\mathcal{U}^+[i,j]$ is an *algebra*, and two consecutive variables always q–commute, $[e_i, e_{i+1}]_q = 0$.

7.2 q–holonomicity of quantum groups

Suppose $A\colon \mathcal{U} \to \mathcal{U}$ is a linear operator. Using the PBW basis of $\mathcal{U}$ (see Equation (4)), we can present A by a matrix:

$$A(x_\sigma) = \sum_{\sigma'} A_\sigma^{\sigma'} x_{\sigma'},$$

with $A_\sigma^{\sigma'} \in \mathbb{Q}(v)$. We call (σ, σ') the coordinates of the matrix entry $A_\sigma^{\sigma'}$.

Definition 7.2 We say that A is *q–holonomic* if the matrix entry $A_\sigma^{\sigma'}$, considered as a function of (σ, σ') is q–holonomic with respect to all the variables.

A priori this definition depends on the reduced word w. But we will soon see that if A is q–holonomic in one PBW basis, then it is so in any other PBW basis.

7.2.1 q–holonomicity of transition matrix

Suppose $x_\sigma(w')$ is another PBW basis associate to another reduced word w' representing the longest element of W. Then we have the transition matrix $M_\sigma^{\sigma'}$ between the two bases, with entries in $\mathbb{Q}(v)$. The next proposition checks that the entries of the transition matrix are q–holonomic, by a standard reduction to the rank 2 case.

Proposition 7.3 *Except for the case of G_2, the matrix entry $M_\sigma^{\sigma'}$ is q–holonomic with respect to all its coordinates.*

Proof Since any two reduced presentations of an element of W are related by a sequence of Coxeter moves, it is enough to consider the case of a single Coxeter move. Since each Coxeter move involves only two fundamental roots and all T_α's are *algebra* isomorphisms, it is enough to considered the case of rank 2 Lie algebras. For all rank 2 Lie algebras (except G_2) we present the proof in Appendix. □

7.2.2 Structure constants

Recall the PBW basis $\{x_\sigma \,|\, \sigma \in S\}$ of the algebra $\mathcal{U}$. The multiplication in $\mathcal{U}$ is determined by the structure constants $c(\sigma, \sigma', \sigma'') \in \mathbb{Q}(v)$ defined by:

$$x_\sigma x_{\sigma'} = \sum_{\sigma''} c(\sigma, \sigma', \sigma'') x_{\sigma''}.$$

We will show the following:

Theorem 7 *The structure constant $c(\sigma, \sigma', \sigma'')$ is q–holonomic with respect to all its variables.*

Proof will be given in subsection 7.4.5.

7.2.3 Actions on Verma modules are q–holonomic

Each Verma module $M(\lambda)$ is naturally isomorphic to $\mathcal{U}^-$, as a vector space, via the map $u \to u \cdot \eta$. Using this isomorphism we identify a PBW basis of $\mathcal{U}^-$ with a basis of $M(\lambda)$, also called a PBW basis. If $u \in \mathcal{U}$, then the action of u on $M(\lambda)$ in a PBW basis can be written by a matrix $u_{\mathbf{n}}^{\mathbf{n}'}$ with entries in $\mathbb{Q}(v)$. We call $(\mathbf{n}, \mathbf{n}') \in \mathbb{N}^t \times \mathbb{N}^t$ the coordinates of the matrix entry.

Proposition 7.4 *For every r with $1 \le r \le t$, the entries of the matrices e_r^k, f_r^k are q–holonomic with respect to k, λ, and the coordinates of the entry.*

This Proposition follows immediately from Theorem 7 and Fact 0.

7.3 Quantum knot invariants

7.3.1 The quasi-R–matrix

Fix a reduced word w representing the longest element of W. For each $r, 1 \le r \le t$, let

$$\Theta_r := \sum_{k \in \mathbb{N}} c_k \, f_r^k \otimes e_r^k,$$

where
$$c_k = (-1)^k v_{\gamma_r}^{-k(k-1)/2} \, \frac{(v_{\gamma_r} - v_{\gamma_r}^{-1})^k}{[k]_{\gamma_r}!}.$$

Here $v_\gamma = v^{(\gamma|\gamma)/2}$, and

$$[k]_\gamma! = \prod_{i=1}^{k} \frac{v_\gamma^i - v_\gamma^{-i}}{v_\gamma - v_\gamma^{-1}}.$$

The main thing to observe is that c_k is q–holonomic with respect to k. Note that although Θ_r is an infinite sum, for every weight $\lambda \in X$, the action of Θ_r on $M(\lambda) \otimes M(\lambda)$ is well-defined. This is because the action of e_r is locally nilpotent, ie, for every $x \in M(\lambda)$, there is k such that $e_r^k \cdot x = 0$.

The *quasi-R–matrix* is:

$$\Theta := \Theta_t \Theta_{t-1} \dots \Theta_1.$$

We will consider Θ as an operator from $M(\lambda) \otimes M(\lambda)$ to itself. There is a natural basis for $M(\lambda) \otimes M(\lambda)$ coming from the PBW basis of $M(\lambda)$.

Proposition 7.5 *The matrix of Θ acting on $M(\lambda)$ in a PBW basis is q–holonomic with respect to all the coordinates of the entry and λ.*

Proof It's enough to prove the statement for each Θ_r. The result for Θ_r follows from the fact that the actions of e_r^k, f_r^k on $M(\lambda)$, as well as c_k, are q–holonomic in k and so are all the coordinates of the matrix entries, by Proposition 7.4 . □

7.3.2 The R–matrix and the braiding

As usual, let us define the weight on $M(\lambda)$ by declaring the weight of $F_{\mathbf{n}} \cdot \eta$ to be $\lambda - \sum n_i \gamma_i$, where $\mathbf{n} = (n_1, \dots, n_t)$. The space $M(\lambda)$ is the direct sum of its weight subspaces.

Let $\mathcal{D}$: $M(\lambda) \otimes M(\lambda) \to M(\lambda) \otimes M(\lambda)$ be the linear operator defined by

$$\mathcal{D}(x \otimes y) = v^{-(|x|,|y|)} x \otimes y.$$

Clearly $\mathcal{D}$ is q–holonomic; it's called the diagonal part of the R–matrix, which is $R := \Theta\mathcal{D}$.

The braiding is $\mathcal{B} := R\sigma$, where $\sigma(x \otimes y) = y \otimes x$. Combining the above results, we get the following:

Theorem 8 *The entry of the matrix of the braiding acting on $M(\lambda)$ is q–holonomic with respect to all the coordinates and λ.*

Remark 7.6 Technically, in order to define the diagonal part $\mathcal{D}$, one needs to extend the ground ring to include a D-th root of v, since (λ, μ), with $\lambda, \mu \in X$, is in general not an integer, but belonging to $\frac{1}{D}\mathbb{Z}$.

7.3.3 q–holonomicity of quantum invariants of knots

First let us recall the definition of quantum knot invariant.

Using the braiding $\mathcal{B}\colon\ M(\lambda) \to M(\lambda)$ one can define a representation of the braid group $\tau\colon\ B_m \to (M(\lambda))^{\otimes m}$ by putting

$$\tau(\sigma_i) := \mathrm{id}^{\otimes i-1} \otimes \mathcal{B} \otimes \mathrm{id}^{\otimes m-i-1}.$$

Let ρ denote the half-sum of positive roots. For an element $x \in \mathcal{U}$ and an $\mathcal{U}$–module V, the quantum trace is defined as

$$\mathrm{tr}_q(x, V) := \mathrm{tr}(xK_{-2\rho}, V).$$

Suppose a framed knot $\mathcal{K}$ is obtained by closing a braid $\beta \in B_m$. We would say that the colored Jones polynomial is the quantum trace of $\tau(\beta)$. However, since $M(\lambda)$ is infinite-dimensional, the trace may not make sense. Instead, we will use a trick of *breaking the knot*. Let $\mathcal{K}'$ denote the *long knot* which is the closure of all but the first strand of β.

Recall that $\tau(\beta)$ acts on $(M(\lambda))^{\otimes m}$. Let

$$\tau(\beta)(\lambda)^{\mathbf{n}'_1,\dots,n'_m}_{\mathbf{n}_1,\dots,\mathbf{n}_m} \in \mathbb{Z}[v^{\pm 1/D}]$$

be the entries of the matrix $\tau(\beta)(\lambda)$. We will take partial trace by first putting $\mathbf{n}_1 = \mathbf{n}'_1 = 0$ and then take the sum over all $\mathbf{n}_2 = \mathbf{n}'_2, \dots, \mathbf{n}_m = \mathbf{n}'_m$. The following lemma shows that the sum is actually finite.

Lemma 7.7 *Suppose $\mathbf{n}_1 = 0$. There are only a finite number of collections of $(\mathbf{n}_2, \mathbf{n}_3, \dots, \mathbf{n}_m) \in \mathbb{N}^{t-1}$ such that*

$$\tau(\beta)(\lambda)^{\mathbf{n}_1,\dots,\mathbf{n}_m}_{\mathbf{n}_1,\dots,\mathbf{n}_m}$$

is not zero.

Proof Let $M'(\lambda)$ be the maximal proper $\mathcal{U}$–submodule of $M(\lambda)$. Then $L(\lambda) = M(\lambda)/M'(\lambda)$ is a finite dimensional vector space. In particular it has only a finite number of non-trivial weights. Hence, all except for a finite number of $\mathbf{f_n}, \mathbf{n} \in \mathbb{N}^t$, are in $M'(\lambda)$.

We present the coefficients $\mathcal{B}_\pm(\lambda)$ graphically as in Figure 3.

Note that if $(\mathcal{B}_\pm)^{\mathbf{m}_1,\mathbf{m}_2}_{\mathbf{n}_1,\mathbf{n}_2}$ is not equal to 0, then $\mathbf{f}_{\mathbf{m}_2}$ can be obtained from $\mathbf{f}_{\mathbf{n}_1}$ by action of an element in $\mathcal{U}$, and similarly, $\mathbf{f}_{\mathbf{m}_1}$ can be obtained from $\mathbf{f}_{\mathbf{n}_2}$ by action of an element in $\mathcal{U}$. Thus if we move upwards along a string of the braid,

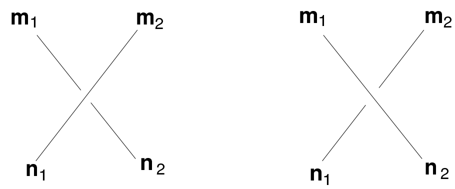


Figure 3: $(\mathcal{B}_+)^{\mathbf{m}_1,\mathbf{m}_2}_{\mathbf{n}_1,\mathbf{n}_2}$ and $(\mathcal{B}_-)^{\mathbf{m}_1,\mathbf{m}_2}_{\mathbf{n}_1,\mathbf{n}_2}$

the basis element at the top can always be obtained from the one at the bottom by an action of $\mathcal{U}$.

Because the closure of β is a knot, by moving around the braid one can get any point from any particular point. Because the basis element $\mathbf{f}_0$ is not in $M'(\lambda)$, we conclude that if

$$\tau(\beta)(\lambda)^{\mathbf{n}_1,\dots,\mathbf{n}_m}_{\mathbf{n}_1,\dots,\mathbf{n}_m}$$

is not 0, with $\mathbf{n}_1 = 0$, then all the basis vectors $\mathbf{f}_{\mathbf{n}_2}, \dots, \mathbf{f}_{\mathbf{n}_m}$ are not in $M'(\lambda)$, and there are only a finite number of such collections. $\square$

R ecall that 2ρ is the sum of all positive roots. Let us define

$$J_{\mathcal{K}'}(\lambda) = \sum_{\mathbf{n}_2,\dots\mathbf{n}_m \in \mathbb{N}^t, \mathbf{n}_1 = 0,} (K_{-2\rho}\,\tau(\beta)(\lambda))^{\mathbf{n}_1,\dots,n_m}_{\mathbf{n}_1,\dots,\mathbf{n}_m}.$$

From q–holonomicity of $\tau(\beta)(\lambda)$ it follows that $J_{\mathcal{K}'}(\lambda)$ is q–holonomic. $J_{\mathcal{K}'}(\lambda)$ is a long knot invariant, and is related to the colored Jones polynomial $J_{\mathcal{K}}$ of the knot $\mathcal{K}$ by

$$J_{\mathcal{K}}(\lambda) = J_{\mathcal{K}'}(\lambda) \times \dim_q(L(\lambda)),$$

where $L(\lambda)$ is the finite-dimensional simple $\mathcal{U}$–module of highest weight λ, and $\dim_q(L(\lambda))$ is its quantum dimension, and is given by the formula

$$\dim_q(L(\lambda)) = \prod_{\alpha>0} \frac{v^{(\lambda+\rho,\alpha)} - v^{-(\lambda+\rho,\alpha)}}{v^{(\rho,\alpha)} - v^{-(\rho,\alpha)}}.$$

Since $\dim_q(L(\lambda))$ is q–holonomic in λ, we see that $J_{\mathcal{K}}(\lambda)$ is q–holonomic. This completes the proof of Theorem 6. $\square$

Remark 7.8 The invariant $J_{\mathcal{K}'}$ of long knots is sometime more convenient. For example, $J_{\mathcal{K}}(\lambda)$ might contain fractional power of q, but (if $\mathcal{K}'$ has framing 0,) $J_{\mathcal{K}'}(\lambda)$ is always in $\mathbb{Z}[q^{\pm 1}]$, see [20]. Also the function $J_{\mathcal{K}'}$ can be extended to the whole weight lattice.

7.4 Proof of Theorem 7

7.4.1 r_α is q–holonomic

We will need the linear maps $r_\alpha, r'_\alpha \colon \mathcal{U}^\pm \to \mathcal{U}^\pm$, as defined in [17, Chapter 6]. Their restriction to $\mathcal{U}^-$ is uniquely characterized by the properties:

$$r_\alpha(xy) = r_\alpha(x)\, y + v^{(\alpha,|x|)}\, x\, r_\alpha(y) \qquad r'_\alpha(xy) = x\, r'_\alpha(y) + v^{(\alpha,|x|)}\, r'_\alpha(x)\, y \quad (5)$$

and for any two fundamental roots α, β, (see [17, Eqn.(6.15.4)]) and

$$r_\alpha(F_\beta^n) = r'_\alpha(F_\beta^n) = \delta_{\alpha,\beta}\frac{1 - v_\alpha^{2n}}{1 - v_\alpha^2} F_\alpha^{n-1}, \tag{6}$$

where $v_\alpha := v^{(\alpha,\alpha)/2}$; see [17, Eqn.(8.26.2)].

Lemma 7.9 *For a fixed $\alpha \in \{\alpha_1, \dots, \alpha_\ell\}$, the matrix entries of the operators $(r_\alpha)^k, (r'_\alpha)^k : \mathcal{U}^- \to \mathcal{U}^-$ are q–holonomic with respect to k and the coordinates of the matrix entry. Similarly, $(r_\alpha)^k, (r'_\alpha)^k \colon \mathcal{U}^+ \to \mathcal{U}^+$ are q–holonomic.*

Proof We give a proof for $r_\alpha^k \colon \mathcal{U}^- \to \mathcal{U}^-$. The other case is similar.

There is a reduced word $w' = s_{i_1} \dots s_{i_t}$ representing the longest element ω_0 of W such that $\alpha_{i_1} = \alpha$. Then $w = s_{i_2} \dots s_{i_t} s_{\bar\alpha}$ is another reduced word representing ω_0, where $\bar\alpha := -\omega_0(\alpha)$.

For the PBW basis of $\mathcal{U}^-$ associated with w it's known that $\gamma_t = \alpha$, and thus $f_t = F_\alpha$. According to [17, 8.26.5], for every x in the algebra $\mathcal{U}^-[t-1, 1]$, one has

$$r_\alpha(x) = 0.$$

Using Equations (5) and (6) and induction, one can easily show that for every $x \in \mathcal{U}^-[t-1, 1]$,

$$(r_\alpha)^k(f_t^{n_t} x) = \prod_{i=1}^{k} \frac{1 - v_\alpha^{2n_t - 2i + 2}}{1 - v_\alpha^2}\, f_t^{n_t - k} x,$$

This formula, applied to $x = f_{t-1}^{n_{t-1}} \dots f_1^{n_1}$, proves the statement. □

7.4.2 $U_q(\mathfrak{sl}_2)$ is q–holonomic

Lemma 7.10 *Theorem 7 holds true for $\mathfrak{g} = \mathfrak{sl}_2$.*

Proof The PBW basis for $\mathcal{U}$ is $F^n K^j E^m$, with $m, n \in \mathbb{N}$ and $j \in \mathbb{Z}$. First of all we know that

$$E_\alpha^m F_\alpha^n = \sum_{i=0}^{\infty} \begin{bmatrix} m \\ i \end{bmatrix}_{v_\alpha} \begin{bmatrix} n \\ i \end{bmatrix}_{v_\alpha} F_\alpha^{n-i}\, b(K_\alpha; 2i-n-m, i)\, E_\alpha^{m-i},$$

where
$$b(K_\alpha; a, i) := \prod_{j=1}^{i} \frac{K_\alpha v_\alpha^{a-j+1} - K_\alpha^{-1} v_\alpha^{-a+j-1}}{v_\alpha - v_\alpha^{-1}}.$$

Here for any root γ, one defines $v_\gamma = v^{(\gamma,\gamma)/2}$, and $\begin{bmatrix} m \\ i \end{bmatrix}_{v_\alpha}$ is the usual quantum binomial coefficient calculated with v replaced by v_α.

Hence

$$(F^m K^k E^n)(F^{m'} K^{k'} E^{m'}) = \sum_{i=0}^{\infty} F^{m+m'-i}\, a(m,k,n,m',k',n',i)\, E^{n+n'-i},$$

where

$$\begin{aligned} &a(m,k,n,m',k',n',i) \\ &\qquad = v^{2k(i-m')+2k'(i-n)} \begin{bmatrix} n \\ i \end{bmatrix} \begin{bmatrix} m' \\ i \end{bmatrix} [i]!\, b(K; 2i-n-m', i)\, K^{k+k'}. \end{aligned}$$

The value of the function a is in $\mathbb{Z}[v^{\pm 1}][K^{\pm 1}]$. Consider the coefficient of K^r in a; one gets a function of $m, n, k, m', n', k', i, r$ with values in $\mathbb{Z}[v^{\pm 1}]$ which is clearly q–holonomic with respect to all variables. The lemma follows. □

7.4.3 $E_\alpha^k, F_\alpha^k \colon \mathcal{U} \to \mathcal{U}$ are q–holonomic in k

Proposition 7.11 *For a fixed fundamental root $\alpha \in \{\alpha_1, \dots, \alpha_\ell\}$, the operators $E_\alpha^k, F_\alpha^k \colon \mathcal{U} \to \mathcal{U}$ of left multiplication are q–holonomic with respect to k and all the coordinates of the matrix entry. Similarly, the right multiplication by E_α^k, F_α^k are q–holonomic with respect to k and all the coordinates of the matrix entry.*

Proof (a) Left multiplication by F_α^k and right multiplication by E_α^k.

Choose w as in the proof of Lemma 7.9. Then $f_t = F_\alpha$ and $e_t = E_\alpha$, and an element of the PBW basis has the form $f_t^{n_t} x K_\beta y e_t^{m_t}$. It's clear that left multiplication by F_α and right multiplication by E_α are q–holonomic.

(b) Left multiplication by E_α^k.

Choose a reduced word $w = s_{i_1} \dots s_{i_t}$ representing the longest element ω_0 that begins with α: $\alpha_{i_1} = \alpha$. We have the corresponding PBW basis $f_i, e_i, i = 1, \dots, t$ with $f_1 = F_\alpha$ and $e_1 = E_\alpha$. Thus a typical element of the PBW basis has the form

$$xF_\alpha^{n_1} K_\beta E_\alpha^{m_1} y, \tag{7}$$

where $x = f_t^{n_t} \dots f_2^{n_2}, y = e_2^{m_2} \dots e_t^{m_t}$. By [17, 8.26.6], since $x \in \mathcal{U}^-[t,2]$, one has $r'_\alpha(x) = 0$. Using formula [17, 6.17.1], one can easily prove by induction that

$$(E_\alpha)^k x = \sum_{i=0}^{\infty} v^{i-ik} \begin{bmatrix} k \\ i \end{bmatrix}_{v_\alpha} \frac{K_\alpha^i}{(v_\alpha - v_\alpha^{-1})^i} (r_\alpha)^i(x) E_\alpha^{k-i}.$$

Using this formula one can move the E_α past x in the expression (7), (there appear r_α and K_α), then one moves E_α past F_α using the $\mathfrak{sl}_2$ case. The last step is moving past K_β is easy, since

$$E_\alpha K_\beta = v^{-(\beta,\alpha)} K_\beta E_\alpha.$$

Using Lemmas 7.9 and 7.10, we see that each "moving step" is q–holonomic. Hence we get the result for the left multiplication by E_α^k.

(c) Right multiplication by F_α^k.

The proof is similar. We use the same basis (7) as in the case b). For y, by Lemma 8.26 of [17], one has $r_\alpha(y) = 0$. Hence using induction based on the formula (6.17.2) of [17] one can show that

$$yF_\alpha^n = \sum_{i=0}^{\infty} \frac{v_\alpha^{i(n-i)}}{(v_\alpha^{-1} - v_\alpha)^i} \begin{bmatrix} n \\ i \end{bmatrix}_{v_\alpha} F_\alpha^{n-i} K_\alpha^{-i} (r'_\alpha)^i(y).$$

Using this formula, and the results for r'_α (Lemma 7.9) and $\mathfrak{sl}_2$ (Lemma 7.10) we can move F_α to the right. □

7.4.4 T_α is q–holonomic

Proposition 7.12 *For a fixed fundamental root $\alpha \in \{\alpha_1, \dots, \alpha_\ell\}$, the braid operator $T_\alpha\colon \mathcal{U} \to \mathcal{U}$ and its inverse T_α^{-1} are q–holonomic.*

Proof By Proposition 7.3 we can use any PBW basis.

Choose a reduced word $w' = s_{i_1} \dots s_{i_t}$ representing the longest element ω_0 that begins with α: $\alpha_{i_1} = \alpha$. Then $w = s_{i_2} \dots s_{i_t} s_{\bar{\alpha}}$ is another reduced word representing ω_0, where $\bar{\alpha}$ is the dual of α: $\bar{\alpha} = -\omega_0(\alpha)$.

We use f_r to denote $f_r(w)$, and f'_r to denote $f_r(w')$. The relation between the two PBW basis of w and w is as follows: For $1 \le r \le t-1$,

$$T_\alpha(f_r) = f'_{r+1}, \qquad T_\alpha(e_r) = e'_{r+1}.$$

Besides, $f_t = F_\alpha = f'_1, e_t = E_\alpha = e'_1$.

We will consider the matrix entry of $T_\alpha\colon \mathcal{U} \to \mathcal{U}$ where the source space is equipped with the PBW corresponding to w, while the target space with the PBW basis corresponding to w'.

From [17, Chapter 8], recall that:

$$T_\alpha(F_\alpha) = -K_\alpha^{-1}E_\alpha, \qquad T_\alpha(E_\alpha) = -F_\alpha K_\alpha.$$

Hence

$$T_\alpha(F_\alpha^n) = (-1)^n v_\alpha^{n(n-1)} K_\alpha^{-n} E_\alpha^n, \qquad T_\alpha(E_\alpha^m) = (-1)^m v_\alpha^{-m(m-1)} F_\alpha^m K_\alpha^m.$$

For a basis element $x_\sigma = f_t^{n_t} \dots f_1^{n_1} K_\beta e_1^{m_1} \dots e_t^{m_t}$, we have

$$\begin{aligned} T_\alpha(x_\sigma) = d_\alpha(n_t, m_t)\, K_\alpha^{-n_t} E_\alpha^{n_t} \times (f'_t)^{m_{t-1}} \\ \dots (f'_1)^{n_2} K_{s_\alpha\beta} (e'_1)^{m_2} \dots (e'_t)^{m_{t-1}} \times F_\alpha^{m_t} K_\alpha^{m_t}, \end{aligned}$$

where $$d_\alpha(n_t, m_t) := (-1)^{n_t+m_t} v_\alpha^{n_t(n_t-1)-m_t(m_1-1)}.$$

The left or right multiplication by K_α^n is q–holonomic with respect to n and all the coordinates. The left multiplication by E^{n_t}, as well as the right multiplication my $F_\alpha^{m_t}$ is q–holonomic with respect to n_t and all coordinates, by Proposition 7.11. One then can conclude that T_α is q–holonomic.

The proof for $T_\alpha^{-1}\colon \mathcal{U} \to \mathcal{U}$ is similar. One should use the PBW basis of w' for the source, and that of w for the target. □

7.4.5 Proof of Theorem 7

It is clear that for each $\mathbf{j} \in \mathbb{Z}^\ell$, the operator $\mathbf{K}^{\mathbf{j}}\colon \mathcal{U} \to \mathcal{U}$ of left multiplication is q–holonomic.

Fix a reduced word w representing the longest element of W. It suffices to show that for each $1 \le r \le t$ the operators $e_r^k, f_r^k\colon \mathcal{U} \to \mathcal{U}$ (left multiplication) are q–holonomic with respect to all variables, including k.

This is true if $e_r = E_\alpha$ and $f_r = F_\alpha$, where α is one of the fundamental roots, by Proposition 7.11. But any e_r or f_r can be obtained from E_α and F_α by actions of product of various T_{α_i}'s. Hence from Proposition 7.12 we get Theorem 7. □

A Appendix: Proof of Proposition 7.3 for A_2 and B_2

In this appendix we will prove Proposition 7.3 for the rank 2 Lie algebras A_2 and B_2. We will achieve this by a brute-force calculation.

First, let us discuss some simplification, due to symmetry. The transition matrix of $\mathcal{U}$ leaves invariant each of $\mathcal{U}^+,\mathcal{U}^-,\mathcal{U}^0$. On $\mathcal{U}^0$ the transition matrix is identity. Hence it's enough to consider the restriction of the transition matrix in $\mathcal{U}^-$ and $\mathcal{U}^+$. Furthermore, the Cartan symmetry (the operator τ of [17]) reduces the case of $\mathcal{U}^+$ to that of $\mathcal{U}^-$.

A.1 The case of A_2

There are two fundamental roots denoted by α and β. The set of positive roots is $\{\alpha,\beta,\alpha+\beta\}$. The reduced representations of the longest element of the Weyl group are $w = s_1s_2s_1$ and $w' = s_2s_1s_2$, where $s_1 = s_\alpha$ and $s_2 = s_\beta$.

The total ordering (see Section 7.1.4) of the set of positive roots corresponding to w and w' are, respectively:

$$\begin{aligned} (\gamma_1,\gamma_2,\gamma_3) &= (\alpha,\alpha+\beta,\beta) \\ (\gamma_{1'},\gamma_{2'},\gamma_{3'}) &= (\beta,\alpha+\beta,\alpha). \end{aligned}$$

Notice that $\gamma_{i'} = \gamma_{3-i}$ for $i = 1,\ldots,3$.

The PBW basis of $\mathcal{U}^-$ (see Section 7.1.5) corresponding to w and w' are, respectively:

$$\{f_3^m f_2^n f_1^p \mid m,n,p \in \mathbb{N}\}, \qquad \{f_{3'}^m f_{2'}^n f_{1'}^p \mid m,n,p \in \mathbb{N}\},$$

where

$$\begin{aligned} (f_3,f_2,f_1) &= (F_\beta, T_\alpha(F_\beta) = -v[F_\beta,F_\alpha]_q = F_\beta F_\alpha - vF_\alpha F_\beta, F_\alpha) \\ (f_{3'},f_{2'},f_{1'}) &= (F_\alpha, T_\beta(F_\alpha) = F_\alpha F_\beta - vF_\beta F_\alpha, F_\beta). \end{aligned}$$

From explicit formulas of [23, section 5] it follows that:

Lemma A.1 *The structure constants of $\mathcal{U}^-$, in the basis of w, is q–holonomic.*

Let us define a scalar product $(\cdot,\cdot)$ on $\mathcal{U}^-$ such that the PBW basis of w is an orthonormal basis. Since

$$f_{3'}^{m'} f_{2'}^{n'} f_{1'}^{p'} = \sum_{m,n,p} (f_{3'}^{m'} f_{2'}^{n'} f_{1'}^{p'}, f_3^m f_2^n f_1^n) f_3^m f_2^n f_1^p$$

Proposition 7.3 is equivalent to showing that

$$(f_{3'}^{m'} f_{2'}^{n'} f_{1'}^{p'}, f_3^m f_2^n f_1^n)$$

is q–holonomic in all variables m, n, p, m', n', p'.

Since multiplication is q–holonomic in the PBW basis of w (see Lemma A.1), it suffices to show that

$$(f_{i'}^k, f_3^m f_2^n f_1^n)$$

is q–holonomic in k, m, n, p for each $i = 1, 2, 3$. This is clear for $i = 1$ or $i = 3$, since $f_{1'} = f_3$ and $f_{3'} = f_1$. As for $f_{2'}$, an easy induction shows that

$$f_{2'}^n = (-v)^{-n} \sum_{k=0}^{\infty} v^{-k(k-3)/2} (v - v^{-1})^k \begin{bmatrix} n \\ k \end{bmatrix} f_3^k f_2^{n-k} f_1^k.$$

and the statement also holds true for $i = 2$. This proves Proposition 7.3 for A_2.

A.2 The case of B_2

There are two fundamental roots denoted here by α and β, where α is the short root. The set of positive roots is $\{\alpha, \beta, 2\alpha + \beta, \alpha + \beta\}$. The reduced representations of the longest element of the Weyl group are $w = s_1 s_2 s_1 s_2$ and $w' = s_2 s_1 s_2 s_1$, where $s_1 = s_\alpha$ and $s_2 = s_\beta$.

The total ordering of the set of positive roots corresponding to w and w' are, respectively:

$$\begin{aligned} (\gamma_1, \gamma_2, \gamma_3, \gamma_4) &= (\alpha, 2\alpha + \beta, \alpha + \beta, \beta) \\ (\gamma_{1'}, \gamma_{2'}, \gamma_{3'}, \gamma_{4'}) &= (\beta, \alpha + \beta, 2\alpha + \beta, \alpha). \end{aligned}$$

Notice that $\gamma_{i'} = \gamma_{4-i}$ for $i = 1, \ldots, 4$.

The PBW basis of $\mathcal{U}^-$ (see Section 7.1.5) corresponding to w and w' are, respectively:

$$\{f_4^l f_3^m f_2^n f_1^p \mid l, m, n, p \in \mathbb{N}\}, \qquad \{f_{4'}^l f_{3'}^m f_{2'}^n f_{1'}^p \mid l, m, n, p \in \mathbb{N}\},$$

where

$$\begin{aligned} (f_4, f_3, f_2, f_1) &= (F_\beta, F_\beta F_\alpha - v^2 F_\alpha F_\beta, \frac{F_\beta F_\alpha^2}{[2]} - v F_\alpha F_\beta F_\alpha + \frac{v^2 F_\alpha^2 F_\beta}{[2]}, F_\alpha) \\ (f_{4'}, f_{3'}, f_{2'}, f_{1'}) &= (F_\alpha, \frac{v^2 F_\beta F_\alpha^2}{[2]} - v F_\alpha F_\beta F_\alpha + \frac{F_\alpha^2 F_\beta}{[2]}, F_\alpha F_\beta - v^2 F_\beta F_\alpha, F_\beta). \end{aligned}$$

It follows from [23] that:

Lemma A.2 *The structure constants of $\mathcal{U}^-$, in the basis of w, is q–holonomic.*

Let us define a scalar product $(\cdot,\cdot)$ on $\mathcal{U}^-$ such that the PBW basis of w is an orthonormal basis. Then Proposition 7.3 is equivalent to

$$(f_{4'}^{l'} f_{3'}^{m'} f_{2'}^{n'} f_{1'}^{p'}, f_4^l f_3^m f_2^n f_1^n)$$

is q–holonomic in all variables $l, m, n, p, l'm', n', p'$.

Since multiplication is q–holonomic in the PBW basis of w (see Lemma A.2), it suffices to show that

$$(f_{i'}^k, f_4^l f_3^m f_2^n f_1^p)$$

is q–holonomic in k, l, m, n, p for each $i = 1, 2, 3, 4$. This is clear for $i' = 1$ or $i' = 4$, since $f_{1'} = f_4$ and $f_{4'} = f_1$. As for $i' = 2$ and $i' = 3$, the formula of [22, Section 37.1] shows that

$$\begin{aligned} f_{2'}^n &= \sum_{i=0}^{n} (-1)^i \frac{v^{2i} F_\beta^i F_\alpha^n F_\beta^{n-i}}{[n-i]_\beta ! [i]_\beta !} \\ f_{3'}^n &= \sum_{i=0}^{2n} (-1)^i \frac{v^{i} F_\alpha^{2n-i} F_\beta^n F_\alpha^{i}}{[2n-i]! [i]!} \end{aligned}$$

and since $F_\alpha = f_{4'}$ and $F_\beta = f_{1'}$, the cases of $i' = 2'$ and $i' = 3'$ reduce to the cases of $i' = 1'$ and $i' = 4'$. This proves Proposition 7.3 for B_2.

Remark A.3 If Lemma A.1 holds for G_2, then we can prove Proposition 7.3 for G_2.

References

[1] **D Bar-Natan**, *ColoredJones.nb*, Mathematica program, part of “KnotAtlas”, April 2003.

[2] **I N Bernšteĭn**, *Modules over a ring of differential operators. An investigation of the fundamental solutions of equations with constant coefficients*, Funkcional. Anal. i Priložen. 5 (1971) 1–16, English translation: 89–101 MathReview

[3] **I N Bernšteĭn**, *Analytic continuation of generalized functions with respect to a parameter*, Funkcional. Anal. i Priložen. 6 (1972) 26–40, English translation: 273–285 MathReview

[4] **J-E Björk**, *Rings of differential operators*, volume 21 of *North-Holland Mathematical Library*, North-Holland Publishing Co., Amsterdam (1979) MathReview

[5] **A Borel**, **P-P Grivel**, **B Kaup**, **A Haefliger**, **B Malgrange**, **F Ehlers**, *Algebraic D-modules*, Perspectives in Mathematics 2, Academic Press, Boston, MA (1987) MathReview

[6] **P Cartier**, *Démonstration "automatique" d'identités et fonctions hypergéométriques (d'après D. Zeilberger)*, Astérisque (1992) Exp. No. 746, 3, 41–91 MathReview

[7] **S C Coutinho**, *A primer of algebraic D-modules*, volume 33 of *London Mathematical Society Student Texts*, Cambridge University Press, Cambridge (1995) MathReview

[8] **F Chyzak**, **B Salvy**, *Non-commutative elimination in Ore algebras proves multivariate identities*, J. Symbolic Comput. 26 (1998) 187–227 MathReview

[9] **S Garoufalidis**, **T T Q Le**, **D Zeilberger**, *The quantum MacMahon Master Theorem*, `arXiv:math.QA/0303319`, to appear in Proc. Nat. Acad. Sci.

[10] **S Garoufalidis**, *Difference and differential equations for the colored Jones function*, `arXiv:math.GT/0306229`

[11] **S Garoufalidis**, *On the characteristic and deformation varieties of a knot*, from: "Proceedings of the CassonFest (Arkansas and Texas 2003)", Geom. Topol. Monogr. 7 (2004) 291–309

[12] **S Garoufalidis**, **X Sun**, *The C–polynomial of a knot*, preprint (2005) `arXiv:math.GT/0504305`

[13] **R Gelca**, *Non-commutative trigonometry and the A-polynomial of the trefoil knot*, Math. Proc. Cambridge Philos. Soc. 133 (2002) 311–323 MathReview

[14] **R Gelca**, **J Sain**, *The noncommutative A-ideal of a $(2, 2p + 1)$-torus knot determines its Jones polynomial*, J. Knot Theory Ramifications 12 (2003) 187–201 MathReview

[15] **K Habiro**, *On the quantum* sl_2 *invariants of knots and integral homology spheres*, from: "Invariants of knots and 3-manifolds (Kyoto, 2001)", Geom. Topol. Monogr. 4 (2002) 55–68 MathReview

[16] **K Habiro**, **T T Q Le**, in preparation

[17] **J Jantzen**, *Lectures on quantum groups*, volume 6 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI (1996) MathReview

[18] **L I Korogodski**, **Y S Soibelman**, *Algebras of functions on quantum groups. Part I*, volume 56 of *Mathematical Surveys and Monographs*, American Mathematical Society, Providence, RI (1998) MathReview

[19] **K Hikami**, *Difference equation of the colored Jones polynomial for torus knot*, Internat. J. Math. 15 (2004) 959–965 MathReview

[20] **T T Q Le**, *Integrality and symmetry of quantum link invariants*, Duke Math. J. 102 (2000) 273–306 MathReview

[21] **T T Q Le**, *The Colored Jones Polynomial and the A–Polynomial of Knots*, `arXiv:math.GT/0407521`

[22] **G Lusztig**, *Introduction to quantum groups*, volume 110 of *Progress in Mathematics*, Birkhäuser Boston Inc., Boston, MA (1993) MathReview

[23] **G Lusztig**, *Quantum groups at roots of* 1, Geom. Dedicata 35 (1990) 89–113 MathReview

[24] **B Malgrange**, *Équations différentielles à coefficients polynomiaux*, volume 96 of *Progress in Mathematics*, Birkhäuser Boston Inc., Boston, MA (1991) MathReview

[25] **G Masbaum**, *Skein-theoretical derivation of some formulas of Habiro*, Algebr. Geom. Topol. 3 (2003) 537–556 MathReview

[26] **P Paule**, **A Riese**, *A Mathematica q-analogue of Zeilberger's algorithm based on an algebraically motivated approach to q-hypergeometric telescoping*, from: "Special functions, q–series and related topics (Toronto, ON, 1995)", Fields Inst. Commun. 14, Amer. Math. Soc., Providence, RI (1997) 179–210 MathReview

[27] **P Paule**, **A Riese**, Mathematica software, available at: `http://www.risc.uni-linz.ac.at/research/combinat/risc/software/qZeil/`

[28] **M Petkovšek**, **H S Wilf**, **Doron Zeilberger**, $A = B$, A K Peters Ltd., Wellesley, MA (1996) MathReview

[29] **N Yu Reshetikhin**, **V G Turaev**, *Ribbon graphs and their invariants derived from quantum groups*, Comm. Math. Phys. 127 (1990) 1–26 MathReview

[30] **A Riese**, *qMultisum–A package for proving q–hypergeometric multiple summation identities*, preprint (2002)

[31] **C Sabbah**, *Systèmes holonomes d'équations aux q-différences*, from: "D–modules and microlocal geometry (Lisbon, 1990)", de Gruyter, Berlin (1993) 125–147 MathReview

[32] **V G Turaev**, *The Yang-Baxter equation and invariants of links*, Invent. Math. 92 (1988) 527–553 MathReview

[33] **H S Wilf**, **D Zeilberger**, *An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities*, Invent. Math. 108 (1992) 575–633 MathReview

[34] **L Yen**, *A two-line algorithm for proving q-hypergeometric identities*, J. Math. Anal. Appl. 213 (1997) 1–14 MathReview

[35] **D Zeilberger**, *A holonomic systems approach to special functions identities*, J. Comput. Appl. Math. 32 (1990) 321–368 MathReview

Geometry & Topology 17 (2013) 1253–1315

The quantum content of the gluing equations

TUDOR DIMOFTE
STAVROS GAROUFALIDIS

The gluing equations of a cusped hyperbolic 3–manifold M are a system of polynomial equations in the shapes of an ideal triangulation $\mathcal{T}$ of M that describe the complete hyperbolic structure of M and its deformations. Given a Neumann–Zagier datum (comprising the shapes together with the gluing equations in a particular canonical form) we define a formal power series with coefficients in the invariant trace field of M that should (a) agree with the asymptotic expansion of the Kashaev invariant to all orders, and (b) contain the nonabelian Reidemeister–Ray–Singer torsion of M as its first subleading "1–loop" term. As a case study, we prove topological invariance of the 1–loop part of the constructed series and extend it into a formal power series of rational functions on the $\mathrm{PSL}(2,\mathbb{C})$ character variety of M. We provide a computer implementation of the first three terms of the series using the standard `SnapPy` toolbox and check numerically the agreement of our torsion with the Reidemeister–Ray–Singer for all 59924 hyperbolic knots with at most 14 crossings. Finally, we explain how the definition of our series follows from the quantization of 3–dimensional hyperbolic geometry, using principles of topological quantum field theory. Our results have a straightforward extension to any 3–manifold M with torus boundary components (not necessarily hyperbolic) that admits a regular ideal triangulation with respect to some $\mathrm{PSL}(2,\mathbb{C})$ representation.

57M25, 57N10

1 Introduction

1.1 The Kashaev invariant and perturbative Chern–Simons theory

The *Kashaev invariant* $\langle K\rangle_N \in \mathbb{C}$ of a knot K in 3–space (for $N = 2, 3, \ldots$) is a powerful sequence of complex numbers determined by the *Jones polynomial* of the knot (see [46]) and its cablings; see Turaev [70] and Witten [72]. The *Volume Conjecture* of Kashaev and Murakami–Murakami [49; 50; 55] relates the Kashaev invariant of a hyperbolic knot K with the hyperbolic volume $\mathrm{Vol}(M)$ of its complement $M = S^3\setminus K$ (see Thurston [66]):

$$\lim_{N\to\infty}\frac{1}{N}\log|\langle K\rangle_N| = \frac{\mathrm{Vol}(M)}{2\pi}. \tag{1-1}$$

Published: 30 May 2013 DOI: 10.2140/gt.2013.17.1253

A generalization of the Volume Conjecture (see Gukov [39]) predicts a full asymptotic expansion of the Kashaev invariant to all orders in $1/N$:

$$\langle K \rangle_N \overset{N\to\infty}{\sim} \mathcal{Z}_M(2\pi i/N) \tag{1-2}$$

for a suitable formal power series

$$\mathcal{Z}_M(\hbar) = \exp\left(\frac{1}{\hbar} S_{M,0} - \frac{3}{2}\log\hbar + S_{M,1} + \sum_{n\geq 2} \hbar^{n-1} S_{M,n}\right), \qquad \hbar = \frac{2\pi i}{N}. \tag{1-3}$$

The formal power series $\mathcal{Z}_M(\hbar)$ in (1-3) is conjectured to coincide with the perturbative partition function of Chern–Simons theory with complex gauge group $\mathrm{SL}(2,\mathbb{C})$ along the discrete faithful representation ρ_0 of the hyperbolic manifold M. Combining such an interpretation with further conjectures of the first author, Gukov, Lenells and Zagier [17] and the second author and Lê [34; 36] one predicts the following.

- $S_{M,0} = i(\mathrm{Vol}_M + i\,\mathrm{CS}_M) \in \mathbb{C}/(4\pi^2\mathbb{Z})$ is the complexified volume of M (cf Thurston [67] and Neumann [56]).
- $S_{M,1}$ is related (see Witten [73], Bar-Natan and Witten [3] and Gukov and Murakami [40]) to the nonabelian Ray–Singer torsion (see De Loera, Rambau and Santos [63]), which ought to equal (cf Müller [54]) the combinatorial nonabelian Reidemeister torsion. More precisely, by Dubois and the second author [19, Conjecture 1.8] we should have

$$\tau_M^{\mathrm{R}} = 4\pi^3 \exp(-2S_{M,1}) \in E_M^*, \tag{1-4}$$

  where τ_M^{R} is the nonabelian Reidemeister–Ray–Singer torsion of M with respect to the meridian (see Porti [62] and Dubois [18]), and E_M is the invariant trace field of M.
- For $n \geq 2$, the *n–loop invariants* $S_{M,n}$ are conjectured to lie in the invariant trace field E_M [17; 34].

The generalization (1-2) of the Volume Conjecture has been numerically verified for a few knots using either state integral formulas for Chern–Simons theory when available [17] or a numerical computation of the Kashaev invariant and its numerical asymptotics, lifted to algebraic numbers; see the second author and Zagier [37; 31; 32].

Our goal is to provide an exact, combinatorial definition of the formal power series $\mathcal{Z}_M(\hbar)$ via formal Gaussian integration using the shape parameters and the Neumann–Zagier matrices of a regular ideal triangulation of M. Our definitions

- express the putative torsion $\exp(-2S_{M,1})$ and the n–loop invariants $S_{M,n}$ manifestly in terms of the shape parameters z_i and the gluing matrices of a regular ideal triangulation $\mathcal{T}$ of M;

- manifestly deduce that the putative torsion and the n–loop invariants for $n \geq 2$ are elements of the invariant trace field;
- explain the difference of $\mathcal{Z}_M(\hbar)$ for pairs of geometrically similar knots studied by Zagier and the second author;
- provide an effective way to compute the n–loop invariants using standard commands of the `SnapPy` toolbox [13], as demonstrated for $n = 1, 2, 3$ for hyperbolic knots with at most 14 crossings;
- allow efficient tests of the asymptotics of the Volume Conjecture (1-2), the "1–loop Conjecture" (1-4) and other conjectures in Quantum Topology.

We note that we only define $\exp(-2S_{M,1})$ up to a sign, and $S_{M,2}$ modulo $\mathbb{Z}/24$. All higher n–loop invariants are defined unambiguously.

Although we give a purely combinatorial definition of $\mathcal{Z}_M(\hbar)$ without any knowledge of state integrals or Chern–Simons theory with complex gauge group, in Section 5 we explain how our definition follows from the state integral model of the first author [15] and its perturbative expansion.

1.2 The Neumann–Zagier datum

All manifolds and all ideal triangulations in this paper will be oriented. The volume of a hyperbolic manifold M, appearing in the Volume Conjecture and contributing to $S_{M,0}$, is already known to have a simple expression in terms of shape parameters of a regular ideal triangulation, ie, one that recovers the complete hyperbolic structure of M. (For extended discussion on regular triangulations, see Section 4.) If $\mathcal{T} = \{\Delta_i\}_{i=1}^N$ is a regular ideal triangulation of M with shape parameters $z_i \in \mathbb{C} \setminus \{0, 1\}$ for $i = 1, \dots, N$, then (cf Dupont and Sah [23] and Neumann and Zagier [58])

$$\mathrm{Vol}(M) = \sum_{i=1}^{N} D(z_i), \tag{1-5}$$

where $D(z) := \mathrm{Im}(\mathrm{Li}_2(z)) + \arg(1-z)\log|z|$ is the Bloch–Wigner dilogarithm function. This formula can also be interpreted as calculating the image of the class $[M] := \sum_i [z_i] \in \mathcal{B}$ of M in the Bloch group $\mathcal{B}$ under the natural map $D\colon \mathcal{B} \to \mathbb{R}$. An analogous formula, using the class of M in the "extended" Bloch group $\widehat{\mathcal{B}}$, gives the full complexified volume $S_{M,0}$; see Neumann [56; 57], Goette and Zickert [38; 75].

It is natural to ask whether the class of M in $\mathcal{B}$ determines not only $S_{M,0}$ but the higher $S_{M,n}$ as well. This question was posed to the authors several years ago by D Zagier. Subsequent computations [37; 32] indicated that a positive answer was not possible.

For example, there is a family of pairs of pretzel knots $((-2,3,3+2p),(-2,3,3-2p))$ for $p=2,3,\dots$, as well as the figure-eight knot and its sister, which all have the same class in the Bloch group (and classes differing by 6–torsion in the extended Bloch group), but different invariants $S_{M,n}$ for $n\geq 1$.

The extra information necessary to determine the $S_{M,n}$ can be described as follows. Recall that if $\mathcal{T}$ is a regular ideal triangulation of M with N tetrahedra, its shapes $z=(z_1,\dots,z_N)$ satisfy a system of polynomial equations, one equation for every edge, and one imposing parabolic holonomy around the meridian of the cusp [66; 58]. Let us set

$$z_i'=(1-z_i)^{-1},\quad z_i''=1-z_i^{-1}. \tag{1-6}$$

The equations can then be written in the form

$$z^{\boldsymbol{A}}z''^{\boldsymbol{B}}:=\prod_{j=1}^{N} z_j^{\boldsymbol{A}_{ij}}(1-z_j^{-1})^{\boldsymbol{B}_{ij}}=\pm 1,\quad i=1,\dots,N, \tag{1-7}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are $N\times N$ square matrices with integer entries, which we call the *Neumann–Zagier matrices* following [58].

Definition 1.1 If $\mathcal{T}$ is a regular ideal triangulation of M, its *Neumann–Zagier datum* (resp. *enhanced Neumann–Zagier datum*) is given by the triple

$$\beta_{\mathcal{T}}=(z,\boldsymbol{A},\boldsymbol{B}),\quad \text{resp.}\quad \widehat{\beta}_{\mathcal{T}}=(z,\boldsymbol{A},\boldsymbol{B},f),$$

where z is a solution to the gluing equations and f is a combinatorial flattening of $\mathcal{T}$, a collection of integers that we define in Section 2.4.

As we will discuss in detail in Section 2, implicit in the above definition is the dependence of $\beta_{\mathcal{T}}$ and $\widehat{\beta}_{\mathcal{T}}$ on the following choices:

(1) a pair of opposite edges for every oriented ideal tetrahedron (a so-called choice of *quad type*).

(2) An edge of $\mathcal{T}$.

(3) A meridian loop in the boundary of M in general position with respect to $\mathcal{T}$.

(4) A combinatorial flattening.

1.3 The 1–loop invariant

Definition 1.2 Given a one-cusped hyperbolic manifold M with regular ideal triangulation $\mathcal{T}$ and enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}}$ we define

$$\tau_{\mathcal{T}} := \pm\frac{1}{2}\det(\boldsymbol{A}\,\Delta_{z''} + \boldsymbol{B}\,\Delta_z^{-1}) z^{f''} z''^{-f} \in E_M/\{\pm 1\}, \tag{1-8}$$

where $\Delta_z := \mathrm{diag}(z_1, \dots, z_N)$ and $\Delta_{z''} := \mathrm{diag}(z''_1, \dots, z''_N)$ are diagonal matrices, and $z^{f''} z''^{-f} := \prod_i z_i{}^{f''_i} z''^{-f_i}_i$.

Note that $\tau_{\mathcal{T}}$ takes value in the invariant trace field of M and is only defined up to a sign. In Section 3 we will show the following.

Theorem 1.3 *$\tau_{\mathcal{T}}$ is independent of the quad type of $\mathcal{T}$, the chosen edge of $\mathcal{T}$, the choice of a meridian loop, and the choice of a combinatorial flattening.*

We now consider the dependence of $\tau_{\mathcal{T}}$ on the choice of a regular ideal triangulation of M. It is well known that the set $\mathcal{X}$ of ideal triangulations of a cusped hyperbolic manifold is nonempty (see Casler [10]) and connected by 2–3 moves; see for example Matveev [52; 53] and Piergallini [59]. That is, a sequence of 2–3 moves can be used to take any one ideal triangulation to any other. The subset $\mathcal{X}_{\rho_0}$ of $\mathcal{X}$ of regular triangulations is also nonempty; see Section 4. Topologically, these are the triangulations without any univalent edges; see Champanerkar [11], Boyd, Dunfield and Rodriguez-Villegas [7], Dunfield and the second author [22] and Tillmann [69]. We will prove the following in Section 3.

Theorem 1.4 *$\tau_{\mathcal{T}}$ is constant on every component of $\mathcal{X}_{\rho_0}$ connected by 2–3 moves.*

1.4 Expectations

We may pose some questions and conjectures about the 1–loop invariant $\tau_{\mathcal{T}}$ and the structure of the set $\mathcal{X}_{\rho_0}$. Let us begin with two questions whose answers are unfortunately unknown.

Question 1.5 Is $\mathcal{X}_{\rho_0}$ connected by 2–3 moves?

Question 1.6 Is $\tau_{\mathcal{T}}$ constant on the set $\mathcal{X}_{\rho_0}$?

Clearly, a positive answer to the first question implies a positive answer to the second.

Despite the unknown answer to the above questions, with additional effort we can still define a distinguished component of $\mathcal{X}_{\rho_0}$, and thus obtain a topological invariant of M. Namely, let $\mathcal{X}_M^{\mathrm{EP}} \subset \mathcal{X}_{\rho_0}$ denote the subset that consists of regular refinements of the canonical (Epstein–Penner) ideal cell decomposition of M [25]. $\mathcal{X}_M^{\mathrm{EP}}$ is canonically associated to a cusped hyperbolic manifold M. A detailed description of $\mathcal{X}_M^{\mathrm{EP}}$ is given by the second author, Hodgson, Rubinstein and Segerman in [35, Section 6]. $\mathcal{X}_M^{\mathrm{EP}}$ generically consists of a single element. In Section 4.2 we will show the following.

Proposition 1.7 *$\mathcal{X}_M^{\mathrm{EP}}$ lies in a connected component of $\mathcal{X}_{\rho_0}$. Consequently, the value of $\tau_{\mathcal{T}}$ on $\mathcal{X}_M^{\mathrm{EP}}$ is a topological invariant τ_M of M.*

Admittedly, it would be more natural to show that $\tau_{\mathcal{T}}$ is constant on all of $\mathcal{X}_{\rho_0}$. Proposition 1.7 appears to be an artificial way to construct a much needed topological invariant of cusped hyperbolic 3–manifolds.

Our next conjecture compares our torsion τ_M with the nonabelian Reidemeister torsion τ_M^{R} of M with respect to the meridian defined in [62; 18].

Conjecture 1.8 *For all hyperbolic knot complements we have $\tau_M^{\mathrm{R}} = \pm\tau_M$.*

Numerical evidence for the above conjecture is presented in Appendix D using Dunfield's computation of τ_M^{R} via `SnapPy` [21]. Observe that both sides of the equation in Conjecture 1.8 are algebraic numbers (defined up to a sign) that are elements of the invariant trace field of M. Moreover, if M has a regular ideal triangulation with N ideal tetrahedra and its fundamental group is generated with r elements, then τ_M and τ_M^{R} are essentially given by the determinant of square matrices of size N and $3r-3$, respectively. It is still unclear to us how to relate these two matrices or their determinants.

By definition, $\tau_M^{\mathrm{R}} \in E_M^*$. Thus, a mild but important corollary of Conjecture 1.8 is that τ_M is nonzero. This is a crucial ingredient, necessary for the definition of the higher loop invariants $S_{M,n}$ using perturbation theory.

1.5 The higher-loop invariants

In this section we define the higher loop invariants $S_{\mathcal{T},n}$ for $n \geq 2$. They are analyzed in detail in Section 5, using a state integral (5-2). The result, however, may be summarized as follows. Let us introduce a formal power series

$$\psi_\hbar(x;z) = \exp\Bigg(\sum_{n,k,2n+k-2>0} \frac{\hbar^{n+k/2-1}(-x)^k B_n}{n!k!}\mathrm{Li}_{2-n-k}(z^{-1})\Bigg) \in \mathbb{Q}(z)[\![x,\hbar^{1/2}]\!], \tag{1-9}$$

where B_n is the n^{th} *Bernoulli number* (with $B_1 = +1/2$), and $L_m(z)$ is the m^{th} polylogarithm. Note that $\mathrm{Li}_m(z) \in (1-z)^{-m-1}\mathbb{Z}[z]$ is a rational function for all nonpositive integers m. This formal series comes from the asymptotic expansion of the quantum dilogarithm function after removal of its two leading asymptotic terms; see Barnes [4], Faddeev and Kashaev [28; 27]. The quantum dilogarithm is the Chern–Simons partition function of a single tetrahedron and its asymptotics are studied in detail in Section 5.

We fix an enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$ of an oriented 1–cusped manifold M and a regular ideal triangulation $\mathcal{T}$ with N tetrahedra. Let $\nu = \boldsymbol{A} f + \boldsymbol{B} f''$. We assume that

$$\det(\boldsymbol{B}) \neq 0, \quad \tau_M \neq 0.$$

The condition $\det(\boldsymbol{B}) \neq 0$ is always satisfied with a suitable labeling of shapes; see Lemma A.3. In that case, Lemma A.2 implies that

$$\mathcal{H} = -\boldsymbol{B}^{-1}\boldsymbol{A} + \Delta_{z'}, \tag{1-10}$$

is a symmetric matrix, where $\Delta_{z'} = \mathrm{diag}(z'_1, \ldots, z'_N)$. We define

$$f_{\mathcal{T},\hbar}(x;z) = \exp\Big(-\frac{\hbar^{1/2}}{2} x^T \boldsymbol{B}^{-1}\nu + \frac{\hbar}{8} f^T \boldsymbol{B}^{-1}\boldsymbol{A} f\Big) \prod_{i=1}^{N} \psi_\hbar(x_i, z_i) \in \mathbb{Q}(z)[\![x, \hbar^{1/2}]\!], \tag{1-11}$$

where $x = (x_1, \ldots, x_N)^T$ and $z = (z_1, \ldots, z_N)$. Assuming that $\mathcal{H}$ is invertible, a formal power series $f_\hbar(x) \in \mathbb{Q}(z)[\![x, \hbar^{1/2}]\!]$ has a *formal Gaussian integration*, given by (cf Bessis, Itzykson and Zuber [6]),

$$\langle f_\hbar(x) \rangle = \frac{\int dx e^{-1/2 x^T \mathcal{H} x} f_\hbar(x)}{\int dx e^{-1/2 x^T \mathcal{H} x}}. \tag{1-12}$$

This integration is defined by expanding $f_\hbar(x)$ as a series in x, and then formally integrating each monomial, using the quadratic form $\mathcal{H}^{-1}$ to contract x–indices pairwise.

Definition 1.9 With the above conventions, we define

$$\exp\Big(\sum_{n=2}^{\infty} S_{\mathcal{T},n}(z)\,\hbar^{n-1}\Big) := \langle f_{\mathcal{T},\hbar}(x;z) \rangle. \tag{1-13}$$

Remark 1.10 Notice that the result involves only integral powers of $\hbar$ and each term is a rational function in the complex numbers z. Moreover, $S_{\mathcal{T},n} \in \tau_{\mathcal{T}}^{-3n+3}\mathbb{Q}[z, z', z'']$

for all $n \geq 2$. This follows from the fact that the connected Feynman diagrams that contribute to $S_{\mathcal{T},n}$ have at most $3n-3$ edges and each edge (contracted by $\mathcal{H}^{-1}$) contributes a factor of $\det(\mathcal{H})^{-1}$. Thus, we can also write

$$\exp\left(\sum_{n=2}^{\infty} S_{\mathcal{T},n}\hbar^{n-1}\right) = 1 + \sum_{n=1}^{\infty} \frac{\widetilde{S}_{\mathcal{T},n}}{\tau_{\mathcal{T}}^{3n}}\,\hbar^n, \tag{1-14}$$

where $\widetilde{S}_{\mathcal{T},n} \in \mathbb{Q}[z,z',z'']$. Experimentally, it appears that $\widetilde{S}_{\mathcal{T},n}$ have lower complexity than $S_{\mathcal{T},n}$; see Appendix D.

1.6 Feynman diagrams

A convenient way to organize the above definition is via Feynman diagrams, using Wick's Theorem to express each term $S_{\mathcal{T},n}$ as a finite sum of connected diagrams with at most n loops, where the number of loops of a connected graph is its first Betti number. This is well known and explained in detail, eg by Hori, Katz, Klemm, Pandharipande, Thomas, Vafa, Vakil and Zaslow in [45, Chapter 9], by Polyak in [60] and in [6].

The Feynman rules for computing the $S_{\mathcal{T},n}$, described in Section 5, turn out to be the following.[1] One draws all connected graphs D with vertices of all valencies, such that

$$L(D) := (\#1\text{–vertices} + \#2\text{–vertices} + \#\text{loops}) \leq n. \tag{1-15}$$

In each diagram, the edges represent an $N \times N$ propagator

$$\text{propagator}: \quad \Pi = \hbar\mathcal{H}^{-1}, \tag{1-16}$$

while each k–vertex comes with an N–vector of factors $\Gamma_i^{(k)}$,

$$\Gamma_i^{(k)} = (-1)^k \sum_{p=\alpha_k}^{\alpha_k+n-L(D)} \frac{\hbar^{p-1}B_p}{p!}\mathrm{Li}_{2-p-k}(z_i^{-1}) + \begin{cases} -\frac{1}{2}(\boldsymbol{B}^{-1}\nu)_i & k=1, \\ 0 & k \geq 2, \end{cases} \tag{1-17}$$

where $\alpha_k = 1$ (resp. 0) if $k = 1,2$ (resp. $k \geq 3$). The diagram D is then evaluated by contracting the *vertex factors* $\Gamma_i^{(k)}$ with propagators, multiplying by a standard *symmetry factor*, and taking the $\hbar^{n-1}$ part of the answer. In the end, $S_{M,n}$ is the sum of evaluated diagrams, plus an additional *vacuum* contribution

$$\Gamma^{(0)} = \frac{B_n}{n!}\sum_{i=1}^{N}\mathrm{Li}_{2-n}(z_i^{-1}) + \begin{cases} \frac{1}{8} f\cdot \boldsymbol{B}^{-1}\boldsymbol{A} f & n=2, \\ 0 & n\geq 3. \end{cases} \tag{1-18}$$

To illustrate the above algorithm, we give the explicit formulas for S_2 and S_3 below.

[1] To derive these from (1-12), one should first rescale $x \to \hbar^{-1/2}x$.

1.7 The 2–loop invariant

The six diagrams that contribute to $S_{M,2}$ are shown in Figure 1, together with their symmetry factors.

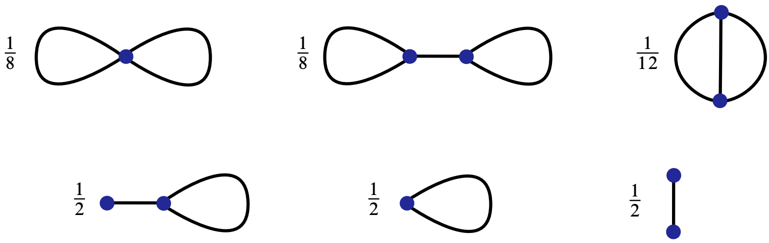


Figure 1: Diagrams contributing to $S_{M,2}$ with symmetry factors. The top row of diagrams have exactly two loops, while the bottom row have fewer loops and additional 1–vertices and 2–vertices.

Their evaluation gives the following formula for $S_{\mathcal{T},2}$:

$$S_{\mathcal{T},2} = \operatorname{coeff}\Big[\frac{1}{8}\Gamma_i^{(4)}(\Pi_{ii})^2 + \frac{1}{8}\Pi_{ii}\Gamma_i^{(3)}\Pi_{ij}\Gamma_j^{(3)}\Pi_{jj} + \frac{1}{12}\Gamma_i^{(3)}(\Pi_{ij})^3\Gamma_j^{(3)} + \frac{1}{2}\Gamma_i^{(1)}\Pi_{ij}\Gamma_j^{(3)}\Pi_{jj} + \frac{1}{2}\Gamma_i^{(2)}\Pi_{ii} + \frac{1}{2}\Gamma_i^{(1)}\Pi_{ij}\Gamma_j^{(1)}, \hbar\Big] + \Gamma^{(0)}, \tag{1-19}$$

where all the indices i and j are implicitly summed from 1 to N and $\operatorname{coeff}[f(\hbar),\hbar]$ denotes the coefficient of $\hbar$ of a power series $f(\hbar)$. Concretely, the 2–loop contribution from the vacuum energy is $\Gamma^{(0)} = \frac{1}{8} f^T \boldsymbol{B}^{-1}\boldsymbol{A} f - \frac{1}{12}\sum_i z_i'$, and the four vertices that appear only contribute at leading order,

$$\Gamma_i^{(1)} = \frac{z_i' - (\boldsymbol{B}^{-1}\nu)_i}{2}, \qquad \Gamma_i^{(2)} = \frac{z_i z_i'^2}{2}, \qquad \Gamma_i^{(3)} = -\frac{z_i z_i'^2}{\hbar}, \qquad \Gamma_i^{(4)} = -\frac{z_i(1+z_i)z_i'^3}{\hbar}. \tag{1-20}$$

We expect $S_{\mathcal{T},2}$ to be well-defined modulo $\mathbb{Z}/24$, and this is exactly what happens in hundreds of examples that we computed.

1.8 The 3–loop invariant

For the next invariant $S_{\mathcal{T},3}$, all the diagrams of Figure 1 contribute, collecting the coefficient of $\hbar^2$ of their evaluation. In addition, there are 34 new diagrams that satisfy the inequality (1-15); they are shown in Figures 2 and 3. Calculations indicate that the 3–loop invariant $S_{\mathcal{T},3}$ is well-defined, independent of the regular triangulation $\mathcal{T}$. The invariants $S_{\mathcal{T},0}, \tau_{\mathcal{T}}, S_{\mathcal{T},2}, S_{\mathcal{T},3}$ have been programmed in `Mathematica` and take as input a Neumann–Zagier datum available from `SnapPy` [13].

For the 4–loop invariant, there are 291 new diagrams. A `python` implementation will be provided in the future. For large n, one expects about $n!^2 C^n$ diagrams to contribute to S_n.

Remark 1.11 Note that the n–loop invariant for $n \geq 3$ is independent of the combinatorial flattening and in fact depends only on $(\boldsymbol{B}^{-1}\boldsymbol{A}, \boldsymbol{B}^{-1}\nu, z)$.

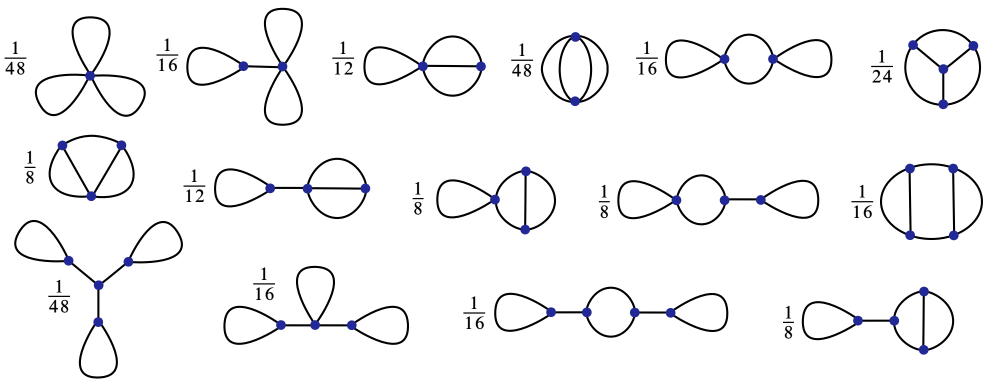


Figure 2: Diagrams with three loops contributing to S_3

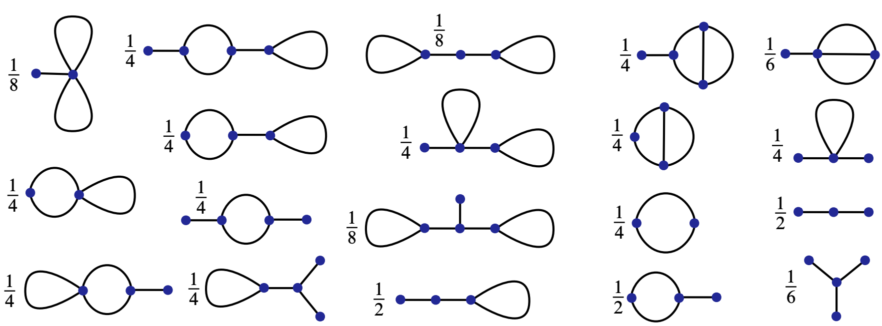


Figure 3: Diagrams with 1–vertices and 2–vertices contributing to S_3

1.9 The Feynman diagrams are stable graphs

During a Master's Class in Aarhus in February 2013, the second author observed that the Feynman graphs of our paper can be identified with the stable graphs which appear in the *Topological Recursion* of Eynard–Orantin and the graphs that appear in the intersection theory of the moduli space of curves; see [26] and Aganagic, Bouchard and Klemm [1, Figure 1]. We thank Bertrand Eynard and Nicolas Orantin for delivering the Master's Class in Aarhus and Jorgen Andersen for organizing it.

Definition 1.12 A *stable graph* G is an abstract connected graph (with no cyclic order of the edges around a vertex) with the property that every vertex v of G is attached a genus g_v and a degree (ie, valency) n_v such that $2g_v - 2 + n_v > 0$. The total degree of a stable graph is given by $\sum_v (2g_v - 2 + n_v) = 2g - 2 + n$ where n is the number of external legs and g is defined to be the genus of G.

Let $\mathcal{G}(g,n)$ denote the (finite) set of stable graphs of genus g with n external legs. Then

$$S_g = \sum_{G \in \mathcal{G}(g,0)} \frac{1}{|\mathrm{Aut}(G)|} \langle G \rangle.$$

To explain where the genus comes in our Feynman graphs, observe that $\Gamma_i^{(k)}$ from Equation (1-17) for each vertex v has $k = n_v$, and $2g_v = p$ that contributes to $\hbar^{2g_v - 1} B_{2g_v}/(2g_v)! \mathrm{Li}_{2-2g_v-n_v}$. In other words, $2g_v$ extracts the monomial h^{p-1} from $\Gamma^{(k)}$. Notice that since $B_{\mathrm{odd}} = 0$ (for odd greater than 1) then we must have $p = \text{even} = 2g_v$.

1.10 Generalizations

There are several natural extensions of the results presented above. First, one could attempt to prove the independence of the all-loop invariants $Z_{\mathcal{T}}(\hbar)$, including the entire series of $S_{\mathcal{T},n}$s, under 2–3 moves and different choices of Neumann–Zagier datum. This was done nonrigorously in [15], but a full mathematical argument in the spirit of Theorems 1.3 and 1.4 is still missing. We hope to address this in future work.

In a different direction, one can extend the formulas for $\tau_{\mathcal{T}}$ and $S_{\mathcal{T},n}$ to

- manifolds with multiple cusps,
- representations other than the discrete faithful,
- representations with nonparabolic meridian holonomy,
- nonhyperbolic manifolds.

The only truly necessary condition is that a 3–manifold M have a topological ideal triangulation $\mathcal{T}$ that — upon solving gluing equations and using a developing map — reproduces some desired representation $\rho\colon \pi_1(M) \to \mathrm{PSL}(2,\mathbb{C})$. We call such an ideal triangulation ρ–regular, and in Section 4 we will briefly discuss most of the above generalizations. In particular, we will demonstrate in Sections 4.6 and 5.5 how to extend $\tau_{\mathcal{T}}, S_{\mathcal{T},n}$ to rational functions on the character variety of a (topologically) cusped manifold. The generalization to multiple cusps is also quite straightforward, but left out mainly for simplicity of exposition.

Acknowledgments The authors wish to thank Nathan Dunfield, Walter Neumann, Josephine Yu and Don Zagier for many extremely enlightening conversations.

The work of S G is supported by NSF grant DMS-11-05678. The work of T D is supported primarily by the Friends of the Institute for Advanced Study, and in part by DOE grant DE-FG02-90ER40542.

2 Mechanics of triangulations

We begin by reviewing the gluing rules for ideal hyperbolic tetrahedra and the equations that determine their shape parameters. We essentially follow the classic [66; 58], but find it helpful to work with additive logarithmic (rather than multiplicative) forms of the gluing equations. Recall that all manifolds and all ideal triangulations are *oriented*.

2.1 Ideal tetrahedra

Combinatorially, an *oriented ideal tetrahedron* Δ is a topological ideal tetrahedron with three complex *shape parameters* (z, z', z'') assigned to pairs of opposite edges (Figure 4). The shapes always appear in the same cyclic order (determined by the orientation) around every vertex, and they satisfy

(2-1a) $$zz'z'' = -1,$$

(2-1b) $$z'' + z^{-1} - 1 = 0.$$

In other words, $z' = 1/(1-z)$ and $z'' = 1 - z^{-1}$. We call the tetrahedron *nondegenerate* if none of the shapes take values in $\{0, 1, \infty\}$, ie, $z, z', z'' \in \mathbb{C}^*\backslash\{1\}$. It is sufficient to impose this on a single one of the shapes.

Borrowing common terminology from the theory of normal surfaces, cf Burton [8], Kang and Rubenstein [47; 48] and Tillman [68], we define the *quadrilateral type* (in short, *quad type*) of Δ to be the distinguished pair of opposite edges labelled by z.

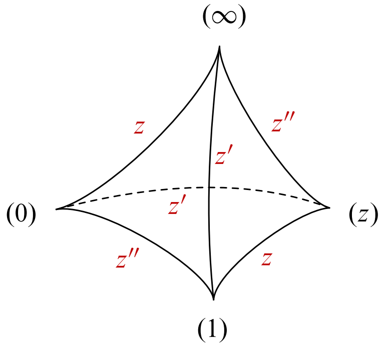


Figure 4: An ideal tetrahedron

Clearly, there is a threefold choice of quad type for any oriented ideal tetrahedron. Different choices correspond to a cyclic permutation of the vector $(z, z,', z'')$, which leaves relations (2-1a) invariant.

Geometrically, the shape parameters determine a $\mathrm{PSL}(2, \mathbb{C})$ structure on Δ. Equivalently, they determine a hyperbolic structure, possibly of negative volume. We can then describe the ideal hyperbolic tetrahedron Δ as the convex hull of four ideal points in hyperbolic three-space $\mathbb{H}^3$, whose cross ratio is z (or z', or z''). Each shape z fixes the complexified dihedral angle on the edge it labels, via

$$z = \exp(\text{torsion} + i\,\text{angle}), \tag{2-2}$$

and similarly for z', z''.

2.2 The gluing matrices

We now discuss an important combinatorial invariant of ideal triangulations, namely the gluing and Neumann–Zagier matrices, their symplectic properties, and the notion of a combinatorial flattening. Although these notions are motivated by hyperbolic geometry (namely the gluing of ideal tetrahedra around their faces and edges to describe a complete hyperbolic structure on a cusped manifold), we stress that these notions make sense for arbitrary 3–manifolds with torus boundary, and for triangulations whose gluing equations may not have solutions in $\mathbb{C} \setminus \{0, 1\}$.

Let M be an oriented one-cusped manifold with an ideal triangulation $\mathcal{T} = \{\Delta_i\}_{i=1}^N$ and a choice of quad type.

The choice of quad, combined with the orientation of $\mathcal{T}$ and M allow us to attach variables (Z_i, Z_i', Z_i'') to each tetrahedron Δ_i. An Euler characteristic argument shows that the triangulation has N edges E_I, $I = 1, \dots, N$. For each edge E_I we introduce a gluing equation of the form

$$(2\text{-}3) \qquad E_I: \quad \sum_{i=1}^{N} (\boldsymbol{G}_{Ii} Z_i + \boldsymbol{G}'_{Ii} Z'_i + \boldsymbol{G}''_{Ii} Z''_i) = 2\pi i, \quad I = 1, \dots, N,$$

where $\boldsymbol{G}_{Ii} \in \{0, 1, 2\}$ (resp., $\boldsymbol{G}'_{Ii}$, $\boldsymbol{G}''_{Ii}$) is the number of times an edge of tetrahedron Δ_i with parameter Z_i (resp., Z'_i, Z''_i) is incident to the edge E_I in the triangulation. In addition, we impose the equations

$$(2\text{-}4) \qquad Z_i + Z'_i + Z''_i = i\pi,$$

for $i = 1, \dots, N$. Equations (2-3) are not all independent. For a one-cusped manifold, every edge begins and ends at the cusp, which implies $\sum_{I=1}^{N} \boldsymbol{G}_{Ii} = \sum_{I=1}^{N} \boldsymbol{G}'_{Ii} = \sum_{I=1}^{N} \boldsymbol{G}''_{Ii} = 2$, and therefore that the sum of the left-hand sides of Equations (2-3) equals $2\pi i N$. This is the only linear dependence in case of one cusp. In general, there is one relation per cusp of M, as follows from [56, Theorem 4.1].

An oriented peripheral simple closed curve μ (such as a meridian) on the boundary of M, in general position with the triangulation of the boundary torus that comes from $\mathcal{T}$, also gives rise to a gluing equation. We assume that the curve is simple (has no self intersections), and set the signed sum of edge parameters on the dihedral angles subtended by the curve to zero. A parameter is counted with a plus sign (resp. minus sign) if an angle is subtended in a counterclockwise (resp. clockwise) direction as viewed from the boundary. These rules are the same as described in [56], and demonstrated in Section 2.6 below.

Let us choose such a peripheral curve μ. We choose a meridian if M is a knot complement. The gluing equation associated to μ then takes the form

$$(2\text{-}5) \qquad \mu: \quad \sum_{i=1}^{N} (\boldsymbol{G}_{N+1,i} Z_i + \boldsymbol{G}'_{N+1,i} Z'_i + \boldsymbol{G}''_{N+1,i} Z''_i) = 0,$$

with $\boldsymbol{G}_{N+1,i}, \boldsymbol{G}'_{N+1,i}, \boldsymbol{G}''_{N+1,i} \in \mathbb{Z}$.

2.3 The Neumann–Zagier matrices

The matrices $\boldsymbol{G}$, $\boldsymbol{G}'$ and $\boldsymbol{G}''$ have both symmetry and redundancy. We have already observed that any one of the edge constraints (2-3) can be removed. Let us then ignore the edge $I = N$. We can also use (2-4) to eliminate one of the three shapes for each tetrahedron. We choose this canonically to be Z'_i, though which pair of edges is labelled Z'_i depends on the choice of quad type for the tetrahedron. Then the first $N-1$ edge equations and the meridian (μ) equation are equivalent to

$$\sum_{j=1}^{N}(\boldsymbol{A}_{ij} Z_j + \boldsymbol{B}_{ij} Z''_j) = i\pi\nu_i, \quad i = 1, \ldots, N, \tag{2-6}$$

where

$$\begin{aligned} \boldsymbol{A}_{ij} &= \begin{cases} \boldsymbol{G}_{ij} - \boldsymbol{G}'_{ij} & I \neq N, \\ \boldsymbol{G}_{N+1,j} - \boldsymbol{G}'_{N+1,j} & i = N, \end{cases} \\ \boldsymbol{B}_{ij} &= \begin{cases} \boldsymbol{G}''_{ij} - \boldsymbol{G}'_{ij} & I \neq N, \\ \boldsymbol{G}''_{N+1,j} - \boldsymbol{G}'_{N+1,j} & i = N, \end{cases} \end{aligned} \tag{2-7}$$

$$\nu_i := \begin{cases} 2 - \sum_{j=1}^{N} \boldsymbol{G}'_{ij} & i \neq N, \\ -\sum_{j=1}^{N} \boldsymbol{G}'_{N+1,j} & i = N. \end{cases} \tag{2-8}$$

We will generally assume Z, Z'' and ν to be column vectors, and we will write $\boldsymbol{A}Z + \boldsymbol{B}Z'' = i\pi\nu$. The matrices $(\boldsymbol{G}, \boldsymbol{G}', \boldsymbol{G}'')$ as well as $(\boldsymbol{A}, \boldsymbol{B}, \nu)$ can easily be obtained from `SnapPy` [13], as is illustrated in Appendix D.

The Neumann–Zagier matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have a remarkable property: they are the top two blocks of a $2N \times 2N$ symplectic matrix [58]. It follows that

$$\boldsymbol{A}\boldsymbol{B}^T = \boldsymbol{B}\boldsymbol{A}^T, \tag{2-9}$$

and that the $N \times 2N$ block $(\boldsymbol{A}\boldsymbol{B})$ has full rank. This symplectic property is crucial for defining the state integral of [15], for defining our formal power series invariant $\mathcal{Z}_M(\hbar)$, and for the combinatorial proofs of topological invariance of the 1–loop invariant. A detailed discussion of the symplectic properties of the Neumann–Zagier matrices $\boldsymbol{A}, \boldsymbol{B}$ is given in Appendix A.

2.4 Combinatorial flattenings

We now have all ingredients to define what is a combinatorial flattening.

Definition 2.1 Given an ideal triangulation $\mathcal{T}$ of M, a *combinatorial flattening* is a collection of $3N$ integers $(f_i, f'_i, f''_i) \in \mathbb{Z}^3$ for $i = 1, \ldots, N$ that satisfy

$$f_i + f'_i + f''_i = 1, \quad i = 1, \ldots, N, \tag{2-10a}$$

$$\sum_{i=1}^{N}(\boldsymbol{G}_{Ii} f_i + \boldsymbol{G}'_{Ii} f'_i + \boldsymbol{G}''_{Ii} f''_i) = \begin{cases} 2 & I = 1, \ldots, N, \\ 0 & I = N+1. \end{cases} \tag{2-10b}$$

Note that if we eliminate f' using Equation (2-10a), a flattening is a pair of vectors $(f, f'') \in \mathbb{Z}^{2N}$ that satisfies

$$\boldsymbol{A} f + \boldsymbol{B} f'' = \nu. \tag{2-11}$$

Evidently, Equation (2-11) is a system of linear Diophantine equations. Neumann proved in [56, Lemma 6.1] that every ideal triangulation $\mathcal{T}$ has a flattening.

Remark 2.2 Combinatorial flattenings should not be confused with the (geometric) flattenings of [57, Definition 3.1]. The latter flattenings are coherent choices of logarithms for the shape parameters z, z', z'' of a complex solution to the gluing equations. On the other hand, our combinatorial flattenings are independent of a solution to the gluing equations. In the rest of the paper, the term *flattening* will mean a *combinatorial flattening* in the sense of Definition 2.1.

2.5 The shape solutions to the gluing equations

If we exponentiate the equations (2-4), and set $(z_i, z_i', z_i'') = (e^{Z_i}, e^{Z_i'}, e^{Z_i''})$, we obtain that (z_i, z_i', z_i'') satisfy Equation (2-1a). If we combine the exponentiated equations (2-6) with the nonlinear relation (2-1b) for each tetrahedron, we obtain

$$z^{\boldsymbol{A}} z''^{\boldsymbol{B}} = z^{\boldsymbol{A}} (1 - z^{-1})^{\boldsymbol{B}} = (-1)^{\nu}, \tag{2-12}$$

where $z^{\boldsymbol{A}} := \prod_j z_j^{\boldsymbol{A}_{ij}}$. These N equations in N variables are just the gluing equations of Thurston [66] and Neumann and Zagier [58], and fully capture the constraints imposed by the gluing. For hyperbolic M, a triangulation $\mathcal{T}$ is regular precisely when one of the solutions to (2-12) corresponds to the complete hyperbolic structure.

2.6 Example: 4_1

As an example, we describe the enhanced Neumann–Zagier datum of the figure-eight knot complement M. It has a well known regular ideal triangulation $\mathcal{T}$ consisting of $N = 2$ tetrahedra, to which we assign logarithmic shape parameters (Z, Z', Z'') and (W, W', W'').

A map of the boundary of the cusp neighborhood is shown in Figure 5. We have chosen one of 3^2 possible cyclic labelings by Zs and Ws (ie one of 3^2 possible quad types). Each of the edges intersects the cusp twice, so it is easy to read off from Figure 5 that the edge constraints (2-3) are

$$\begin{aligned} E_1 &: \quad 2Z + Z'' + 2W + W'' = 2\pi i, \\ E_2 &: \quad 2Z' + Z'' + 2W' + W'' = 2\pi i. \end{aligned}$$

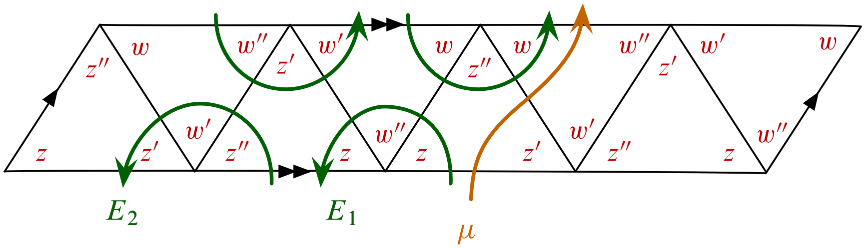


Figure 5: The boundary of the cusp neighborhood for the figure-eight knot

The sum of the left-hand sides is automatically $4\pi i$, so we can choose to ignore the second constraint. If we choose the meridian path μ as in Figure 5, the meridian constraint (2-5) is

$$\mu : \quad -Z' + W = 0.$$

Putting together the first edge constraint and the meridian into matrices, we have

$$\begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Z \\ W \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} Z' \\ W' \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Z'' \\ W'' \end{pmatrix} = i\pi \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Using $Z + Z' + Z'' = W + W' + W'' = i\pi$ to eliminate Z' and W', we get

$$\begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} Z \\ W \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Z'' \\ W'' \end{pmatrix} = i\pi \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

From this last expression, we can read off

$$\boldsymbol{A} = \begin{pmatrix} 2 & 2 \\ 1 & 1 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \nu = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{2-13}$$

The two gluing equations (2-12) are then

$$z^2 w^2 z'' w'' = 1, \quad z w z'' = -1, \tag{2-14}$$

with $z'' = 1 - z^{-1}$ and $w'' = 1 - w^{-1}$. The solution for the complete hyperbolic structure is $z = w = e^{i\pi/3}$.

Finally, a flattening $(f_z, f_z', f_z''; f_w, f_w', f_w'') \in \mathbb{Z}^6$, ie an integer solution to the equations $\boldsymbol{A} f + \boldsymbol{B} f'' = \nu$ and $f + f' + f'' = 1$, is given by

$$(f_z, f_z', f_z''; f_w, f_w', f_w'') = (0, 1, 0; 1, 0, 0). \tag{2-15}$$

It is easy to see that every flattening has the form $(a, b, 1-a-b; b, a, 1-a-b)$ for integers a, b.

3 Topological invariance of our torsion

Given a one-cusped hyperbolic manifold M with regular triangulation $\mathcal{T} = \{\Delta_i\}_{i=1}^N$ and Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$, we have proposed the nonabelian torsion is given by

$$\tau_{\mathcal{T}} := \pm\frac{1}{2}\det(\boldsymbol{A}\,\Delta_{z''} + \boldsymbol{B}\,\Delta_z^{-1}) z^{f''} z''^{-f} \in E_M/\{\pm 1\}, \tag{3-1}$$

where $\Delta_z = \operatorname{diag}(z_1, \dots, z_N)$, and similarly for $\Delta_{z''}$. Since $(z, z', z'') \in E_M$ we must have $\tau_{\mathcal{T}} \in E_M$ as well.

After a brief example of how the formula (3-1) works, we will proceed to prove Theorems 1.3 and 1.4 on the topological invariance of $\tau_{\mathcal{T}}$. We saw in Section 2 that the Neumann–Zagier datum depends not only on a triangulation $\mathcal{T}$, but also on a choice of

(1) quad type for $\mathcal{T}$,

(2) one edge of $\mathcal{T}$ whose gluing equation is redundant,

(3) normal meridian path,

(4) flattening f.

We will begin by showing $\tau_{\mathcal{T}}$ is independent of these four choices, and then show it is invariant under 2–3 moves, so long as the 2–3 moves connect two regular triangulations.

The four choices here are independent, and can be studied in any order. However, in order to prove independence of flattening, it is convenient to use a quad type for which the matrix $\boldsymbol{B}$ is nondegenerate. Such a quad type can always be found (Lemma A.3), but is not automatic. Therefore, we will first show invariance under change of quad type, and then proceed to the other choices. It is interesting to note that of all the arguments that follow (including the 2–3 move), independence of flattening is the only one that requires the use of the full gluing equations $z^{\boldsymbol{A}} z''^{\boldsymbol{B}} = (-1)^{\nu}$.

3.1 Example: 4_1 continued

To illustrate the Equation (1-8), consider the figure-eight knot complement again. From Section 2.6, we already have one possible choice for the Neumann–Zagier matrices (2-13) and a generic flattening (2-15). We use them to obtain

$$\begin{aligned}
\pm\tau_{\mathbf{4_1}} &= \frac{1}{2}\det\left[\begin{pmatrix}2&2\\1&1\end{pmatrix}\begin{pmatrix}z''&0\\0&w''\end{pmatrix}+\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}z^{-1}&0\\0&w^{-1}\end{pmatrix}\right]w''^{-1}\\
&= \frac{1}{2w''}\det\begin{pmatrix}2z''+z^{-1}&2w''+w^{-1}\\z''+z^{-1}&w''\end{pmatrix}\\
&= \frac{1}{2w''}\det\begin{pmatrix}z''+1&w''+1\\1&w''\end{pmatrix}\\
&= \frac{1}{2}(z''-w''^{-1}) = \frac{1}{2}\sqrt{-3},
\end{aligned} \tag{3-2}$$

where at intermediate steps we used $z''+z^{-1}-1=w''+w^{-1}-1=0$, and at the end we substituted the discrete faithful solution $z=z''=w=w''=e^{i\pi/3}$.

The invariant $\tau_{\mathbf{4_1}}$ belongs to the invariant trace field $E_{\mathbf{4_1}}=\mathbb{Q}(\sqrt{-3})$, and agrees with the torsion of the figure-eight knot complement [21].

3.2 Independence of a choice of quad type

Now, let us fix a manifold M, a triangulation $\mathcal{T}$ with N tetrahedra, and an enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}}=(z,\boldsymbol{A},\boldsymbol{B},f)$.

To prove independence of quad type, it is sufficient to check that $\tau_{\mathcal{T}}$ is invariant under a cyclic permutation of the first triple of shape parameters (z_1,z_1',z_1''), while holding fixed the choice of meridian loop and redundant edge. Let us write $z=(z_1,\dots,z_N)$, $\nu=(\nu_1,\dots,\nu_N)^T$, $f=(f_1,\dots,f_N)^T$ and

$$\boldsymbol{A}=(a_1,a_2,\dots,a_N),\quad \boldsymbol{B}=(b_1,b_2,\dots,b_N), \tag{3-3}$$

in column notation. After the permutation, a new Neumann–Zagier datum is given by $(\widetilde{z},\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{B}},\widetilde{f})$ where

$$\widetilde{z}=(z_1',z_2,\dots,z_N),\quad \widetilde{z}'=(z_1'',z_2',\dots,z_N'),\quad \widetilde{z}''=(z_1,z_2'',\dots,z_N''), \tag{3-4}$$

$$\begin{gathered}
\widetilde{\boldsymbol{A}}=(-b_1,a_2,\dots,a_N),\quad \widetilde{\boldsymbol{B}}=(a_1-b_1,b_2,\dots,b_N),\\
\widetilde{\nu}=(n_1-b_1,n_2,\dots,n_N)^T.
\end{gathered} \tag{3-5}$$

The new shapes satisfy $\widetilde{z}^{\widetilde{\boldsymbol{A}}}\widetilde{z}''^{\widetilde{\boldsymbol{B}}}=(-1)^{\widetilde{\nu}}$. We also naturally obtain a new flattening $(\widetilde{f},\widetilde{f}',\widetilde{f}'')$ by permuting

$$\begin{gathered}
\widetilde{f}=(f_1',f_2,\dots,f_N)^T,\quad \widetilde{f}'=(f_1'',f_2',\dots,f_N')^T,\\
\widetilde{f}''=(f_1,f_2'',\dots,f_N'')^T;
\end{gathered} \tag{3-6}$$

this is an integer solution to $\widetilde{\boldsymbol{A}}\widetilde{f}+\widetilde{\boldsymbol{B}}\widetilde{f}''=\widetilde{\nu}$ and $\widetilde{f}+\widetilde{f}'+\widetilde{f}''=1$.

The torsion $\tau_{\mathcal{T}}$ (1-8) consists of two parts, a determinant and a monomial correction. By making use of the relations $z_1 + z_1'^{-1} - 1 = 0$ and $z_1 z_1' z_1'' = -1$, we find the determinant with the permuted Neumann–Zagier datum to be

$$\begin{aligned}
\text{(3-7)} \quad \det(\widetilde{\boldsymbol{A}}\Delta_{\tilde{z}''} + \widetilde{\boldsymbol{B}}\Delta_{\tilde{z}}^{-1}) &= \det(-b_1 z_1 + (a_1 - b_1) z_1'^{-1}, a_2 z_2'' + b_2 z_2^{-1}, \dots, a_N z_N'' + b_N z_N^{-1}) \\
&= \det(a_1 z_1'^{-1} - b_1, a_2 z_2'' + b_2 z_2^{-1}, \dots, a_N z_N'' + b_N z_N^{-1}) \\
&= -z_1 \det(a_1 z_1'' + b_1 z_1^{-1}, a_2 z_2'' + b_2 z_2^{-1}, \dots, a_N z_N'' + b_N z_N^{-1}) \\
&= -z_1 \det(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1}).
\end{aligned}$$

By simply using $z_1 z_1' z_1'' = -1$ and $f_1 + f_1' + f_1'' = 1$, we also see that the monomial correction transforms as

$$\text{(3-8)} \qquad \begin{aligned}
\tilde{z}^{\tilde{f}''} \tilde{z}''^{-\tilde{f}} &= z^{f''} z''^{-f} \frac{z_1'^{\,f_1} z_1^{\,-f_1'}}{z_1^{\,f_1''} z_1''^{-f_1}} \\
&= z^{f''} z''^{-f} (-1)^{f_1} \frac{(z_1 z_1'')^{-f_1} z_1^{\,f_1 + f_1'' - 1}}{z_1^{\,f_1''} z_1''^{-f_1}} \\
&= z^{f''} z''^{-f} (-1)^{f_1} z_1^{-1}.
\end{aligned}$$

The extra factors $z_1^{\pm 1}$ in the two parts of the torsion precisely cancel each other, leading in the end to

$$\text{(3-9)} \quad \det(\widetilde{\boldsymbol{A}}\Delta_{\tilde{z}''} + \widetilde{\boldsymbol{B}}\Delta_{\tilde{z}}^{-1}) \tilde{z}^{\tilde{f}''} \tilde{z}''^{-\tilde{f}} = (-1)^{f_1 + 1} \det(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1}) z^{f''} z''^{-f}.$$

This is just as desired, showing that the torsion is invariant up to a sign.

3.3 Independence of a choice of edge

We fix $M, \mathcal{T}, \hat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$. In order to choose matrices $\boldsymbol{A}, \boldsymbol{B}$, we must ignore the redundant gluing equation corresponding to an edge of $\mathcal{T}$. This was discussed in Section 2.3. Suppose, then, that we choose a different edge to ignore. For example, if we choose the $(N-1)^{\text{st}}$ rather than the N^{th} (and keep the same quad type and meridian path), then we obtain new Neumann–Zagier matrices $\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}$, which are related to the original ones as

$$\text{(3-10)} \qquad \widetilde{\boldsymbol{A}} = P_{(N-1,N)} \boldsymbol{A}, \quad \widetilde{\boldsymbol{B}} = P_{(N-1,N)} \boldsymbol{B},$$

where

(3-11)
$$P_{(N-1,N)} := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 \\ & & \ddots & & & \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ -1 & -1 & \cdots & -1 & -1 & -1 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

Similarly, eliminating the I^{th} rather than the N^{th} edge constraint is implemented by multiplying with a matrix $P_{(I,N)}$ whose I^{th} row is filled with -1s. Any such matrix satisfies $\det P_{(I,N)} = -1$.

In the formula for $\tau_{\mathcal{T}}$, only the determinant part is affected by a change of edge. Then

(3-12)
$$\begin{aligned} \det(\widetilde{\boldsymbol{A}}\Delta_{z''} + \widetilde{\boldsymbol{B}}\Delta_z^{-1}) &= \det(P_{(I,N)}(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1})) \\ &= -\det(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1}), \end{aligned}$$

leading to invariance of $\tau_{\mathcal{T}}$, up to the usual sign.

3.4 Independence of a choice of meridian path

Recall that an ideal triangulation on M induces a triangulation of its boundary torus ∂M. Consider two simple closed meridian loops in ∂M in general (normal) position with respect to the triangulation of ∂M. Recall that these paths are drawn on the triangulated 2–dimensional torus ∂M where faces of tetrahedra correspond to edges in the 2–dimensional triangulation, and edges of tetrahedra to vertices. In particular, for a one-cusped manifold M, every edge of the triangulation intersects a pair of vertices on the boundary ∂M.

We can deform one of our meridian paths into the other by using repeated applications of the fundamental move shown in Figure 6, locally pushing a section of the path across a vertex of ∂M. Thus, it suffices to assume that the two paths only differ by one such move. Suppose that we cross the I^{th} edge (by Section 3.3 we may assume that $I \neq N$), which has a combinatorial gluing constraint

(3-13)
$$X_I := \sum_{i=1}^{N} (\boldsymbol{G}_{Ii} Z_i + \boldsymbol{G}'_{Ii} Z'_i + \boldsymbol{G}''_{Ii} Z''_i) = 2\pi i,$$

and that the two tetrahedra where the paths enter and exit the vicinity of the edge have parameters (Z, Z', Z'') and (W, W', W''), as in the figure. We do not exclude the possibility that (Z, Z', Z'') and (W, W', W'') both coincide with the same triple

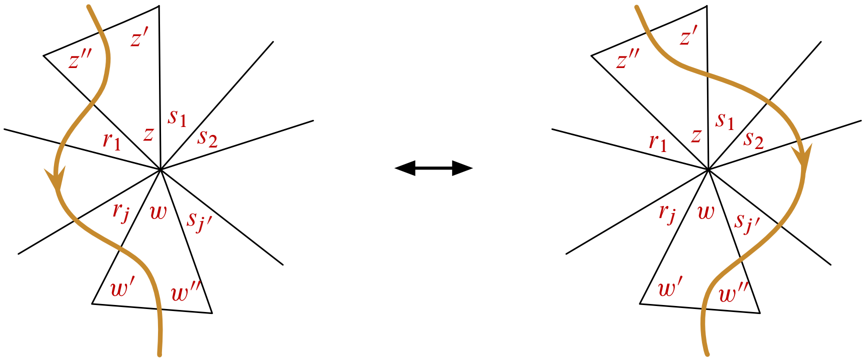


Figure 6: The fundamental move for changing a meridian path. Here, we deform through an edge E_I with gluing constraint $X_I = Z + W + R_1 + \cdots + R_j + S_1 + \cdots + S_{j'} = 2\pi i$.

(Z_i, Z_i', Z_i''), in some cyclic permutation. Then the difference in the logarithmic meridian equations (2-5) for the two paths will be

$$\delta_\mu = \pm(X_I - (Z + Z' + Z'') - (W + W' + W'')). \tag{3-14}$$

Note that two logarithmic meridian constraints that differ by (3-14) are compatible and equivalent, since upon using the additional equations $X_I = 2\pi i$ and $Z + Z' + Z'' = W + W' + W'' = i\pi$, we find that $\delta_\mu = 0$. A discretized version of this observation demonstrates that the same flattening satisfies both discretized meridian constraints.[2]

If we compute matrices $\boldsymbol{A}, \boldsymbol{B}$ using one meridian path and $\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}$ using the other — keeping quad type, flattening, and edge the same — the change (3-14) implies

$$\widetilde{\boldsymbol{A}} = P_I^{(\mu)\pm 1}\boldsymbol{A}, \quad \widetilde{\boldsymbol{B}} = P_I^{(\mu)\pm 1}\boldsymbol{B}, \tag{3-15}$$

where $P_I^{(\mu)}$ is the $\mathrm{SL}(N,\mathbb{Z})$ matrix

$$P_I^{(\mu)} = I + E_{NI}, \tag{3-16}$$

ie the identity plus an extra entry "1" in the N^{th} (meridian) row and I^{th} column. Since $\det P_I^{(\mu)} = 1$, this immediately shows that $\det(\widetilde{\boldsymbol{A}}\Delta_{z''} + \widetilde{\boldsymbol{B}}\Delta_z^{-1}) = \det(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1})$, and so the a change in the meridian path cannot affect $\tau_{\mathcal{T}}$.

[2]Note that this would not be the case if we allowed self-intersections of the meridian loops.

3.5 Independence of a choice of flattening

Now suppose that we choose two flattenings (f, f', f'') and $(\tilde{f}, \tilde{f}', \tilde{f}'')$, both satisfying

$$\boldsymbol{A} f + \boldsymbol{B} f'' = \nu, \quad f + f' + f'' = 1, \tag{3-17}$$

$$\boldsymbol{A} \tilde{f} + \boldsymbol{B} \tilde{f}'' = \nu, \quad \tilde{f} + \tilde{f}' + \tilde{f}'' = 1. \tag{3-18}$$

We may assume that we have a quad type with $\boldsymbol{B}$ nondegenerate. Indeed, by the result of Section 3.2, flattening invariance in one quad type implies flattening invariance in any quad type. Moreover, by Lemma A.3 of Appendix A, a quad type with nondegenerate $\boldsymbol{B}$ always exists. We also note that when $\boldsymbol{B}$ is invertible the matrix $\boldsymbol{B}^{-1}\boldsymbol{A}$ is symmetric (Lemma A.2).

The determinant in $\tau_{\mathcal{T}}$ is insensitive to the change of flattening. The monomial, on the other hand, can be manipulated as follows. Let us choose logarithms (Z, Z', Z'') of the shape parameters such that $\boldsymbol{A} Z + \boldsymbol{B} Z'' = i\pi\nu$. Then, assuming that $\boldsymbol{B}$ is nondegenerate, we compute

$$\begin{aligned}
\frac{z^{\tilde{f}''} z''^{-\tilde{f}}}{z^{f''} z''^{-f}} &= \exp[Z \cdot (f'' - \tilde{f}'') - Z'' \cdot (f - \tilde{f})] \\
&= \exp[-Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A}(f - \tilde{f}) - (i\pi \boldsymbol{B}^{-1}\nu - \boldsymbol{B}^{-1}\boldsymbol{A} Z) \cdot (f - \tilde{f})] \\
&= \exp[-i\pi \boldsymbol{B}^{-1}\nu \cdot (f - \tilde{f})] \\
&= \exp[-i\pi f'' \cdot (f - \tilde{f}) - i\pi \boldsymbol{B}^{-1}\boldsymbol{A} f \cdot (f - \tilde{f})] \\
&= \exp[-i\pi f'' \cdot (f - \tilde{f}) + i\pi f \cdot (f'' - \tilde{f}'')] \\
&= \exp[i\pi (f'' \cdot \tilde{f} - f \cdot \tilde{f}'')] = \pm 1.
\end{aligned}$$

Therefore, the monomial can change at most by a sign, and $\tau_{\mathcal{T}}$ is invariant as desired. This completes the proof of Theorem 1.3. □

3.6 Invariance under 2–3 moves

We finally come to the proof of Theorem 1.4, ie the invariance of $\tau_{\mathcal{T}}$ under 2–3 moves. We set up the problem as in Figure 7. Namely, we suppose that M has two different regular triangulations $\mathcal{T}$ and $\widetilde{\mathcal{T}}$, with N and $N+1$ tetrahedra, respectively, which are related by a local 2–3 move. Let us denote the respective (triples of) shape parameters as

$$Z := (X_1, X_2, Z_3, \dots, Z_N), \quad \widetilde{Z} := (W_1, W_2, W_3, Z_3, \dots, Z_N). \tag{3-19}$$

We fix a quad type, labeling the five tetrahedra involved in the 2–3 move as in the figure. We will also assume that when calculating Neumann–Zagier matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we choose to ignore an edge that is *not* the central one of the 2–3 *bipyramid*.

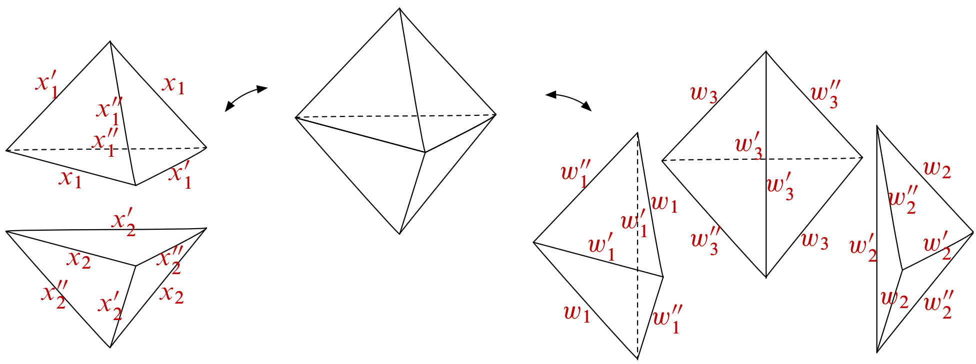


Figure 7: The geometry of the 2–3 move: a bipyramid split into two tetrahedra for triangulation $\bigcup_{i=1}^{N} \Delta_i$, and three for triangulation $\bigcup_{i=1}^{N+1} \widetilde{\Delta}_i$.

There are nine linear relations among the shapes of the tetrahedra involved in the move; three come from adding dihedral angles on the equatorial edges of the bipyramid

$$W_1' = X_1 + X_2, \quad W_2' = X_1' + X_2'', \quad W_3' = X_1'' + X_2', \tag{3-20}$$

and six from the longitudinal edges

$$\begin{aligned} X_1 &= W_2 + W_3'', & X_1' &= W_3 + W_1'', & X_1'' &= W_1 + W_2'', \\ X_2 &= W_2'' + W_3, & X_2' &= W_1'' + W_2, & X_2'' &= W_3'' + W_1. \end{aligned} \tag{3-21}$$

Moreover, due to the central edge of the bipyramid, there is an extra gluing constraint in $\widetilde{\mathcal{T}}$:

$$W_1' + W_2' + W_3' = 2\pi i. \tag{3-22}$$

After exponentiating the relations (3-20)–(3-22), and also using $z_i z_i' z_i'' = -1$ and $z_i'' + z_i^{-1} - 1 = 0$ for every tetrahedron Δ_i and $\widetilde{\Delta}_i$, we find a birational map between the shape parameters in the two triangulations. Explicitly,

$$\begin{gathered} \left\{ w_1' = x_1 x_2,\ w_2' = \frac{1 - x_2^{-1}}{1 - x_1},\ w_3' = \frac{1 - x_1^{-1}}{1 - x_2} \right\} \quad \text{or} \\ \left\{ x_1 = \frac{1 - w_2'^{-1}}{1 - w_3'},\ x_2 = \frac{1 - w_3'^{-1}}{1 - w_2'} \right\}. \end{gathered} \tag{3-23}$$

Note that the birational map is well-defined and one-to-one as long as no shape parameters $(x_1, x_2, w_1, w_2, w_3)$ equal 0, 1, or ∞. This condition is satisfied *so long as* triangulations $\mathcal{T}$ and $\widetilde{\mathcal{T}}$ are both regular. (A necessary condition is that no univalent edges are created on one side or the other of the 2–3 move; this is also sufficient when considering the discrete faithful representation of M.)

We must also choose a flattening in the two triangulations. Let us suppose for $\bigcup_{i=1}^{N+1} \widetilde{\Delta}_i$ we have a flattening with (triples of) integer parameters $\widetilde{f} = (d_1, d_2, d_3, f_3, \dots, f_N)$. This automatically determines a flattening $f = (e_1, e_2, f_3, \dots, f_N)$ for the $\bigcup_{i=1}^{N} \Delta_i$ triangulation, by simply setting

$$e_1 = d_2 + d_3'', \quad e_1' = d_3 + d_1'', \quad e_1'' = d_1 + d_2'', \\ e_2 = d_2'' + d_3, \quad e_2' = d_1'' + d_2, \quad e_2'' = d_3'' + d_1. \tag{3-24}$$

This is a discretized version of the six longitudinal relations (3-21). One can check that expected relations such as $e_1 + e_1' + e_1'' = 1$ are satisfied by virtue of the discretized edge constraint $d_1' + d_2' + d_3' = 2$ (cf (3-22)).

We have all the data needed to calculate $\tau_{\mathcal{T}}$. Let us start with determinants. In the triangulation $\bigcup_{i=1}^{N} \Delta_i$, we write the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ schematically in columns as

$$\boldsymbol{A} = (a_1, a_2, a_i), \quad \boldsymbol{B} = (b_1, b_2, b_i), \tag{3-25}$$

with a_i meaning $(a_3, a_4, \dots, a_N)$ and similarly for b_i. This leads to a determinant

$$\det\big(\boldsymbol{A}\,\Delta_{z''} + \boldsymbol{B}\,\Delta_z^{-1}\big) = \det(a_1 x_1'' + \tfrac{b_1}{x_1}, a_2 x_2'' + \tfrac{b_2}{x_2}, a_i z_i'' + \tfrac{b_i}{z_i}). \tag{3-26}$$

Alternatively, in the triangulation $\bigcup_{i=1}^{N+1} \widetilde{\Delta}_i$, the matrices $\widetilde{\boldsymbol{A}}$ and $\widetilde{\boldsymbol{B}}$ have one extra row and one extra column. The extra gluing condition (3-22) causes the extra row in both $\widetilde{\boldsymbol{A}}$ and $\widetilde{\boldsymbol{B}}$ to contain three -1s. Altogether, the matrices take the form

$$\widetilde{\boldsymbol{A}} = \begin{pmatrix} -1 & -1 & -1 & 0 \\ b_1 + b_2 & a_1 & a_2 & a_i \end{pmatrix}, \quad \widetilde{\boldsymbol{B}} = \begin{pmatrix} -1 & -1 & -1 & 0 \\ 0 & a_2 + b_1 & a_1 + b_2 & b_i \end{pmatrix}, \tag{3-27}$$

so that

$$\begin{aligned} \widetilde{\boldsymbol{A}}\,\Delta_{\widetilde{z}''} + \widetilde{\boldsymbol{B}}\,\Delta_{\widetilde{z}}^{-1} &= \begin{pmatrix} -w_1'' - \frac{1}{w_1} & -w_2'' - \frac{1}{w_2} & -w_3'' - \frac{1}{w_3} & 0 \\ (b_1 + b_2) w_1'' & a_1 w_2'' + \frac{a_2 + b_1}{w_2} & a_2 w_3'' + \frac{a_1 + b_2}{w_3} & a_i z_i'' + \frac{b_i}{z_i} \end{pmatrix} \\ &= \begin{pmatrix} -1 & -1 & -1 & 0 \\ (b_1 + b_2) w_1'' & a_1 w_2'' + \frac{a_2 + b_1}{w_2} & a_2 w_3'' + \frac{a_1 + b_2}{w_3} & a_i z_i'' + \frac{b_i}{z_i} \end{pmatrix}. \end{aligned}$$

It is then straightforward to check, using the map (3-23), that

$$(\widetilde{\boldsymbol{A}}\Delta_{\tilde{z}''}+\widetilde{\boldsymbol{B}}\,\Delta_{\tilde{z}}^{-1})\begin{pmatrix}1&-1&-1&0\\0&1&0&0\\0&0&1&0\\0&0&0&\mathbf{1}\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&\frac{w_3'x_2''}{x_1''}&0\\0&\frac{w_2'x_1''}{x_2''}&1&0\\0&0&0&\mathbf{1}\end{pmatrix}=\begin{pmatrix}-1&0\\ *&\boldsymbol{A}\,\Delta_{z''}+\boldsymbol{B}\,\Delta_z^{-1}\end{pmatrix}. \tag{3-28}$$

The determinant of the last matrix on the left hand side is $1-w_2'w_3'=1-w_1'^{-1}=w_1$. Therefore,

$$\det(\widetilde{\boldsymbol{A}}\Delta_{\tilde{z}''}+\widetilde{\boldsymbol{B}}\,\Delta_{\tilde{z}}^{-1})=-w_1^{-1}\det(\boldsymbol{A}\,\Delta_{z''}+\boldsymbol{B}\,\Delta_z^{-1}). \tag{3-29}$$

We should also consider the monomial correction. However, with flattenings related as in (3-24), and with shapes related by the exponentiated version of (3-21), it is easy to check that

$$\tilde{z}^{\tilde{f}''}\tilde{z}''^{-\tilde{f}}=w_1(-1)^{d_2''-e_1''}z^{f''}z''^{-f}. \tag{3-30}$$

We have thus arrived at the desired result; by combining (3-29) and (3-30), we find

$$\det(\widetilde{\boldsymbol{A}}\Delta_{\tilde{z}''}+\widetilde{\boldsymbol{B}}\,\Delta_{\tilde{z}}^{-1})\tilde{z}^{\tilde{f}''}\tilde{z}''^{-\tilde{f}}=\pm\det(\boldsymbol{A}\,\Delta_{z''}+\boldsymbol{B}\,\Delta_z^{-1})z^{f''}z''^{-f}, \tag{3-31}$$

so we have that $\tau_{\mathcal{T}}$ is invariant under the 2–3 move. This completes the proof of Theorem 1.4. □

4 Torsion on the character variety

Having given a putative formula for the nonabelian torsion of a cusped hyperbolic manifold M at the discrete faithful representation ρ_0, it is natural to ask whether the formula generalizes to other settings. In this section, we extend the torsion formula to general representations $\rho\colon \pi_1(M)\longrightarrow \mathrm{(P)SL}(2,\mathbb{C})$ for manifolds M with torus boundary, essentially by letting the shapes z be functions of ρ. We also find that some special results hold when M is hyperbolic and the representations lie on the geometric component X_M^{geom} of the $\mathrm{SL}(2,\mathbb{C})$ character variety.

We will begin with a short review of what it means for a combinatorial ideal triangulation to be *regular* with respect to a general representation ρ. We will also finally prove Proposition 1.7. Recall that Proposition 1.7 identified a canonical connected subset

$\mathcal{X}_M^{\mathrm{EP}}$ of the set of regular triangulations $\mathcal{X}_{\rho_0}$ of a hyperbolic 3–manifold M. This result allowed us to construct the topological invariant τ_M.

We then proceed to define an enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$ suitable for a general representation ρ, and propose a generalization of the torsion formula:

$$\tau_{\mathcal{T}}(\rho) := \pm\frac{1}{2}\det(\boldsymbol{A}\,\Delta_{z''} + \boldsymbol{B}\,\Delta_z^{-1}) z^{f''} z''^{-f}. \tag{4-1}$$

This formula looks identical to (1-8). However, the shape parameters here are promoted to functions $z \to z(\rho)$ of the representation ρ, which satisfy a well known deformed version of the gluing equations. Moreover, the flattening f is slightly more restricted than it was previously. We will prove the following in Section 4.5.

Theorem 4.1 *The formula for $\tau_{\mathcal{T}}(\rho)$ is independent of the choice of enhanced Neumann–Zagier datum, and is invariant under 2–3 moves connecting ρ–regular triangulations.*

When M is hyperbolic, it turns out that ρ_0–regular triangulations are ρ–regular for all but finitely many representations $\rho \in X_M^{\mathrm{geom}}$. Then we can create a topological invariant τ_M that is a function on X_M^{geom} just as in Proposition 1.7, by evaluating $\tau_{\mathcal{T}}(\rho)$ on any triangulation in the canonical subset $\mathcal{X}_M^{\mathrm{EP}} \subseteq \mathcal{X}_{\rho_0}$.

In general, there is a rational map from the character variety X_M to the zero-locus Y_M of the A–polynomial $A_M(\ell, m) = 0$ (see Cooper, Culler, Gillet, Long and Shalen [12]), for any M with torus boundary. Therefore, the shapes z and the torsion $\tau_{\mathcal{T}}$ are algebraic functions on components of the A–polynomial curve Y_M. When M is hyperbolic and $\rho \in X_M^{\mathrm{geom}}$, somewhat more is true: the shapes are *rational* function on the geometric component Y_M^{geom} (Proposition B.1). Then

$$\tau_M \in C(Y_M^{\mathrm{geom}}) = \mathbb{Q}(m)[\ell]/(A_M^{\mathrm{geom}}(\ell, m)), \tag{4-2}$$

where $A_M^{\mathrm{geom}}(\ell, m)$ is the geometric factor of the A–polynomial. We will give a simple example of the function τ_M for the figure-eight knot in Section 4.6.

4.1 A review of ρ–regular ideal triangulations

In this section we discuss the ρ–regular ideal triangulations that are needed to generalize our torsion invariant. Let M denote a 3–manifold with nonempty boundary and let $\rho\colon \pi_1(M) \longrightarrow \mathrm{PSL}(2, \mathbb{C})$ be a $\mathrm{PSL}(2, \mathbb{C})$ representation of its fundamental group. Let $\mathcal{X}$ denote the set of combinatorial ideal triangulations $\mathcal{T}$ of M. Matveev and Piergallini independently showed that every two elements of $\mathcal{X}$ with at least two ideal tetrahedra are connected by a sequence of 2–3 moves (and their inverses) [52; 59]. For a detailed exposition, see [53] and Benedetti and Petronio [5].

Given an ideal triangulation $\mathcal{T}$, let $V_{\mathcal{T}}$ denote the affine variety of nondegenerate solutions (ie, solutions in $\mathbb{C}\setminus\{0,1\}$) of the gluing equations of $\mathcal{T}$ corresponding to its edges. There is a developing map

$$V_{\mathcal{T}} \longrightarrow X_M, \tag{4-3}$$

where $X_M := \mathrm{Hom}(\pi_1(M), \mathrm{PSL}(2,\mathbb{C}))/\mathrm{PSL}(2,\mathbb{C})$ denotes the affine variety of all $\mathrm{PSL}(2,\mathbb{C})$ representations of $\pi_1(M)$.

Definition 4.2 Fix a $\mathrm{PSL}(2,\mathbb{C})$–representation of M. We say that $\mathcal{T}\in\mathcal{X}$ is ρ–*regular* if ρ is in the image of the developing map (4-3).

Let $\mathcal{X}_\rho \subset \mathcal{X}$ denote the set of all ρ–regular ideal triangulations of M. When M is hyperbolic, let ρ_0 denote its discrete faithful representation ρ_0 and let $X_M^{\mathrm{geom}} \subset X_M$ denote the geometric component of its character variety [66; 58]. We then have the following result.

Lemma 4.3 (a) *$\mathcal{T}\in\mathcal{X}_{\rho_0}$ if and only if $\mathcal{T}$ has no homotopically peripheral (ie, univalent) edges.*

(b) *If $\mathcal{T}\in\mathcal{X}_{\rho_0}$, then $\mathcal{T}\in\mathcal{X}_\rho$ for all but finitely many $\rho\in X_M^{\mathrm{geom}}$.*

Proof Part (a) has been observed several times; see [11], [7, Section 10.3], [69, Theorem 2.3] and also [22, Remark 3.4]. For part (b), fix $\mathcal{T}\in\mathcal{X}_{\rho_0}$. Observe that $\mathcal{T}$ is ρ–regular if the image of every edge[3] of $\mathcal{T}$ under ρ does not commute with the image under ρ of the peripheral subgroup of M. This is an algebraic condition on ρ, and moreover, when $\rho\in X_M^{\mathrm{geom}}$ is analytically nearby ρ_0, the condition is satisfied. It follows that the set of points of X_M^{geom} that satisfy the above condition is Zariski open. On the other hand, X_M^{geom} is an affine curve [66; 58]. It follows that $\mathcal{T}$ is ρ–regular for all but finitely many $\rho\in X_M^{\mathrm{geom}}$. □

4.2 The Epstein–Penner cell decomposition and its triangulations

Now we consider the canonical ideal cell decomposition of a hyperbolic manifold M with cusps [25], and finally prove Proposition 1.7. It is easy to see that every convex ideal polyhedron can be triangulated into ideal tetrahedra with nondegenerate shapes; see for instance Hodgson, Rubinstein and Segerman [44]. One wishes to know that

[3]Note that every edge can be completed to a closed loop by adding a path on the boundary T^2. The choice of completion does not matter for studying commutation with the peripheral subgroup.

every two such triangulations are related by a sequence of 2–3 moves. This is a combinatorial problem of convex geometry which we summarize below. For a detailed discussion, the reader may consult the book by De Loera, Rambau and Santos [14] and references therein.

Fix a convex polytope P in $\mathbb{R}^d$. One can consider the set of triangulations of P. When $d=2$, P is a polygon and it is known that every two triangulations are related by a sequence of flips. For general d, flips are replaced by *geometric bistellar moves*. When $d \geq 5$, it is known that the graph of triangulations (with edges given by geometric bistellar flips) is not connected, and has isolated vertices. For $d=3$, it is not known whether the graph is connected.

The situation is much better when one considers *regular triangulations* of P. In that case, the corresponding graph of regular triangulations is connected, an in fact it is the edge set of the *secondary polytope* of P. When $d=3$ and P is convex and in general position, then the only geometric bistellar move is the 2–3 move where the added edge that appears in the move is an edge that connects two vertices of P. When $d=3$ and P is not in general position, the same conclusion holds as long as one allows for tetrahedra that are flat, ie, lie on a 2–dimensional plane.

Returning to the Epstein–Penner ideal cell decomposition, let $\mathcal{X}_M^{\mathrm{EP}}$ denote the set of regular (in the sense of polytopes and in the sense of ρ_0) ideal triangulations of the ideal cell decomposition. The above discussion together with the fact that no edge of the ideal cell decomposition is univalent, implies that $\mathcal{X}^{\mathrm{EP}}$ is a connected subset of $\mathcal{X}_{\rho_0}$. This concludes the proof of Proposition 1.7.

A detailed discussion on the canonical set $\mathcal{X}_M^{\mathrm{EP}}$ of ideal triangulations of a cusped hyperbolic 3–manifold M is given in [35, Section 6].

4.3 Neumann–Zagier datum and the geometric component

Let M be a manifold with torus boundary and $\mathcal{T}$ a (combinatorial) ideal triangulation. The Neumann–Zagier datum $\beta_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B})$ may be generalized for representations $\rho \in X_M$ besides the discrete faithful.

To begin, choose a representation $\rho\colon \pi_1(M) \longrightarrow \mathrm{PSL}(2,\mathbb{C})$, and, if desired, a lift to $\mathrm{SL}(2,\mathbb{C})$. Let (μ,λ) be meridian and longitude cycles[4] on ∂M, and let $(m^{\pm 1}, \ell^{\pm 1})$ be the eigenvalues of $\rho(\mu)$ and $\rho(\lambda)$, respectively. For example, for the lift of the

[4] Recall again that these cycles are only canonically defined for knot complements. In general there is some freedom in choosing them, but the torsion depends in a predictable way on the choice, cf Yamaguchi [74].

discrete faithful representation to $\mathrm{SL}(2,\mathbb{C})$, we have $(m,\ell)=(1,-1)$; see Calegari [9]. These eigenvalues define a map

$$X_M \longrightarrow (\mathbb{C}^*)^2/\mathbb{Z}_2, \tag{4-4}$$

whose image is a curve Y_M, the zero-locus of the A–polynomial $A_M(\ell,m)=0$ [12]. We will denote the representation ρ as ρ_m to emphasize its meridian eigenvalue.

Now, given a triangulation $\mathcal{T}$, and with $\boldsymbol{A}$, $\boldsymbol{B}$, and ν defined as in Section 2.3, the gluing equations (2-12) can easily be deformed to account for $m\neq 1$. Namely, we find [58]

$$\prod_{j=1}^{N} z_j{}^{\boldsymbol{A}_{ij}} z_j''{}^{\boldsymbol{B}_{ij}} = (-1)^{\nu_i} m^{2\delta_{iN}}. \tag{4-5}$$

The developing map (4-3) maps every solution of these equations to a representation $\pi_1(M)\longrightarrow \mathrm{PSL}(2,\mathbb{C})$ with meridian eigenvalue $\pm m$. The triangulation $\mathcal{T}$ is ρ_m–regular if and only if ρ_m is in the image of this map. We can similarly express the longitude eigenvalue as a product of shape parameters

$$\prod_{j=1}^{N} z_j{}^{2C_j} z_j''{}^{2D_j} = (-1)^{2\nu_\lambda}\ell^2, \tag{4-6}$$

for some $2C_j, 2D_j, 2\nu_\lambda\in\mathbb{Z}$. Then, if $\mathcal{T}$ is a ρ_m–regular triangulation, the irreducible component of Y_M containing ρ_m is explicitly obtained by eliminating all shapes z_j from (4-5)–(4-6).

In general, the shapes z_j are algebraic functions on components of the variety Y_M. However, if M is hyperbolic and $\mathcal{T}$ is regular for all but finitely many representations on the geometric component Y_M^{geom}, then the shapes z_j become rational functions, $z_j\in C(Y_M^{\mathrm{geom}})$. We provide a proof of this fact in Appendix B. The field of functions $C(Y_M^{\mathrm{geom}})$ may be identified with $\mathbb{Q}(m)[\ell]/(A^{\mathrm{geom}}(\ell,m))$, and the functions $z_j(\ell,m)$ can easily be obtained from equations (4-5)–(4-6).

4.4 Flattening compatible with a longitude

In this section we define a restricted combinatorial flattening that is compatible with a longitude.

Recall what is a combinatorial flattening of an ideal triangulation $\mathcal{T}$ from Definition 2.1. Given a simple peripheral curve λ on the boundary of M that represents a longitude — in particular, having intersection number one with the chosen meridian μ — we can construct the sum of combinatorial edge parameters along λ, just as in Section 2.2. It takes the form

$$\lambda: \quad \sum_{i=1}^{N}(\boldsymbol{G}_{N+2,i}Z_i+\boldsymbol{G}'_{N+2,i}Z'_i+\boldsymbol{G}''_{N+2,i}Z''_i), \tag{4-7}$$

for integer vectors $\boldsymbol{G}_{N+1}$, $\boldsymbol{G}'_{N+2}$, $\boldsymbol{G}''_{N+2}$. Just as we obtained $\boldsymbol{A}$, $\boldsymbol{B}$, and ν from the edge and meridian equations (with or without deformation), we may also now define

$$\begin{gathered} C_i=\frac{1}{2}(\boldsymbol{G}_{N+2,i}-\boldsymbol{G}'_{N+2,i}), \quad D_i=\frac{1}{2}(\boldsymbol{G}''_{N+2,i}-\boldsymbol{G}'_{N+2,i}), \\ \nu_\lambda=-\frac{1}{2}\sum_{i=1}^{N}\boldsymbol{G}'_{N+2,i}. \end{gathered} \tag{4-8}$$

Definition 4.4 A combinatorial flattening (f,f',f'') is compatible with a longitude if in addition to Equations (2-10a)–(2-10b), it also satisfies

$$\boldsymbol{G}_{N+2,i}f+\boldsymbol{G}'_{N+2,i}f'+\boldsymbol{G}''_{N+2,i}f''=0. \tag{4-9}$$

Equivalently, a combinatorial flattening compatible with the longitude is a vector $(f,f'')\in\mathbb{Z}^{2N}$ that satisfies

$$\boldsymbol{A}f+\boldsymbol{B}f''=\nu, \quad C\cdot f+D\cdot f''=\nu_\lambda. \tag{4-10}$$

A combinatorial flattening compatible with a longitude always exists [56, Lemma 6.1].

In the context of functions on the character variety, it is natural to deform the meridian gluing equation, and simultaneously to introduce a longitude gluing equation, in the form

$$\mu: \quad \sum_{i=1}^{N}(\boldsymbol{G}_{N+1,i}Z_i+\boldsymbol{G}'_{N+1,i}Z'_i+\boldsymbol{G}''_{N+1,i}Z''_i)=2u, \tag{4-11a}$$

$$\lambda: \quad \sum_{i=1}^{N}(\boldsymbol{G}_{N+2,i}Z_i+\boldsymbol{G}'_{N+2,i}Z'_i+\boldsymbol{G}''_{N+2,i}Z''_i)=2v, \tag{4-11b}$$

for some complex parameters u and v. Upon exponentiation, these equations reduce to the expected (4-5)–(4-6) if

$$m^2=e^{2u}, \quad \ell^2=e^{2v}. \tag{4-12}$$

(It is easy to show, following [58], that C and D as defined by (4-10) are the correct exponents for the exponentiated longitude Equation (4-6).) If M is a knot complement and we want to lift from $\mathrm{PSL}(2,\mathbb{C})$ to $\mathrm{SL}(2,\mathbb{C})$ representations, we should take $m = e^u$ and $\ell = -e^v$ and divide (4-11a) by two before exponentiating. This provides the correct way to take a square root of the exponentiated gluing equation, cf [9].

The remarkable symplectic property of $\boldsymbol{A}$ and $\boldsymbol{B}$ may be extended to C and D, even though in general C, D are vectors of half-integers rather than integers. Namely, there exists a completion of $(\boldsymbol{A}\,\boldsymbol{B})$ to a full symplectic matrix $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{smallmatrix}\right)$ such that the *bottom rows* of $\boldsymbol{C}$ and $\boldsymbol{D}$ are the vectors C and D [58]. In particular, this means that

$$\boldsymbol{A}_N \cdot D - \boldsymbol{B}_N \cdot C = 1, \tag{4-13}$$

where $\boldsymbol{A}_N, \boldsymbol{B}_N$ are the bottom (meridian) rows of $\boldsymbol{A}, \boldsymbol{B}$.

4.5 Invariance of the generalized torsion

We finally have all the required ingredients for the generalized torsion formula. Let M be a three-manifold with torus boundary, and $\rho_m \colon \pi_1(M) \to \mathrm{(P)SL}(2,\mathbb{C})$ a representation with meridian eigenvalue m. Let $\mathcal{T}$ be a ρ_m–regular triangulation of M, which exists by Lemma 4.3 at least for a dense set of representations on the geometric component of the character variety. Choose an enhanced Neumann–Zagier datum $(z, \boldsymbol{A}, \boldsymbol{B}, f)$, with $z = z(\rho_m)$ satisfying the deformed gluing equations (4-5) and f satisfying (4-10). Then, as in (4-1), we define

$$\tau_{\mathcal{T}}(\rho_m) := \pm\frac{1}{2}\det(\boldsymbol{A}\,\Delta_{z''} + \boldsymbol{B}\,\Delta_z^{-1}) z^{f''} z''^{-f}.$$

We can now prove Theorem 4.1.

Repeating verbatim the arguments of Section 3, it is easy to see that $\tau_{\mathcal{T}}$ is independent of a choice of quad type, a choice of an edge of $\mathcal{T}$ and a choice of a meridian loop. The crucial observation is that the equations $\boldsymbol{A}Z + \boldsymbol{B}Z'' = i\pi\nu$ (including the meridian equation) are never used in the respective proofs. Therefore, deforming the meridian equation by $u \neq 0$ does not affect anything. For the same reason, it is not hard to see that the formula is invariant under ρ_m–regular 2–3 moves, by repeating the argument of Section 3.6.

The only nontrivial verification required is that $\tau_{\mathcal{T}}$ is independent of the choice of flattening. This does use the gluing equations in a crucial way. We check it now for $m \neq 1$.

Choose logarithms (Z, Z', Z'') of the shape parameters and a logarithm u of m such that $Z + Z' + Z'' = i\pi$ and

$$\boldsymbol{A}Z + \boldsymbol{B}Z'' = 2\boldsymbol{u} + i\pi\nu, \tag{4-14}$$

where $\boldsymbol{u}$ denote the N–dimensional vector $(0,0,\dots,0,u)^T$. By independence of quad type and Lemma A.3, we may assume we are using a quad type with nondegenerate $\boldsymbol{B}$. Now, suppose that (f,f',f'') and $(\tilde{f},\tilde{f}',\tilde{f}'')$ are two different generalized flattenings. Then

$$
\begin{aligned}
&(Z\cdot f''-Z''\cdot f)-(Z\cdot\tilde{f}''-Z''\cdot\tilde{f})\\
&\qquad = Z\cdot(f''-\tilde{f}'')+\boldsymbol{B}^{-1}(\boldsymbol{A}Z-i\pi\nu-2\boldsymbol{u})\cdot(f-\tilde{f})\\
&\qquad = Z\cdot(f''-\tilde{f}'')+Z\cdot\boldsymbol{B}^{-1}\boldsymbol{A}(f-\tilde{f})-i\pi\boldsymbol{B}^{-1}\nu\cdot(f-\tilde{f})-2\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f})\\
&\qquad = -i\pi\boldsymbol{B}^{-1}\nu\cdot(f-\tilde{f})-2\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f})\\
&\qquad = i\pi(f''\cdot\tilde{f}-f\cdot\tilde{f}'')-2\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f}),
\end{aligned}
$$

by manipulations similar to those of Section 3.5. The new term $2\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f})$ is now dealt with by completing the Neumann–Zagier matrices $(\boldsymbol{A}\,\boldsymbol{B})$ to a full symplectic matrix $\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{smallmatrix}\right)\in\mathrm{Sp}(2N,\mathbb{Q})$, whose bottom row agrees with (C,D). The symplectic condition implies that $\boldsymbol{A}\boldsymbol{D}^T-\boldsymbol{B}\boldsymbol{C}^T=I$, or $\boldsymbol{B}^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{D}^T-\boldsymbol{C}^T$. Then

$$
\begin{aligned}
\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f}) &= \boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{D}^T\boldsymbol{u}\cdot(f-\tilde{f})-\boldsymbol{C}^T\boldsymbol{u}\cdot(f-\tilde{f})\\
&= \boldsymbol{u}\cdot\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{A}(f-\tilde{f})-\boldsymbol{u}\cdot\boldsymbol{C}(f-\tilde{f})\\
&= -\boldsymbol{u}\cdot(\boldsymbol{C}(f-\tilde{f})+\boldsymbol{D}(f''-\tilde{f}''))\\
&= -u(C\cdot(f-\tilde{f})+D\cdot(f''-\tilde{f}'')).
\end{aligned}
$$

In this last equation, only the bottom row of $\boldsymbol{C}$ and $\boldsymbol{D}$ appears, due to the contraction with $\boldsymbol{u}=(0,0,\dots,0,u)$. But this bottom row is precisely what enters the generalized flattening equations (4-10); since both flattenings satisfy these equations, we must have $\boldsymbol{B}^{-1}\boldsymbol{u}\cdot(f-\tilde{f})=0$. Therefore, upon exponentiating, we find

$$
z^{f''}z''^{-f}=(-1)^{f''\cdot\tilde{f}-f\cdot\tilde{f}''}z^{\tilde{f}''}z''^{-\tilde{f}}=\pm z^{\tilde{f}''}z''^{-\tilde{f}}, \tag{4-15}
$$

which demonstrates that $\tau_{\mathcal{T}}$ is independent of the choice of flattening. Theorem 4.1 follows. □

4.6 Example: 4_1 continued

We briefly demonstrate the generalized torsion formula, using representations on the geometric component of the character variety X_{4_1} for the figure-eight knot complement.

We may consider the same triangulation as in Section 2.6. The edge and meridian equations (2-14) are deformed to

$$
z^2w^2z''w''=1,\quad zwz''=-m^2, \tag{4-16}
$$

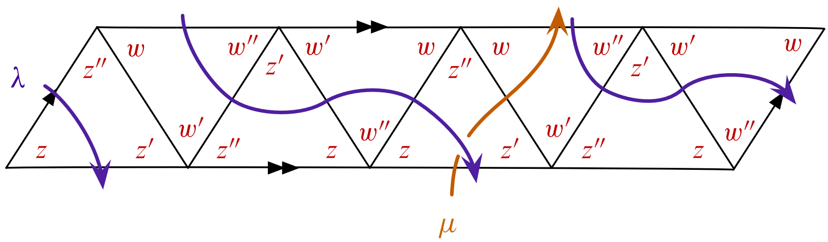


Figure 8: Longitude path for the figure-eight knot complement

with $z'' = 1 - z^{-1}$, $w'' = 1 - w^{-1}$ as usual. In addition, there is a longitude equation that may be read off from the longitude path in Figure 8. In logarithmic (combinatorial) form, we have $-2Z + 2Z' = 2v$, or

$$-4Z - 2Z'' = 2v - 2\pi i, \tag{4-17}$$

from which we identify

$$C = (-2, 0), \quad D = (-1, 0), \quad \nu_\lambda = -1. \tag{4-18}$$

Dividing (4-17) by two and exponentiating, we find

$$z^{-2} z''^{-1} = \ell, \tag{4-19}$$

with $\ell = -e^v$. This is the appropriate square root of (4-6) for lifting the geometric representations to $\mathrm{SL}(2,\mathbb{C})$. We can easily check it: by eliminating shape parameters from (4-16) and (4-19), we recover the geometric $\mathrm{SL}(2,\mathbb{C})$ A–polynomial for the figure-eight knot,

$$A^{\mathrm{geom}}_{\mathbf{4_1}}(\ell, m) = m^4 - (1 - m^2 - 2m^4 - m^6 + m^8)\ell + m^4\ell^2. \tag{4-20}$$

We may also use equations (4-16)–(4-19) to express the shape parameters as functions of ℓ and m. We find

$$z = -\frac{m^2 - m^{-2}}{1 + m^2\ell}, \quad w = \frac{m^2 + \ell}{m^2 - m^{-2}}. \tag{4-21}$$

These are functions on the curve $Y^{\mathrm{geom}}_{\mathbf{4_1}} = \{A^{\mathrm{geom}}_{\mathbf{4_1}}(\ell, m) = 0\}$.

The flattening (2-15) does *not* satisfy the new longitude constraint $C \cdot f + D \cdot f'' = \nu_\lambda$, so we must find one that does. The choice

$$(f_z, f'_z, f''_z; f_w, f'_w, f''_w) = (0, 0, 1; 0, 0, 1) \tag{4-22}$$

will work. Repeating the calculation of Section 3.1 with the same $\boldsymbol{A}$ and $\boldsymbol{B}$ but the new generalized flattening, we now obtain

$$\begin{aligned}\tau_{\mathbf{4_1}}(\rho_m) &= \pm\frac{1}{2}\det\begin{pmatrix} z''+1 & w''+1 \\ 1 & w''\end{pmatrix} zw \\ &= \pm\frac{1}{2}(z''w''-1)zw \\ &= \pm\frac{1-m^2-2m^4-m^6+m^8-2m^4\ell}{2m^4(m^2-m^{-2})}.\end{aligned} \tag{4-23}$$

This is in full agreement with the torsion found by [40; 17]. Note that for fixed m there are *two* choices of representation ρ_m on the geometric component of the character variety; they correspond to the two solutions of $A^{\mathrm{geom}}_{\mathbf{4_1}}(\ell,m)=0$ in ℓ.

Remark 4.5 It is interesting to observe that the numerator of Equation (4-23) is exactly $\partial A^{\mathrm{geom}}_{\mathbf{4_1}}/\partial\ell$. That the numerator of the geometric torsion typically carries a factor of $\partial A^{\mathrm{geom}}_{\mathbf{4_1}}/\partial\ell$ might be gleaned from the structure of "$\widehat{A}$–polynomials" in [17] and Gukov and Sułkowski [41], and will also be explored elsewhere.

5 The state integral and higher loops

Our explicit formulas for the torsion $\tau_{\mathcal{T}}$, as well as higher invariants $S_{\mathcal{T},n}$, have been obtained from a state integral model for analytically continued $\mathrm{SL}(2,\mathbb{C})$ Chern–Simons theory. In this section, we will review the state integral, and analyze its asymptotics in order to rederive the full asymptotic expansion

$$\mathcal{Z}_{\mathcal{T}}(\hbar)=\hbar^{-3/2}\exp\left[\frac{1}{\hbar}S_{\mathcal{T},0}+S_{\mathcal{T},1}+\hbar S_{\mathcal{T},2}+\hbar^2 S_{\mathcal{T},3}+\cdots\right], \tag{5-1}$$

and to unify the formulas of previous sections. We should point out that this section is *not* analytically rigorous, but serves as a motivation for our definition of the all-loop invariants, and provides a glimpse into the calculus of (complex, finite-dimensional) state integrals.

The basic idea of a state integral is to cut a manifold M into canonical pieces (ideal tetrahedra); to assign a simple partition function to each piece (a quantum dilogarithm); and then to multiply these simple partition functions together and integrate out over boundary conditions in order to obtain the partition function of the glued manifold M. A state integral provides a *finite-dimensional reduction* of the full Feynman path integral on M.

Currently, there are two flavors of $\mathrm{SL}(2,\mathbb{C})$ state integrals in the literature. The first, introduced by Hikami in [42; 43], studied in [17], and made mathematically rigorous by Andersen and Kashaev in [2], is based on a 3–dimensional lift of the 2–dimensional quantum Teichmüller theory in Kashaev's formalism [51]. It uses variables associated to faces of tetrahedra. The second, developed in [15], explicitly uses shape parameters — associated to edges of tetrahedra — and constitutes a 3d lift of Teichmüller theory in the Fock–Chekhov formalism [30]. The two types of state integrals should be equivalent, though this has only been demonstrated in isolated examples so far; see Spiridonov and Vartanov [65].

It is the second state integral that we employ in this paper, due to its explicit dependence on shape parameters. Indeed, suppose that M is an oriented one-cusped hyperbolic manifold with a ρ_0–regular triangulation $\mathcal{T}$ and enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$, with $\boldsymbol{A} f + \boldsymbol{B} f'' = \nu$. We must also assume that $\boldsymbol{B}$ is nondegenerate, which (Lemma A.3) is always possible. Then we will[5] show in Appendix C that the state integral of [15] takes the form

$$\mathcal{Z}_{\mathcal{T}}(\hbar) = \sqrt{\frac{8\pi^3}{\hbar^3 \det \boldsymbol{B}}} \int \frac{d^N Z}{(2\pi\hbar)^{N/2}} e^x \prod_{i=1}^{N} \psi_\hbar(Z_i), \tag{5-2}$$

where $x = \frac{1}{\hbar}[\frac{1}{2}(i\pi + \frac{\hbar}{2})^2 f \cdot \boldsymbol{B}^{-1}\nu - (i\pi + \frac{\hbar}{2}) Z \cdot \boldsymbol{B}^{-1}\nu + \frac{1}{2} Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A} Z]$ and $\psi_\hbar(Z)$ is a noncompact quantum dilogarithm [4; 27], the Chern–Simons partition function of a single tetrahedron. The integration variables Z_i are, literally, the logarithmic shape parameters of $\mathcal{T}$.

The integration contour of (5-2) is unspecified. A complete, nonperturbative definition of $\mathcal{Z}_{\mathcal{T}}(\hbar)$ requires a choice of contour, and the choice leading to invariance under 2–3 moves (etc.) may be quite subtle. However, a formal asymptotic expansion of the state integral as in (5-1) *does not require* a choice of contour. It simply requires a choice of critical point for the integrand. Then the asymptotic series may be developed via *formal Gaussian integration* in an infinitesimal neighborhood of the critical point.

We will show in Section 5.1 that all the leading order critical points of (5-2) are logarithmic solutions to the gluing equations

$$\text{critical points} \quad \longleftrightarrow \quad z^{\boldsymbol{A}}(1-z^{-1})^{\boldsymbol{B}} = (-1)^{\nu}, \tag{5-3}$$

with $z = \exp(Z)$. In particular, the critical points are isolated. Then, choosing the discrete faithful solution to (5-3), we formally expand the state integral to find that

[5] Here we multiply (C-23) (at $u = 0$) by an extra, canonical normalization factor $(2\pi/\hbar)^{3/2}$, in order to precisely match the asymptotics of the Kashaev invariant at the discrete faithful representation.

- $S_{\mathcal{T},0}$, the evaluation of leading order part of the integrand at the critical point, is the complex volume of M;
- $\exp(-2S_{\mathcal{T},1})$ is expressed as the determinant of a Hessian matrix
$$\mathcal{H} = -\boldsymbol{B}^{-1}\boldsymbol{A} + \Delta_{1-z}^{-1},$$
with a suitable monomial correction, and reproduces the torsion (1-8);
- the higher $S_{\mathcal{T},n}$ are obtained via a finite-dimensional Feynman calculus, and explicitly appear as rational functions of shape parameters.

It follows from the formalism of [15], reviewed in Appendix C, that the state integral (5-2) is only well-defined up to multiplicative prefactors of the form

$$\exp\left(\frac{\pi^2}{6\hbar}a + \frac{i\pi}{4}b + \frac{\hbar}{24}c\right), \quad a, b, c \in \mathbb{Z}. \tag{5-4}$$

This means that we only obtain $(S_{\mathcal{T},0}, \tau_{\mathcal{T}} = 4\pi^3 e^{-2S_{\mathcal{T},1}}, S_{\mathcal{T},2})$ modulo $(\frac{\pi^2}{6}\mathbb{Z}, i, \frac{1}{24}\mathbb{Z})$, respectively; however, all the higher invariants $S_{\mathcal{T},n\geq 3}$ should be unambiguous. Moreover, in Section 3 we saw that the ambiguity in $\tau_{\mathcal{T}}$ could be lifted[6] to a sign ± 1. Although the construction of the asymptotic series (5-1) appears to depend on $\mathcal{T}$, we certainly expect the following.

Conjecture 5.1 *The invariants $\{S_{\mathcal{T},n}\}_{n=0}^{\infty}$ are independent of the choice of regular triangulation and Neumann–Zagier datum (including the choice of quad type with* $\det \boldsymbol{B} \neq 0$*, etc.), up to the ambiguity* (5-4)*, and thus constitute topological invariants of M.*

We now proceed to analyze the critical points and asymptotics of (5-2) in greater detail. In Section 5.5, we will also generalize the state integral to arbitrary representations, with nonunit meridian eigenvalue $m = e^u \neq 1$, and give an example of $S_{\mathcal{T},2}(m)$, $S_{\mathcal{T},3}(m)$ as functions on the character variety Y_M^{geom} for the figure-eight knot.

5.1 Critical points

We begin by showing that the critical points of (5-2) are indeed solutions to the gluing equations. For this purpose, we need to know the quantum dilogarithm $\psi_\hbar(Z)$. The latter is given by [17, Equation 3.22]

$$\psi_\hbar(Z) = \prod_{r=1}^{\infty} \frac{1 - q^r e^{-Z}}{1 - ({}^L q)^{-r+1} e^{-{}^L Z}}, \tag{5-5}$$

[6] It may also be possible to lift the ambiguities in $S_{\mathcal{T},0}$ and $S_{\mathcal{T},2}$ by using ordered triangulations, as in [57; 75].

for $|q|<1$, where

$$q := \exp\hbar, \quad {}^{L}q := \exp\frac{-4\pi^2}{\hbar}, \quad {}^{L}Z := \frac{2\pi i}{\hbar}Z. \tag{5-6}$$

The quantum dilogarithm $\psi_\hbar(Z)$ coincides with the restriction to $|q|<1$ of Faddeev's quantum dilogarithm [27], as follows from [17, Equation 3.23]. $\psi_\hbar(Z)$ is the Chern–Simons wavefunction of a single tetrahedron [15]. The quantum dilogarithm has an asymptotic expansion as $\hbar\to 0$, given by (cf [17, Equation 3.26])

$$\begin{aligned}\psi_\hbar(Z) &\overset{\hbar\to 0}{\sim} \exp\sum_{n=0}^{\infty}\frac{B_n\hbar^{n-1}}{n!}\widetilde{\mathrm{Li}}_{2-n}(e^{-Z}) \\ &= \exp\Big[\frac{1}{\hbar}\widetilde{\mathrm{Li}}_2(e^{-Z})+\frac{1}{2}\widetilde{\mathrm{Li}}_1(e^{-Z})-\frac{\hbar}{12}z'+\frac{\hbar^3}{720}z(1+z)z'^3+\cdots\Big],\end{aligned} \tag{5-7}$$

where B_n is the n^{th} Bernoulli number, with $B_1=1/2$.

The coefficients of strictly positive powers of $\hbar$ (ie $n\geq 2$) in the expansion are rational functions of $z=e^Z$, but the two leading asymptotics — the logarithm and dilogarithm — are multivalued and have branch cuts. In contrast, the function $\psi_\hbar(Z)$ itself is a meromorphic function on $\mathbb{C}$ for any fixed $\hbar\neq 0$. Branch cuts in its asymptotics arise when families of poles collide in the $\hbar\to 0$ limit. In the case of purely imaginary $\hbar$ with $\operatorname{Im}\hbar>0$ (a natural choice in the analytic continuation of $\mathrm{SU}(2)$ Chern–Simons theory), a careful analysis of this pole-collision process leads to branch cuts for $\widetilde{\mathrm{Li}}_2$ and $\widetilde{\mathrm{Li}}_1$ that are different from the standard ones (Figure 9). We indicate the modified analytic structure of these two functions (really functions of Z rather than e^{-Z}) with an extra tilde.

Now, the critical points of the integrand, at leading order[7] in the $\hbar$ expansion, are solutions to

$$\begin{aligned}0&=\frac{\partial}{\partial Z_i}\Big(-\frac{\pi^2}{2}f\cdot \boldsymbol{B}^{-1}\nu-i\pi Z\cdot\boldsymbol{B}^{-1}\nu+\frac{1}{2}Z\cdot\boldsymbol{B}^{-1}\boldsymbol{A}Z+\sum_i\widetilde{\mathrm{Li}}_2(e^{-Z_i})\Big)\\ &=-i\pi(\boldsymbol{B}^{-1}\nu)_i+(\boldsymbol{B}^{-1}\boldsymbol{A}Z)_i-\widetilde{\mathrm{Li}}_1(e^{-Z_i}),\end{aligned}$$

in other words,

$$\boldsymbol{A}Z+\boldsymbol{B}(-\widetilde{\mathrm{Li}}_1(e^{-Z}))=i\pi\nu. \tag{5-8}$$

Since $\exp[-\widetilde{\mathrm{Li}}_1(e^{-Z_i})]=1-z_i^{-1}$, we see that every solution to (5-8) is a particular logarithmic lift of a solution to the actual gluing equations $z^{\boldsymbol{A}}(1-z^{-1})^{\boldsymbol{B}}=(-1)^\nu$.

[7] We treat all subleading terms as perturbations. The exact location of the critical point will acquire perturbative corrections, described in Section 5.4.

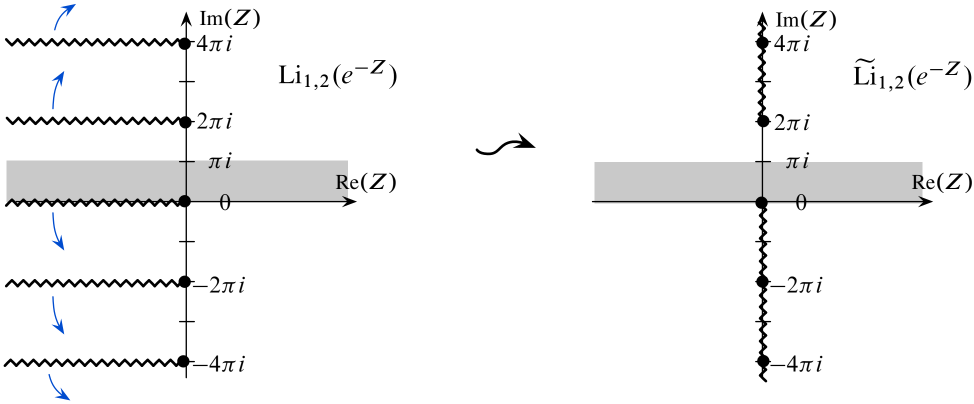


Figure 9: Rotating the standard branch cuts of $\mathrm{Li}_2(e^{-Z})$ and $\mathrm{Li}_1(e^{-Z})$ to produce $\widetilde{\mathrm{Li}}_2(e^{-Z})$ and $\widetilde{\mathrm{Li}}_1(e^{-Z})$, as functions of Z. The shaded region indicates where the standard logarithms of shape parameters for the discrete faithful representation lie.

It is a lift that precisely satisfies the logarithmic constraints (2-6) of Section 2, with $Z_i'' = -\widetilde{\mathrm{Li}}_1(e^{-Z_i})$.

When $0 \le \operatorname{Im} Z_i \le \pi$, the branches of the standard logarithms and dilogarithms agree with those of the modified ones. In particular, given the discrete faithful solution to $z^{\boldsymbol{A}}(1-z^{-1})^{\boldsymbol{B}} = (-1)^{n}$, taking standard logarithms immediately produces a solution to (5-8). Therefore, the discrete faithful representation always corresponds to a critical point of the state integral.

5.2 Volume

By substituting a solution to (5-8) back into the $\hbar^{-1}$ (leading order) part of the integrand, we obtain the following formula for the complex volume of a representation:

$$S_{\mathcal{T},0} = -\frac{\pi^2}{2} f \cdot \boldsymbol{B}^{-1}\nu - i\pi Z \cdot \boldsymbol{B}^{-1}\nu + \frac{1}{2} Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A} Z + \sum_i \widetilde{\mathrm{Li}}_2(e^{-Z_i}) \quad \left(\mathrm{mod}\ \frac{\pi^2}{6}\right). \tag{5-9}$$

Some manipulation involving the flattening can be used to recast this as

$$\boxed{S_{\mathcal{T},0} = -\frac{1}{2}(Z - i\pi f) \cdot (Z'' + i\pi f'') + \sum_i \widetilde{\mathrm{Li}}_2(e^{-Z_i})} \quad \left(\mathrm{mod}\ \frac{\pi^2}{6}\right), \tag{5-10}$$

where $Z_i'' := -\widetilde{\mathrm{Li}}_1(e^{-Z_i})$. It is straightforward to verify that this formula is independent of the choice of quad type, choice of edge of $\mathcal{T}$, choice of meridian loop, choice of flattening, and 2–3 moves defines a topological invariant, which agrees with the complex Chern–Simons invariant of M. Since the complex volume in this form has already been studied at length in the literature, we suppress the details here.

At the discrete faithful representation, we can remove the "tildes" from the logarithm and dilogarithm. If we consider the discrete faithful solution to $z^{\boldsymbol{A}}(1-z^{-1})^{\boldsymbol{B}} = (-1)^{\boldsymbol{n}}$, and take standard logarithms

$$Z_i = \log z_i, \quad Z_i'' = \log(1-z_i^{-1}) \quad (\text{with } 0 \le \operatorname{Im} Z, \operatorname{Im} Z'' \le \pi),$$

we find

$$\begin{aligned} S_{\mathcal{T},0} &= i(\mathrm{Vol}(M) - i\,\mathrm{CS}(M)) \\ &= -\frac{1}{2}(Z - i\pi f)\cdot(Z'' + i\pi f'') + \sum_i \mathrm{Li}_2(e^{-Z_i}) \quad \Big(\mathrm{mod}\ \frac{\pi^2}{6}\Big). \end{aligned} \tag{5-11}$$

This is a version of the simple formula for the complex volume given in [56]. It is known that the ambiguity in the volume can be lifted from $\pi^2/6$ to $2\pi^2$ using more refined methods; see [57; 38; 75] and Dupont and Zickert [24].

5.3 Torsion revisited

Next, we can derive our torsion formula (1-8). The torsion comes from the $\hbar^0$ part in the asymptotic expansion of the state integral, which has several contributions.

From formal Gaussian integration around a critical point (5-8), we get a determinant $(2\pi\hbar)^{N/2}(\det \mathcal{H})^{-1/2}$, where

$$\begin{aligned} \mathcal{H}_{ij} &= -\frac{\partial^2}{\partial Z_i \partial Z_j}\Big(-\frac{\pi^2}{2} f\cdot \boldsymbol{B}^{-1}\nu - i\pi Z\cdot \boldsymbol{B}^{-1}\nu \\ &\qquad\qquad + \frac{1}{2} Z\cdot \boldsymbol{B}^{-1}\boldsymbol{A} Z + \sum_i \widetilde{\mathrm{Li}}_2(e^{-Z_i})\Big) \\ &= \boxed{(-\boldsymbol{B}^{-1}\boldsymbol{A} + \Delta_{z'})_{ij}} \end{aligned} \tag{5-12}$$

is the Hessian matrix of the exponent (at leading order $\hbar^{-1}$). Here we define that $\Delta_{z'} := \mathrm{diag}(z_1', \dots, z_N')$, with $z_i' = (1-z_i)^{-1}$ as usual. Multiplying the determinant is the $\hbar^0$ piece of the integrand, evaluated at the critical point. From the $\hbar^0$ part of the quadratic exponential, we get

$$\begin{aligned}
\text{(5-13)}\quad \exp\Big(\frac{i\pi}{2} f\cdot \boldsymbol{B}^{-1}\nu - \frac{1}{2} Z\cdot \boldsymbol{B}^{-1}\nu\Big) &= \exp\Big(\frac{1}{2} f\cdot(\boldsymbol{B}^{-1}\boldsymbol{A} Z + Z'') - \frac{1}{2} Z\cdot(\boldsymbol{B}^{-1}\boldsymbol{A} f + f'')\Big)\\
&= \exp\Big(-\frac{1}{2} Z\cdot f'' + \frac{1}{2} Z''\cdot f\Big)\\
&= (z^{f''} z''^{-f})^{-1/2},
\end{aligned}$$

whereas from the quantum dilogarithm at order $\hbar^0$ we find

$$\text{(5-14)}\qquad \exp\Big(\frac{1}{2}\sum_i \widetilde{\mathrm{Li}}_1(e^{-Z_i})\Big) = \pm\prod_i \frac{1}{\sqrt{1-z_i^{-1}}} = \pm\det\Delta_{z''}^{-1/2}.$$

Combining the determinant $(2\pi\hbar)^{N/2}(\det\mathcal{H})^{-1/2}$, the corrections (5-13)–(5-14), and the overall prefactor $\sqrt{8\pi^3/\det\boldsymbol{B}}\,(2\pi\hbar)^{-N/2}$ in the integral (5-2) itself, we finally obtain

$$\begin{aligned}
\text{(5-15)}\qquad e^{S_1} &= \sqrt{\frac{8\pi^3}{\det\boldsymbol{B}\,\det(-\boldsymbol{B}^{-1}\boldsymbol{A}+\Delta_{z'})\det\Delta_{z''} z^{f''} z''^{-f}}}\\
&= \sqrt{\frac{-8\pi^3}{\det(\boldsymbol{A}\Delta_{z''}+\boldsymbol{B}\Delta_z^{-1}) z^{f''} z''^{-f}}},
\end{aligned}$$

up to multiplication by a power of i; or

$$\text{(5-16)}\qquad \tau_M := 4\pi^3 e^{-2S_1} = \pm\tfrac{1}{2}\det(\boldsymbol{A}\Delta_{z''}+\boldsymbol{B}\Delta_z^{-1}) z^{f''} z''^{-f},$$

just as in (1-8). Despite the fact that the original state integral only made sense for nondegenerate $\boldsymbol{B}$, the final formula for the torsion is well-defined for any $\boldsymbol{B}$.

5.4 Feynman diagrams and higher loops

The remainder of the invariants $S_{\mathcal{T},n}$ can be obtained by continuing the saddle-point (stationary phase) expansion of the state integral to higher order. The calculation can be systematically organized into a set of Feynman rules (cf [45, Chapter 9], [6] and [60]). The resulting formulas — summarized in Section 1 — are explicit algebraic functions of the exponentiated shape parameters z_i, and belong to the invariant trace field E_M.

To proceed, we should first recenter the integration around a critical point. Thus, we replace $Z\to Z+\zeta$ and integrate over ζ, assuming Z to be a solution to (5-8). Using

[17, Equation 3.26], we expand as follows:

$$Z_{\mathcal{T}}(\hbar) = \sqrt{\frac{8\pi^3}{\hbar^3 \det \boldsymbol{B}}} \int \frac{d^N\zeta}{(2\pi\hbar)^{N/2}} \prod_{i=1}^{N} \psi_\hbar(Z_i+\zeta_i) \times e^{\frac{1}{\hbar}[\frac{1}{2}(i\pi+\frac{\hbar}{2})^2 f\cdot\boldsymbol{B}^{-1}\nu - (i\pi+\frac{\hbar}{2})(Z+\zeta)\cdot\boldsymbol{B}^{-1}\nu + \frac{1}{2}(Z+\zeta)\cdot\boldsymbol{B}^{-1}\boldsymbol{A}(Z+\zeta)]}$$
$$\sim \sqrt{\frac{8\pi^3}{\hbar^3 \det \boldsymbol{B}}} e^{\Gamma^{(0)}(Z)} \times \int \frac{d^N\zeta}{(2\pi\hbar)^{N/2}} \exp\Big[-\frac{1}{2\hbar}\zeta\cdot\mathcal{H}(Z)\cdot\zeta + \sum_{k=1}^{\infty}\sum_{i=1}^{N}\frac{\Gamma_i^{(k)}(Z)}{k!}\zeta_i^k\Big]. \tag{5-17}$$

In this form, the first coefficient $\Gamma^{(0)}(Z)$ can be identified with an overall *vacuum energy*, while the rest of the $\Gamma_i^{(k)}(Z)$ are *vertex factors*.

Every $\Gamma^{(k)}(Z)$ here is a series in $\hbar$, in general starting with a $1/\hbar$ term. However, $\Gamma_i^{(1)}$ must vanish at leading order $\hbar^{-1}$ precisely because Z is a solution to the leading order critical point equations; and we have also already extracted the leading $\hbar^{-1}$ piece of $\Gamma_i^{(2)}$ as the Gaussian integration measure $-(1/2\hbar)\zeta\mathcal{H}\zeta$. Typically, 1–vertices and 2–vertices are absent from a Feynman calculus. Here, however, they appear because our critical point equation and the Hessian (respectively) are only accurate at leading order, and incur $\hbar$–corrections. (Note that the 1–vertices and 2–vertices are counted separately in (5-19) below.)

The vacuum energy $\Gamma^{(0)}$ contributes to every $S_{\mathcal{T},n}$, $n \ge 0$. Its leading order $\hbar^{-1}$ term is just the complex volume (5-10), while the $\hbar^0$ piece contains the corrections (5-13)–(5-14) to the torsion. At higher order in $\hbar$, we have

$$\Gamma^{(0)}(Z) = \frac{1}{\hbar}S_0 + \hbar^0(\cdots) + \frac{\hbar}{8} f\cdot\boldsymbol{B}^{-1}\boldsymbol{A}f + \sum_{n=2}^{\infty}\frac{\hbar^{n-1}B_n}{n!}\sum_{i=1}^{N}\mathrm{Li}_{2-n}(z_i^{-1}) \ \Big(\mathrm{mod}\ \frac{\hbar}{24}\Big). \tag{5-18}$$

Each S_n, $n \ge 2$, is calculated by taking the $\hbar^{n-1}$ part of $\Gamma^{(0)}$, and adding to it an appropriate sum of Feynman diagrams. The rules for the diagrams are derived from (5-17) as follows. There are vertices of all valencies $k = 1, 2, \ldots$, with a vertex factor given by $\Gamma_i^{(k)}$. One draws all connected diagrams (graphs) with

$$\boxed{\#\,\text{loops} + \#\,1\text{–vertices} + \#\,2\text{–vertices} \le n.} \tag{5-19}$$

Each k–vertex is assigned a factor $\Gamma_i^{(k)}$, and each edge is assigned a propagator

$$\text{propagator}: \quad \Pi_{ij} := \hbar \mathcal{H}_{ij}^{-1} = \hbar(-\boldsymbol{B}^{-1}\boldsymbol{A} + \Delta_{z'})_{ij}^{-1}. \tag{5-20}$$

The diagrams are then evaluated by contracting the vertex factors with propagators, and multiplying by a standard *symmetry factor*. In each diagram, one should restrict to the $\hbar^{n-1}$ term in its evaluation.

Explicitly, using the asymptotic expansion (5-7) of the quantum dilogarithm, we find that the vertices are

$$\text{1–vertex:} \quad \Gamma_i^{(1)} = -\frac{1}{2}(\boldsymbol{B}^{-1}\nu)_i - \sum_{n=1}^{\infty} \frac{\hbar^{n-1} B_n}{n!} \mathrm{Li}_{1-n}(z_i^{-1}) = -\frac{1}{2}(\boldsymbol{B}^{-1}\nu)_i + \frac{z_i'}{2} + \cdots, \tag{5-21a}$$

$$\text{2–vertex:} \quad \Gamma_i^{(2)} = \sum_{n=1}^{\infty} \frac{\hbar^{n-1} B_n}{n!} \mathrm{Li}_{-n}(z_i^{-1}) = \frac{z_i z_i'^2}{2} - \frac{\hbar}{12} z_i(1+z_i) z_i'^3 + \cdots, \tag{5-21b}$$

$$k\text{–vertex:} \quad \Gamma_i^{(k)} = (-1)^k \sum_{n=0}^{\infty} \frac{\hbar^{n-1} B_n}{n!} \mathrm{Li}_{2-n-k}(z_i^{-1}) \qquad (k \geq 3). \tag{5-21c}$$

Note that in $\Gamma_i^{(1)}$ we could also write $\boldsymbol{B}^{-1}\nu = \boldsymbol{B}^{-1}\boldsymbol{A} f + f''$. When the inequality (5-19) is saturated, only the leading order ($\hbar^{-1}$ or $\hbar^0$) terms of the vertex factors (5-21) need be considered. Otherwise, subleading $\hbar$–corrections may be necessary.

Examples of 2–loop and 3–loop Feynman diagrams were given in Figures 1–3 of Section 1, along with the entire evaluated expression for $S_{\mathcal{T},2}$.

5.5 *n*–loop invariants on the character variety

Just as we extended the torsion formula to general representations $\rho \in X_M$ in Section 4, we may now generalize the entire state integral. The basic result for the higher invariants $S_{\mathcal{T},n}$ is that their formulas remain completely unchanged. The shapes z_i simply become functions of the representation ρ, and satisfy the deformed gluing equations (4-5)–(4-6). One must also make sure to use a generalized flattening whenever it occurs, just as in Section 4.

We note that, for a hyperbolic knot complement $M = S^3 \backslash K$, the generalized Chern–Simons state integral $\mathcal{Z}_M(u;\hbar)$ is expected to match the asymptotic expansion of the colored Jones polynomials $J_N(K;q)$. Specifically, one should consider the limit

$$N \to \infty, \quad \hbar \to 0, \quad q^N = e^{N\hbar} = e^{2u} \text{ fixed,} \tag{5-22}$$

where $m = e^u$ is the meridian eigenvalue for a geometric representation ρ_m in the neighborhood of the discrete faithful. This is the full *Generalized Volume Conjecture* of [39].

To see how formulas for the generalized invariants $S_{\mathcal{T},n}$, $n \ge 0$, come about, consider the state integral at general meridian eigenvalue $m = e^u$. From (C-23) of Appendix C, we find

$$\mathcal{Z}_{\mathcal{T}}(u;\hbar) = \sqrt{\frac{8\pi}{\hbar^3 \det \boldsymbol{B}}} \int \frac{d^N Z}{(2\pi\hbar)^{N/2}} \prod_{i=1}^{N} \psi_\hbar(Z_i) e^{-\frac{1}{2\hbar} Z\cdot \boldsymbol{B}^{-1}\boldsymbol{A} Z} e^x, \tag{5-23}$$

where

$$x = \tfrac{1}{\hbar}[2\boldsymbol{u}\cdot\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{u} + (2\pi i+\hbar) f\cdot\boldsymbol{B}^{-1}\boldsymbol{u} + \tfrac{1}{2}(i\pi+\tfrac{\hbar}{2})^2 f\cdot\boldsymbol{B}^{-1}\nu - Z\cdot\boldsymbol{B}^{-1}(2\boldsymbol{u}+(i\pi+\tfrac{\hbar}{2})\nu)],$$

$\boldsymbol{u} := (0,\dots,0,u)$ and $\boldsymbol{D}$ is the block appearing in any completion of the Neumann–Zagier matrices $(\boldsymbol{A}\,\boldsymbol{B})$ to $\left(\begin{smallmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{smallmatrix}\right) \in \mathrm{Sp}(2N,\mathbb{Q})$, such that the bottom row D of $\boldsymbol{D}$ appears in the longitude gluing equation $C\cdot Z + D\cdot Z'' = v + 2\pi i\,\nu_\lambda$ (Section 4.4). Indeed, since we are contracting with $\boldsymbol{u}$, only this bottom row of $\boldsymbol{D}$ really matters in (5-23).

The critical points of the state integral are now given by

$$\boldsymbol{A}Z + \boldsymbol{B}Z'' = 2\boldsymbol{u} + i\pi\nu, \tag{5-24}$$

with $Z'' := -\widetilde{\mathrm{Li}}_1(e^{-Z})$. As expected, this is the logarithmic form of the deformed gluing equation (4-5). Thus, all critical points correspond to representations $\rho = \rho_m \in X_M$. The multivalued nature of this equation must be carefully studied to make sure desired solutions actually exist. However, for example, representations on the geometric component X_M^{geom} always exist in a neighborhood of the discrete faithful representation, if we choose $u = \log m$ to be close to zero (and use a regular triangulation).

We then start expanding the state integral around a critical point, setting

$$\mathcal{Z}_{\mathcal{T}}(u;\hbar) \sim \hbar^{-\frac{3}{2}} \exp\left[\frac{1}{\hbar} S_{\mathcal{T},0}(u) + S_{\mathcal{T},1}(u) + \hbar S_{\mathcal{T},2}(m) + \hbar^2 S_{\mathcal{T},3}(m) + \cdots\right]. \tag{5-25}$$

The leading contribution $S_{\mathcal{T},0}(\rho)$ is given, following some standard manipulations using the generalized flattening, by

$$S_{\mathcal{T},0}(u) = uv(u) - \frac{1}{2}(Z - i\pi f)\cdot(Z'' + i\pi f'') + \sum_{i=1}^{N} \widetilde{\mathrm{Li}}_2(e^{-Z}) \quad \left(\mathrm{mod}\ \frac{\pi^2}{6}\right). \tag{5-26}$$

Here we write $S_{\mathcal{T},0}$ as a function of the logarithmic meridian eigenvalue u, though a fixed choice of representation ρ will implicitly fix the choice of longitude eigenvalue $v = \log(-\ell)$ as well. Expression (5-26) is a holomorphic version of the complex volume of a cusped manifold with deformed cusp. Explicitly,

$$S_{\mathcal{T},0}(u) = i\,(\mathrm{Vol}_M(u) + i\,\mathrm{CS}_M(u)) - 2v\Re(u). \tag{5-27}$$

This is the correct form of the complex volume to use in the Generalized Volume Conjecture; cf [40].

At first subleading order, we rederive the generalized torsion formula. The calculation is identical to that of Section 5.3, with the exception of the correction (5-13) coming from the $\hbar^0$ part of the exponential. This correction now becomes

$$\exp\Big[f\cdot \boldsymbol{B}^{-1}\boldsymbol{u} + \frac{i\pi}{2} f\cdot \boldsymbol{B}^{-1}\nu - \frac{1}{2} Z\cdot \boldsymbol{B}^{-1}\nu\Big]. \tag{5-28}$$

To simplify this correction, we must use $\boldsymbol{A} f + \boldsymbol{B} f'' = \nu$ and a deformed gluing equation $\boldsymbol{A} Z + \boldsymbol{B} Z'' = 2\boldsymbol{u} + i\pi\nu$. The $\boldsymbol{u}$–dependent part of the gluing equation cancels the new $\boldsymbol{u}$–dependent term in (5-28), ultimately leading to the same result

$$\exp\Big[f\cdot \boldsymbol{B}^{-1}\boldsymbol{u} + \frac{i\pi}{2} f\cdot \boldsymbol{B}^{-1}\nu - \frac{1}{2} Z\cdot \boldsymbol{B}^{-1}\nu\Big] = (z^{f''} z''^{-f})^{-1/2},$$

and therefore the same torsion[8]

$$\tau_{\mathcal{T}} = 4\pi e^{-2S_{\mathcal{T},1}} = \frac{1}{2}\det(\boldsymbol{A}\Delta_{z''} + \boldsymbol{B}\Delta_z^{-1}) z^{f''} z''^{-f}. \tag{5-29}$$

Finally, we can produce a generalized version of the Feynman rules of Section 5.4. We note, however, that the u–dependent terms in (5-23) do not contribute to either the vacuum energy $\Gamma^{(0)}$ (at order $\hbar^1$ or higher), the propagator, or the vertex factors $\Gamma_i^{(k)}$. Therefore, the Feynman rules must look exactly the same. The only difference is that the critical point Equation (5-24) requires us to use shape parameters that satisfy the generalized gluing equations.

5.6 Example: 4_1 completed

We may demonstrate the power of the Feynman diagram approach by computing the first two subleading corrections $S_{\mathcal{T},2}$ and $S_{\mathcal{T},3}$ for the figure-eight knot complement.

[8] The normalization of the torsion here differs from the torsion at the discrete faithful by a factor of π^2. In fact, we intentionally changed the normalization of the entire state integral (5-23) by π^2. This is because we wanted the state integral to match the asymptotics of the colored Jones polynomials exactly, and the asymptotics happen to jump by π^2 when $u \neq 0$, cf [40].

We can use the same Neumann–Zagier datum described in Section 2.6, along with the generalized flattening of Section 4.6. Let us specialize to representations ρ_m on the geometric component of the character variety. Then the two shapes z, w are expressed as functions on the A–polynomial curve,

$$z = -\frac{m^2 - m^{-2}}{1 + m^2\ell}, \quad w = \frac{m^2 + \ell}{m^2 - m^{-2}}, \tag{5-30}$$

as in (4-21).

The 2–loop invariant is explicitly given in (1-19) of Section 1. Evaluating this expression in `Mathematica`, we find

$$S_{\mathbf{4_1},2} = -\frac{w^3(z+1) + w^2((11-8z)z-4) + w(z-1)(z(z+12)-5) + (z-2)(z-1)^2}{12(w+z-1)^3}.$$

Upon using (5-30) to substitute rational functions for z and w, the answer may be most simply expressed as

$$\widetilde{S}_{\mathbf{4_1},2} = \frac{S_{\mathbf{4_1},2} + 1/8}{\tau_{\mathbf{4_1}}^3} = -\frac{1}{192}(m^{-6} - m^{-4} - 2m^{-2} + 15 - 2m^2 - m^4 + m^6), \tag{5-31}$$

where we have divided by a power of the torsion as suggested in (1-14) of Section 1. We have also absorbed a constant $1/8$, recalling that our formula is only well-defined modulo $\mathbb{Z}/24$.

In a similar way, we may calculate the 3–loop invariant, finding unambiguously

$$\widetilde{S}_{\mathbf{4_1},3} = \frac{S_{\mathbf{4_1},3}}{\tau_{\mathbf{4_1}}^6} = \frac{1}{128}(m^{-6} - m^{-4} - 2m^{-2} + 5 - 2m^2 - m^4 + m^6). \tag{5-32}$$

These answers agree perfectly with the findings of [17], and the comparison there to the asymptotics of the colored Jones polynomials at general u. Moreover, at the discrete faithful representation we obtain

$$S_{\mathbf{4_1},2} = \frac{11i}{72\sqrt{3}} = -\frac{11}{192\tau_{\mathbf{4_1}}^3}, \quad S_{\mathbf{4_1},3} = -\frac{1}{54} = \frac{1}{128\tau_{\mathbf{4_1}}^6}, \tag{5-33}$$

in agreement with known asymptotics of the Kashaev invariant.

Appendix A: Symplectic properties of A and B

The $N \times N$ Neumann–Zagier matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ form the top half of a symplectic matrix $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{smallmatrix}\right) \in \mathrm{Sp}(2N, \mathbb{Q})$ [58]. In this section we discuss some elementary properties of symplectic matrices.

Lemma A.1 *The $N \times 2N$ matrix $(\boldsymbol{A}\,\boldsymbol{B})$ is the upper half of a symplectic matrix if and only if $\boldsymbol{A}\boldsymbol{B}^T$ is symmetric and $(\boldsymbol{A}\,\boldsymbol{B})$ has maximal rank N.*

Proof It is easy to see that the rows of $(\boldsymbol{A}\,\boldsymbol{B})$ have zero symplectic product (with respect to the standard symplectic form on $\mathbb{Q}^{2N}$ if and only if $\boldsymbol{A}\boldsymbol{B}^T$ is symmetric. In addition they span a vector space of rank N if and only if $(\boldsymbol{A}\,\boldsymbol{B})$ has maximal rank N. The result follows. □

Lemma A.2 *If $(\boldsymbol{A}\,\boldsymbol{B})$ is the upper half of a symplectic matrix and $\boldsymbol{B}$ is nondegenerate, then $\boldsymbol{B}^{-1}\boldsymbol{A}$ is symmetric.*

Proof Lemma A.1 implies that $\boldsymbol{A}\boldsymbol{B}^T$ is symmetric, and so is $(\boldsymbol{B})^{-1}\boldsymbol{A}\boldsymbol{B}^T((\boldsymbol{B})^{-1})^T$. □

It is not true in general that $\boldsymbol{B}$ is invertible. However, after a possible change of quad type, we can assume that $\boldsymbol{B}$ is invertible. This is the content of the next lemma.

Lemma A.3 (a) *Suppose $(\boldsymbol{A}\,\boldsymbol{B})$ is the upper half of a symplectic $2N \times 2N$ matrix. If $\boldsymbol{A}$ has rank r, then any r linearly independent columns of $\boldsymbol{A}$ and their complementary $N-r$ columns in $\boldsymbol{B}$ form a basis for the column space of $(\boldsymbol{A}\,\boldsymbol{B})$.*

(b) *There always exists a choice of quad type for which $\boldsymbol{B}$ is nondegenerate (for any fixed choice of redundant edge and meridian path).*

Proof For (a) let $\operatorname{rank}(\boldsymbol{A}) = r \le N$. Without loss of generality, we may suppose that the first r columns of $\boldsymbol{A}$ are linearly independent. We want to show that, together with the last $N-r$ columns of $\boldsymbol{B}$, they form a matrix of rank N.

If we simultaneously multiply both $\boldsymbol{A}$ and $\boldsymbol{B}$ on the left by any nonsingular matrix $U \in \mathrm{GL}(N,\mathbb{R})$, both the symplectic condition and the columns are preserved. This follows from the fact that $\left(\begin{smallmatrix} U & 0 \\ 0 & U^{-1,T} \end{smallmatrix}\right) \in \mathrm{Sp}(2N,\mathbb{R})$. By allowing such a transformation, we may assume that $\boldsymbol{A}$ takes the block form

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_{r\times r} & A_2 \\ 0 & 0 \end{pmatrix} \tag{A-1}$$

for some A_2. Similarly, we split $\boldsymbol{B}$ into blocks of size r and $N-r$,

$$\boldsymbol{B} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}. \tag{A-2}$$

Since $(\boldsymbol{A}\,\boldsymbol{B})$ has full (row) rank, we see that the bottom $N-r$ rows of $\boldsymbol{B}$ must be linearly independent, ie $\operatorname{rank}(B_3 B_4) = N-r$. From the symplectic condition of

Lemma A.1, we also find that $B_3 + B_4 A_2^T = 0$, so that $\operatorname{rank}(B_3 B_4) \le \operatorname{rank}(B_4)$. This then implies that B_4 itself must have maximal rank $N-r$. Therefore, the last $N-r$ columns of $\boldsymbol{B}$ are linearly independent, and also independent of the first N columns of $\boldsymbol{A}$; ie the matrix $\left(\begin{smallmatrix} \boldsymbol{I}_{r\times r} & B_2 \\ 0 & B_4 \end{smallmatrix}\right)$ has maximal rank as desired. This concludes the proof of part (a).

For part (b) let us denote the columns of $\boldsymbol{A}$ and $\boldsymbol{B}$ as a_i and b_i. A change of quad type corresponding to a cyclic permutation $Z_i \mapsto Z_i' \mapsto Z_i'' \mapsto Z_i$ on the i^{th} tetrahedron permutes the i^{th} columns of $\boldsymbol{A}$ and $\boldsymbol{B}$ as $(a_i, b_i) \mapsto (b_i - a_i, -a_i)$. Therefore, given N complementary columns of $(\boldsymbol{A}\,\boldsymbol{B})$ that have full rank, we can use such permutations to move all the columns (up to a sign) into $\boldsymbol{B}$. □

Appendix B: The shape parameters are rational functions on the character variety

In this Appendix, we prove that the shape parameters of a regular ideal triangulation are rational functions on Y_M^{geom}, the geometric component of the $\mathrm{SL}_2(\mathbb{C})$ A–polynomial curve.

Proposition B.1 *Fix a regular ideal triangulation $\mathcal{T}$ of a one-cusped hyperbolic manifold M. Then every shape parameter of $\mathcal{T}$ is a rational function on Y_M^{geom}.*

Proof The proof is a little technical, and follows from work of Dunfield [20, Corollary 3.2], partially presented in [7, Appendix]. For completeness, we give the details of the proof here. We thank N Dunfield for a careful explanation of his proof to us.

Consider the affine variety $R(M, \mathrm{SL}(2,\mathbb{C})) = \mathrm{Hom}(\pi_1, \mathrm{SL}(2,\mathbb{C}))$ and its algebrogeometric quotient $X_{M,\mathrm{PSL}(2,\mathbb{C})}$ by the conjugation action of $\mathrm{PSL}(2,\mathbb{C})$. Following Dunfield from the Appendix to [7], let $\bar{R}(M, \mathrm{SL}(2,\mathbb{C}))$ denote the subvariety of $R(M, \mathrm{SL}(2,\mathbb{C})) \times P^1(\mathbb{C})$ consisting of pairs (ρ, z) where z is a fixed point of $\rho(\pi_1(\partial M))$. Let $\bar{X}_{M,\mathrm{SL}(2,\mathbb{C})}$ denote the algebrogeometric quotient of $\bar{R}(M, \mathrm{SL}(2,\mathbb{C}))$ under the diagonal action of $\mathrm{SL}(2,\mathbb{C})$ by conjugation and Möbius transformations respectively. We will call elements $(\rho, z) \in \bar{R}(M, \mathrm{SL}(2,\mathbb{C}))$ *augmented representations*. Their images in the augmented character variety $\bar{X}(M, \mathrm{SL}(2,\mathbb{C}))$ will be called *augmented characters* and will be denoted by square brackets $[(\rho, z)]$. Likewise, replacing $\mathrm{SL}(2,\mathbb{C})$ by $\mathrm{PSL}(2,\mathbb{C})$, we can define the character variety $X_{M,\mathrm{PSL}(2,C)}$ and its augmented version $\bar{X}_{M,\mathrm{PSL}(2,\mathbb{C})}$.

The advantage of the augmented character variety $\bar{X}_{M,\mathrm{SL}(2,\mathbb{C})}$ is that given $\gamma \in \pi_1(\partial M)$ there is a regular function e_γ that sends $[(\rho, z)]$ to the eigenvalue of $\rho(\gamma)$ corresponding to z, using Lemma B.3 below. In contrast, in $X_{M,\mathrm{SL}(2,\mathbb{C})}$ only the trace $e_\gamma + e_\gamma^{-1}$ of $\rho(\gamma)$ is well-defined. Likewise, in $\bar{X}_{M,\mathrm{PSL}(2,C)}$ (resp. $X_{M,\mathrm{SL}(2,\mathbb{C})}$) only e_γ^2 (resp. $e_\gamma^2 + e_\gamma^{-2}$) is well-defined.

From now on, we will restrict to the geometric component of the character variety $X_{M,\mathrm{PSL}(2,\mathbb{C})}$ and we will fix a regular ideal triangulation $\mathcal{T}$. In [20, Theorem 3.1] Dunfield proves that the natural restriction map

$$X_{M,\mathrm{PSL}(2,\mathbb{C})} \longrightarrow X_{\partial M,\mathrm{PSL}(2,\mathbb{C})}$$

of affine curves is of degree 1. The variety $X_{\partial M,\mathrm{PSL}(2,\mathbb{C})}$ is an affine curve in $(\mathbb{C}^*)^2/\mathbb{Z}_2$ and let $V_{M,\mathrm{PSL}(2,\mathbb{C})} \subset (\mathbb{C}^*)^2$ denote the preimage of $X_{\partial M,\mathrm{PSL}(2,\mathbb{C})}$ of the $2:1$ map $(\mathbb{C}^*)^2 \longrightarrow (\mathbb{C}^*)^2/\mathbb{Z}_2$. The commutative diagram

(B-1)
$$\begin{tikzcd}
\bar{X}_{M,\mathrm{PSL}(2,\mathbb{C})} \arrow[r] \arrow[d] & V_{M,\mathrm{PSL}(2,\mathbb{C})} \arrow[d] \\
X_{M,\mathrm{PSL}(2,\mathbb{C})} \arrow[r] & X_{\partial M,\mathrm{PSL}(2,\mathbb{C})}
\end{tikzcd}$$

has both vertical maps of degree 2, and the bottom horizontal map of degree 1. Thus, it follows that the top horizontal map is of degree 1. In [7, Section 10.3] Dunfield constructs a degree 1 developing map

$$V_{\mathcal{T}} \longrightarrow \bar{X}_{M,\mathrm{PSL}(2,\mathbb{C})},$$

which combined with the previous discussion gives a chain of birational curve isomorphisms

(B-2)
$$V_{\mathcal{T}} \longrightarrow \bar{X}_{M,\mathrm{PSL}(2,\mathbb{C})} \longrightarrow V_{M,\mathrm{PSL}(2,\mathbb{C})}.$$

Since the shape parameters are rational (in fact coordinate) functions on $V_{\mathcal{T}}$, it follows that they are rational functions on $V_{M,\mathrm{PSL}(2,\mathbb{C})}$. Using the regular map $V_{M,\mathrm{SL}(2,\mathbb{C})} \longrightarrow V_{M,\mathrm{PSL}(2,\mathbb{C})}$, we obtain that the shape parameters are rational functions on $V_{M,\mathrm{SL}(2,\mathbb{C})}$. □

Proposition B.1 has the following concrete corollary.

Corollary B.2 *Given a regular ideal triangulation $\mathcal{T}$ with N tetrahedra, there is a solution of the shape parameters in $\mathbb{Q}(m,\ell)/(A(m,\ell))$.*

Lemma B.3 *Suppose*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C})$$

and $c \neq 0$. Then, λ is an eigenvalue of A if and only if $z = (\lambda - 2d)/(2c)$ is a fixed point of the corresponding Möbius transformation in $P^1(\mathbb{C})$.

Appendix C: Deriving the state integral

In this Appendix, we explain the connection between the quantization formalism of [15] and the special state integrals (5-2) and (5-23) that led to all the formulas in the present paper. We will first review classical "symplectic gluing" of tetrahedra, then extend gluing to the quantum setting and construct the state integral. There are multiple points in the construction that have yet to be made mathematically rigorous, which we will try to indicate.

C.1: Symplectic gluing

The main idea of [15] is that gluing of tetrahedra should be viewed, both classically and quantum mechanically, as a process of symplectic reduction.

Suppose we have a one-cusped manifold M with a triangulation $\mathcal{T} = \{\Delta_i\}_{i=1}^N$. Classically, each tetrahedron Δ_i comes with a phase space

(C-1)
$$\begin{aligned}\mathcal{P}_{\partial\Delta_i} &= \{\text{flat } \mathrm{SL}(2,\mathbb{C}) \text{ connections on } \partial\Delta_i\}\\ &\approx \{(Z_i, Z_i', Z_i'') \in \mathbb{C}\backslash(2\pi i\mathbb{Z}) \mid Z_i + Z_i' + Z_i'' = i\pi\},\end{aligned}$$

with (holomorphic) symplectic structure

(C-2)
$$\omega_{\partial\Delta_i} = dZ \wedge dZ'',$$

and a Lagrangian submanifold[9]

(C-3)
$$\begin{aligned}\mathcal{L}_{\Delta_i} &= \{\text{flat } \mathrm{SL}(2,\mathbb{C}) \text{ connections that extend to } \Delta_i\}\\ &= \{e^{Z''} + e^{-Z} - 1 = 0\} \subset \mathcal{P}_{\partial\Delta_i}.\end{aligned}$$

When gluing the tetrahedra together, we first form a product

(C-4)
$$\mathcal{L}_\times = \mathcal{L}_{\Delta_1} \times \cdots \times \mathcal{L}_{\Delta_N} \subset \mathcal{P}_\times = \mathcal{P}_{\partial\Delta_1} \times \cdots \times \mathcal{P}_{\partial\Delta_N}.$$

The edge constraints $X_I := \sum_{i=1}^N (\boldsymbol{G}_{Ii} Z_i + \boldsymbol{G}_{Ii}' Z_i' + \boldsymbol{G}_{Ii}'' Z_i'') - 2\pi i$ from (2-3) are functions on the product phase space $\mathcal{P}_\times$, and can be used as (holomorphic) moment maps to generate $N-1$ independent translation actions t_I. Recall [58] that the

[9] Explicitly, $\mathcal{P}_{\partial\Delta_i}$ is a space of flat connections on a 4–punctures sphere with parabolic holonomy at the four punctures; while $\mathcal{L}_{\Delta_i}$ is the subspace with trivial holonomy, hence connections that extend into the bulk of the tetrahedron. See, eg, the first author, Gaiotto and Gukov [16, Section 2].

logarithmic meridian and longitude holonomies (u,v) are also functions on $\mathcal{P}_\times$, which Poisson-commute with all the edges X_I, and so are fixed under these translations. Then the phase space of M is a symplectic quotient,

$$\begin{aligned}\mathcal{P}_{\partial M} &= \{\text{flat } \mathrm{SL}(2,\mathbb{C}) \text{ connections on } \partial M \simeq T^2\} \approx \{(u,v)\in\mathbb{C}\}\\ &= \mathcal{P}_\times /\!/ (t_I),\end{aligned} \tag{C-5}$$

and the A–polynomial of M (more properly, components of the A–polynomial for which the triangulation is regular) is the result of pulling the Lagrangian $\mathcal{L}_\times$ through the quotient,

$$\mathcal{L}_M = \text{“}\mathcal{L}_\times /\!/ (t_I)\text{”} \approx \{A_M(e^v, e^u) = 0\} \subset \mathcal{P}_{\partial M}. \tag{C-6}$$

This is quite easy to check using equations (4-5) and (4-6).

C.2: Quantization

Quantum mechanically, we have that each tetrahedron has a Hilbert space $\mathcal{H}_{\Delta_i}$, a wavefunction $\mathcal{Z}_{\Delta_i}(Z_i)$ and a quantum operator $\widehat{\mathcal{L}}_{\partial\Delta_i}$ that annihilates the wavefunction. The symplectic-gluing procedure extends to the quantum setting, with appropriate quantum generalizations of all the above operations. Roughly, one forms a product wavefunction

$$\mathcal{Z}_\times(Z_1,\dots,Z_N) = \mathcal{Z}_{\Delta_1}\otimes\cdots\otimes\mathcal{Z}_{\Delta_N} \in \mathcal{H}_\times = \mathcal{H}_{\partial\Delta_1}\otimes\cdots\otimes\mathcal{H}_{\partial\Delta_N}, \tag{C-7}$$

and restricts the product Hilbert space using $N-1$ new polarizations coming from the edge constraints. The resulting restricted wavefunction is $\mathcal{Z}_M(u)$, and it is annihilated by a quantized version of the A–polynomial; see the second author [39; 33].

To make this more precise, let M again be an oriented one-cusped manifold, and choose a triangulation $\mathcal{T} = \{\Delta_i\}_{i=1}^N$ (regular with respect to some desired family of representations), a quad type, a redundant edge, and a meridian path, just as in Section 2.

To each tetrahedron Δ_i we associate a boundary Hilbert space $\mathcal{H}_{\partial\Delta_i}$. It is some extension[10] of $L^2(\mathbb{R})$ that includes the wavefunction

$$\mathcal{Z}_{\Delta_i}(Z_i;\hbar) := \psi_\hbar(Z_i), \tag{C-8}$$

where $\psi_\hbar(Z_i)$ is Faddeev's quantum dilogarithm (5-5) [27]. We also associate to Δ_i an algebra of operators

$$\widehat{\mathcal{A}}_{\partial\Delta_i} = \mathbb{C}\langle \widehat{Z}_i, \widehat{Z}'_i, \widehat{Z}''_i\rangle / (\widehat{Z}_i + \widehat{Z}'_i + \widehat{Z}''_i = i\pi + \tfrac{\hbar}{2}), \tag{C-9}$$

[10] This space has not been mathematically defined yet; constructions of (eg) [2] might prove useful for achieving this.

with commutation relations

$$[\widehat{Z}_i, \widehat{Z}'_i] = [\widehat{Z}'_i, \widehat{Z}''_i] = [\widehat{Z}''_i, \widehat{Z}_i] = \hbar. \tag{C-10}$$

Then the quantization of the Lagrangian (C-3) annihilates the wavefunction,

$$\widehat{\mathcal{L}}_{\Delta_i} := e^{\widehat{Z}''_i} + e^{-\widehat{Z}_i} - 1, \quad \widehat{\mathcal{L}}_{\Delta_i} \mathcal{Z}_{\Delta_i} = 0, \tag{C-11}$$

where the operators act in the representation

$$\begin{aligned} \widehat{Z}_i &= Z_i, & \widehat{Z}''_i &= \hbar \partial_{Z_i}; \quad \text{or} \\ e^{\widehat{Z}_i} \mathcal{Z}(Z_i) &= e^{Z_i} \mathcal{Z}(Z_i), & e^{\widehat{Z}''_i} \mathcal{Z}(Z_i) &= \mathcal{Z}(Z_i + \hbar). \end{aligned} \tag{C-12}$$

In order to glue the tetrahedra together, we start by forming the product wavefunction $\mathcal{Z}_\times(Z_1, \dots, Z_N) = \mathcal{Z}_{\Delta_1}(Z_1) \cdots \mathcal{Z}_{\Delta_N}(Z_N)$. This is an element of a product Hilbert space (C-7). Acting on this product Hilbert space is the product $\widehat{\mathcal{A}}_\times$ of algebras (C-9), which is simply generated by all the $\widehat{Z}_i, \widehat{Z}'_i, \widehat{Z}''_i$, with canonical commutation relations (C-10) (and operators from distinct tetrahedra always commuting).

Now, following the notation of Sections 2.2 and 4.4, we can define N operators $\widehat{X}_I \in \widehat{\mathcal{A}}_\times$, one for each independent edge, and one for the meridian:

$$\widehat{X}_I := \begin{cases} \sum_{i=1}^N (\boldsymbol{G}_{Ii} \widehat{Z}_i + \boldsymbol{G}'_{Ii} \widehat{Z}'_i + \boldsymbol{G}''_{Ii} \widehat{Z}''_i) - 2\pi i - \hbar & I = 1, \dots, N-1, \\ \boldsymbol{G}_{N+1,i} \widehat{Z}_i + \boldsymbol{G}'_{N+1,i} \widehat{Z}'_i + \boldsymbol{G}''_{N+1,i} \widehat{Z}''_i & I = N. \end{cases} \tag{C-13}$$

Similarly, we may define an operator

$$\widehat{P}_N := \tfrac{1}{2} (\boldsymbol{G}_{N+2,i} \widehat{Z}_i + \boldsymbol{G}'_{N+2,i} \widehat{Z}'_i + \boldsymbol{G}''_{N+2,i} \widehat{Z}''_i) \tag{C-14}$$

corresponding to the longitude. Due to the symplectic structure found in [58], we know that we may complete the set $\{\widehat{X}_1, \dots, \widehat{X}_N, \widehat{P}_N\}$ to a full canonical basis of the algebra $\widehat{\mathcal{A}}_\times$. We do this by adding $N-1$ additional operators $\widehat{P}_I$, which are linear combinations of the $\widehat{Z}$s, such that

$$[\widehat{P}_I, \widehat{X}_j] = \delta_{Ij} \hbar, \quad [\widehat{P}_I, \widehat{P}_j] = [\widehat{X}_I, \widehat{X}_j] = 0, \quad 1 \le I, j \le N. \tag{C-15}$$

The operators $\widehat{X}_I, \widehat{P}_I$ have a simple interpretation in terms of a generalized Neumann–Zagier datum. Namely, if we complete $(\boldsymbol{A}\,\boldsymbol{B})$ and the rows C, D (of Section 4.4) to a full symplectic matrix $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{smallmatrix}\right)$, then

$$\begin{pmatrix} \widehat{X} \\ \widehat{P} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix} \begin{pmatrix} \widehat{Z} \\ \widehat{Z}'' \end{pmatrix} - (i\pi + \tfrac{\hbar}{2}) \begin{pmatrix} \nu \\ \nu_P \end{pmatrix}. \tag{C-16}$$

Here ν is precisely the vector of N integers which was introduced in (2-8), while $\nu_P = (*, \dots, *, \nu_\lambda)$, with ν_λ from (4-8). The first $N-1$ entries of ν_P depend on the precise completion of the canonical basis (or the symplectic matrix), and ultimately drop out of the gluing construction.

C.3: Quantum reduction

Classically, in order to glue we would want to set the $N-1$ edge constraints $X_I \to 0$, and the meridian $X_N \to 2u$. In Section C.1, these functions were actually used as moment maps to perform a symplectic reduction. Now we should do the same thing quantum mechanically. In order to reduce the product wavefunction $\mathcal{Z}_\times(Z_1, \dots, Z_N)$ of (C-7) to the final wavefunction $\mathcal{Z}_M(u)$ of the glued manifold M, we must transform the wavefunction to a representation (or "polarization") in which the operators $\widehat{X}_I$ act diagonally (by multiplication). In this representation, the wavefunction depends explicitly on the X_I. The "reduction" then simply requires fixing $X_I \to (0, \dots, 0, 2u)$. Schematically,

$$\mathcal{Z}_\times(Z_1, \dots, Z_N) \xrightarrow{\text{transform}} \widetilde{\mathcal{Z}}_\times(X_1, \dots, X_N) \overset{\text{fix}}{\mapsto} \mathcal{Z}_M(u) = \widetilde{\mathcal{Z}}_\times(0, \dots, 0, 2u). \tag{C-17}$$

The transformation from $\mathcal{Z}_\times$ to $\widetilde{\mathcal{Z}}_\times$ is accomplished— formally—with the Weil representation $\mathcal{R}$ of the affine symplectic group; see Shale [64] and Weil [71]. In particular, we need $\mathcal{R}(\alpha)$ for the affine symplectic transformation α in (C-16). In [15, Section 6], it was discussed in detail how to find $\mathcal{R}(\alpha)$ by factoring the matrix of (C-16) into generators. Then, for example, an "S–type" element of the symplectic group acts via Fourier transform

$$\mathcal{R}\left(\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\right)\colon f(Z) \mapsto \tilde{f}(W) = \int \frac{d^N Z}{(2\pi i \hbar)^{N/2}} e^{\frac{1}{\hbar} Z\cdot W} f(Z), \tag{C-18}$$

whereas a "T–type" element acts as multiplication by a quadratic exponential

$$\mathcal{R}\left(\begin{pmatrix} I & 0 \\ \boldsymbol{T} & I \end{pmatrix}\right)\colon f(Z) \mapsto \tilde{f}(W) = e^{\frac{1}{2\hbar} W^T \boldsymbol{T} W} f(W). \tag{C-19}$$

Affine shifts act either by translation or multiplication by a linear exponential.

In the present case, there is a convenient trick that allows us to find $\mathcal{R}(\alpha)$ without decomposing α into generators. We assume that the block $\boldsymbol{B}$ of the symplectic matrix is nondegenerate, since we know we can always choose a quad type with this property. For the moment, let us also suppose that the affine shifts vanish, $\nu = \nu_P = 0$. Then the Weil action is

$$\mathcal{R}(\alpha)\colon \mathcal{Z}_\times(Z) \mapsto \widetilde{\mathcal{Z}}_\times(X) = \frac{1}{\sqrt{\det \boldsymbol{B}}} \int \frac{d^N Z}{(2\pi i \hbar)^{N/2}} e^{x} \mathcal{Z}_\times(Z), \tag{C-20}$$

where $x = \frac{1}{2\hbar}(X \cdot \boldsymbol{D}\boldsymbol{B}^{-1} X - 2Z \cdot \boldsymbol{B}^{-1} X + Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A} Z)$. In particular, it can easily be verified that this correctly intertwines an action of operators $(\widehat{Z}_i = Z_i, \widehat{Z}''_i = \hbar\partial_{Z_i})$ on $\mathcal{Z}_\times(Z)$ with an action of operators $(\widehat{X}_I = X_I, \widehat{P}_I = \hbar\partial_{X_I})$ on $\widetilde{\mathcal{Z}}_\times(X)$. For example,

$$\begin{aligned}\int d^N Z e^x (\boldsymbol{A}\widehat{Z} + \boldsymbol{B}\widehat{Z}'')\mathcal{Z}_\times(Z) &= \int d^N Z e^x (\boldsymbol{A}Z + \hbar\boldsymbol{B}\partial_Z)\mathcal{Z}_\times(Z) \\ &= \int d^N Z[(\boldsymbol{A}Z - \hbar\boldsymbol{B}\partial_Z)e^x]\mathcal{Z}_\times(Z) \\ &= \int d^N Z X e^x \mathcal{Z}_\times(Z) \\ &= \widehat{X}\widetilde{\mathcal{Z}}_\times(X).\end{aligned}$$

Nonzero affine shifts ν and ν_P further modify the result to

$$\begin{aligned}\widetilde{\mathcal{Z}}_\times(X) = \frac{1}{\sqrt{\det \boldsymbol{B}}} \int \frac{d^N Z}{(2\pi i\hbar)^{N/2}} \exp\Big[-\frac{1}{\hbar} X \cdot \Big(i\pi + \frac{\hbar}{2}\Big)\nu_P \\ + \frac{1}{2\hbar}\Big(\Big(X + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) \cdot \boldsymbol{D}\boldsymbol{B}^{-1}\Big(X + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) \\ - 2Z \cdot \boldsymbol{B}^{-1}\Big(X + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) + Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A}Z\Big)\Big]\mathcal{Z}_\times(Z),\end{aligned}$$

and then, after setting $X \to 2\boldsymbol{u} = (0, \dots, 0, 2u)$ as in (C-17), we find

$$\begin{aligned}\text{(C-21)} \quad \mathcal{Z}_M(u) = \frac{1}{\sqrt{\det \boldsymbol{B}}} \int \frac{d^N Z}{(2\pi i\hbar)^{N/2}} \exp\Big[-\frac{1}{\hbar}(2\pi i + \hbar)\nu_\lambda u \\ + \frac{1}{2\hbar}\Big(\Big(2\boldsymbol{u} + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) \cdot \boldsymbol{D}\boldsymbol{B}^{-1}\Big(2\boldsymbol{u} + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) \\ - 2Z \cdot \boldsymbol{B}^{-1}\Big(2\boldsymbol{u} + \Big(i\pi + \frac{\hbar}{2}\Big)\nu\Big) + Z \cdot \boldsymbol{B}^{-1}\boldsymbol{A}Z\Big)\Big] \prod_{i=1}^{N} \psi_\hbar(Z_i).\end{aligned}$$

This is the partition function of the one-cusped manifold M, modulo a multiplicative ambiguity of the form $\exp[\frac{\pi^2}{6}a + \frac{i\pi}{4}b + \frac{1}{24}c]$ for $a, b, c \in \mathbb{Z}$, which we will say more about in Section C.5. By *construction*, this partition function is annihilated by the quantum $\widehat{A}$–polynomial of M.

C.4: Introducing a flattening

In order to obtain the state integral (5-23) appearing in the paper, we can introduce a generalized flattening (as in Section 4.4) and use it to simplify (C-21). Note that the discrete-faithful state integral (5-2) follows immediately from (5-23) upon setting $\boldsymbol{u} = (0, \dots, 0, u) \to 0$.

Suppose, then, that we have integers (f, f'') that satisfy

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}\begin{pmatrix} f \\ f'' \end{pmatrix} = \begin{pmatrix} \nu \\ \nu_{\boldsymbol{P}} \end{pmatrix}, \tag{C-22}$$

for some $\nu_{\boldsymbol{P}}$ whose last entry is ν_λ. We will *assume* that a completed symplectic matrix $\left(\begin{smallmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{smallmatrix}\right)$ can be chosen in $\mathrm{Sp}(2N,\mathbb{Z})$ rather than in $\mathrm{Sp}(2N,\mathbb{Q})$. In that case, since f and f'' are vectors with integer entries, it follows that $\nu_{\boldsymbol{P}} \in \mathbb{Z}^N$. Then,

$$-\nu_\lambda u + \nu\cdot \boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{u} = -(\boldsymbol{C} f + \boldsymbol{D} f'')\cdot \boldsymbol{u} + (\boldsymbol{D}^T\boldsymbol{A} f + \boldsymbol{D}^T\boldsymbol{B} f'')\cdot \boldsymbol{B}^{-1}\boldsymbol{u} = f\cdot \boldsymbol{B}^{-1}\boldsymbol{u},$$

where we used the symplectic identities $\boldsymbol{D}^T\boldsymbol{B} = \boldsymbol{B}^T\boldsymbol{D}$ and $\boldsymbol{D}^T\boldsymbol{A} = I + \boldsymbol{B}^T\boldsymbol{C}$; and

$$\begin{aligned} \boldsymbol{\nu}\cdot \boldsymbol{D}\boldsymbol{B}^{-1}\nu &= \nu\cdot \boldsymbol{D}(f'' + \boldsymbol{B}^{-1}\boldsymbol{A} f) \\ &= \nu\cdot(\boldsymbol{D} f'' + \boldsymbol{C} f + \boldsymbol{B}^{-1,T} f) \\ &= f\cdot \boldsymbol{B}^{-1}\nu + \nu\cdot\nu_{\boldsymbol{P}} \\ &= f\cdot \boldsymbol{B}^{-1}\nu \pmod{\mathbb{Z}}, \end{aligned}$$

in a similar way. These relations allow us to write the state integral (C-21) as

$$\mathcal{Z}_M(u) = \frac{1}{\sqrt{\det \boldsymbol{B}}}\int \frac{d^N Z}{(2\pi\hbar)^{N/2}} e^x \prod_{i=1}^N \psi_\hbar(Z_i), \tag{C-23}$$

where $x = \frac{1}{\hbar}[2\boldsymbol{u}\cdot\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{u} + (2\pi i+\hbar) f\cdot \boldsymbol{B}^{-1}\boldsymbol{u} + \frac{1}{2}(i\pi+\frac{\hbar}{2})^2 f\cdot\boldsymbol{B}^{-1}\nu - Z\boldsymbol{B}^{-1}(2\boldsymbol{u} + (i\pi+\frac{\hbar}{2})\nu)]$, just as in (5-23). (We drop a factor of $\sqrt{i}$ from the measure, since it can be absorbed in the overall normalization ambiguity.)

C.5: Normalization and invariance

The normalization of Chern–Simons state integrals has always been a subtle issue. For the integral of [15], ambiguities in the normalization come from two sources: the projectivity of the Weil representation, and the incomplete invariance of the integral (even formally) under a change of "quad type" and a 2–3 move.

Let us consider the Weil representation first. We will assume that all symplectic matrices are in $\mathrm{Sp}(2N,\mathbb{Z})$, and that all shifts involve integers (like ν and $\nu_{\boldsymbol{P}}$) times $i\pi+\frac{\hbar}{2}$. This assumption (which, again, is only an observed property) allows us to improve on the estimates of [15, Equation (6.6)]. The Weil representation becomes a projective unitary representation of $\mathrm{ISp}(2N,\mathbb{Z}) \simeq \mathrm{Sp}(2N,\mathbb{Z})\ltimes[(i\pi+\frac{\hbar}{2})\mathbb{Z}]^{2N}$ on $L^2(\mathbb{R}^N)$, for $\hbar$ pure imaginary. Our Hilbert space $\mathcal{H}_\Delta^{\otimes 2N}$ is very close to $L^2(\mathbb{R}^N)$, so we may hope that the Weil representation is also unitary projective there. The most severe projective ambiguity arises from a violation of expected commutation relations between shifts and T–type transformations such as (C-19). This leads to projective factors of the form

$$\exp\Big[\frac{1}{2\hbar}\Big(i\pi+\frac{\hbar}{2}\Big)^2 a\Big]=\exp\Big[\Big(-\frac{\pi^2}{2\hbar}+\frac{i\pi}{2}+\frac{\hbar^2}{8}\Big)a\Big],\quad a\in\mathbb{Z}. \tag{C-24}$$

With the exception of factors like this, unitarity with respect to the norm

$$\|f\|^2=\int\frac{d^N Z}{(\pm 2\pi i\hbar)^{N/2}}|f(Z)|^2$$

may be used to normalize Weil transformations. For example, the factor in (C-20), $[(2\pi i\hbar)^N\det\boldsymbol{B}]^{-1/2}$, follows easily from formal manipulations on the integral transformation to demonstrate unitarity.

The lack of complete invariance under a change of quad type (cyclic permutation invariance) and a 2–3 move can also ruin the normalization of the state integral. The change of quad type was analyzed, formally, in [15, Section 6.2.1]. A cyclic permutation of a tetrahedron is accomplished by an affine version of the element $ST\in\mathrm{Sp}(2N;\mathbb{Z})$, under the Weil representation. The single-tetrahedron wavefunction transforms as

$$\psi_\hbar(Z)\mapsto\int\frac{dZ}{\sqrt{2\pi i\hbar}}e^{\frac{1}{2\hbar}(Z^2+2ZZ'-(2\pi i+\hbar)Z)}\psi_\hbar(Z) = e^{\frac{\pi^2}{6\hbar}\pm\frac{i\pi}{4}-\frac{\hbar}{24}}\psi_\hbar(Z'). \tag{C-25}$$

The last equality follows from the Fourier transform of the quantum dilogarithm; see Faddeev, Kashaev and Volkov [29] and Ponsot and Teschner [61]. This shows that the tetrahedron wavefunction is *invariant* under permutations, up to a factor

$$\exp\Big[\Big(\frac{\pi^2}{6\hbar}\pm\frac{i\pi}{4}-\frac{\hbar}{24}\Big)a\Big],\quad a\in\mathbb{Z}. \tag{C-26}$$

The analysis of the 2–3 move is slightly more involved. It was done in terms of operator algebra in [15], and then explained in terms of wavefunctions in [16, Section 6.2]. The main idea is that a 2–3 move can be done locally during the gluing procedure, by performing a formal, "local" transformation on the state integral. The crucial property involved is the Ramanujan-like identity for the quantum dilogarithm [29; 61], which expresses three quantum dilogarithms as an integral of two; for example,

$$\psi_\hbar(W_1')\psi_\hbar(W_2')\psi_\hbar(W_3')|_{W_1'+W_2'+W_3'=2\pi i+\hbar} \sim\int\frac{dZ}{\sqrt{2\pi i\hbar}}e^{\frac{1}{2\hbar}(Z^2+2W_2'Z-(2\pi i+\hbar)(W_1'+W_2'+Z))}\psi_\hbar(-Z)\psi_\hbar(Z-W_1') \tag{C-27}$$

which holds up to a factor that is again of the type (C-26).

Putting together all three effects, we find that we might be able to control the overall normalization of the state integral up to a factor of the form

$$\exp\Big[\frac{\pi^2}{6\hbar}a+\frac{i\pi}{4}b+\frac{\hbar}{24}c\Big],\quad a,b,c\in\mathbb{Z}. \tag{C-28}$$

Appendix D: Computer implementation and computations

An enhanced Neumann–Zagier datum is a tuple $(z, \boldsymbol{A}, \boldsymbol{B}, f)$ attached to a regular ideal triangulation of a cusped hyperbolic manifold M. The program `SnapPy` [13] in its `python` and `sage` implementation computes the gluing matrices $\boldsymbol{G}, \boldsymbol{G}', \boldsymbol{G}''$ of Sections 2.4 and 4.4; and therefore it can easily compute an enhanced Neumann–Zagier datum $\widehat{\beta}_{\mathcal{T}} = (z, \boldsymbol{A}, \boldsymbol{B}, f)$. The shape parameters z are algebraic numbers computed numerically to arbitrary precision (eg, 10000 digits) or exactly as algebraic numbers.

A `Mathematica` module of the authors computes (numerically or exactly) the n–loop invariants $S_{\mathcal{T},n}$ for $n=0,2,3$ as well as our torsion $\tau_{\mathcal{T}}$ given as input the Neumann–Zagier datum. As an example, consider the hyperbolic knot $\mathbf{9_{12}}$ with volume $8.836642343\dots$ and the `SnapPy` ideal triangulation with 10 tetrahedra. Its invariant trace field $E_{\mathbf{9_{12}}}$ is $\mathbb{Q}(x)$ where $x=-0.06265158\dots+i\,1.24990458\dots$ is a root of

$$\begin{aligned}x^{17}-8x^{16}+32x^{15}-89x^{14}+195x^{13}-353x^{12}+542x^{11}-719x^{10}+834x^9\\ -851x^8+764x^7-605x^6+421x^5-253x^4+130x^3-55x^2+18x-3=0.\end{aligned}$$

$E_{\mathbf{9_{12}}}$ is of type $[1,8]$ with discriminant $3\cdot 298171\cdot 5210119\cdot 156953399$. Our torsion is

$$\begin{aligned}\tau_{\mathbf{9_{12}}} &= \frac{1}{2}(15-7x-15x^2+55x^3-67x^4+81x^5-43x^6-112x^7+303x^8-488x^9\\ &\qquad +606x^{10}-595x^{11}+464x^{12}-289x^{13}+143x^{14}-49x^{15}+8x^{16})\\ &= -3.133657804174628986\dots+14.061239582208047255\dots i.\end{aligned}$$

The two and three-loop invariants simplify considerably when multiplied by $\tau_{\mathbf{9_{12}}}^3$ and $\tau_{\mathbf{9_{12}}}^6$ respectively and are given by

$$\begin{aligned}S_{\mathbf{9_{12}},2}\tau_{\mathbf{9_{12}}}^3 &= \frac{1}{2^6\cdot 3}(36263-194718x+503316x^2-971739x^3+1582041x^4\\ &\qquad -2152164x^5+2372779x^6-2109742x^7+1426659x^8\\ &\qquad -484152x^9-374803x^{10}+836963x^{11}-859483x^{12}\\ &\qquad +621288x^{13}-326550x^{14}+109607x^{15}-16840x^{16})\\ &= 398.62270435384630954\dots+948.91209325049603870\dots i,\end{aligned}$$

$$
\begin{aligned}
S_{\mathbf{9_{12}},3}\tau^6_{\mathbf{9_{12}}}
&= \tfrac{1}{2^7}(2320213 - 19092785x^1 + 72589953x^2 - 186402605x^3 + 382362100x^4 \\
&\qquad - 661985976x^5 + 982969902x^6 - 1258919324x^7 + 1402544816x^8 \\
&\qquad - 1359436057x^9 + 1134208276x^{10} - 803313515x^{11} + 473961630x^{12} \\
&\qquad - 225394732x^{13} + 80872920x^{14} - 19104127x^{15} + 2161102x^{16}) \\
&= 71793.64335382669630\ldots + 204530.00105728258992\ldots i.
\end{aligned}
$$

The norm $(N_1, N_2, N_3) = (N(\tau_{\mathbf{9_{12}}}), N(S_{\mathbf{9_{12}},2}\tau^3_{\mathbf{9_{12}}}), N(S_{\mathbf{9_{12}},3}\tau^6_{\mathbf{9_{12}}}))$ of the above algebraic numbers is given by

$$
\begin{aligned}
N_1 &= \frac{3\cdot 298171\cdot 5210119\cdot 156953399}{2^{17}}, \\
N_2 &= \frac{173137\cdot 2497646101\ldots\ldots 5575954409 \quad \text{(70 digits)}}{2^{102}\cdot 3^{17}}, \\
N_3 &= \frac{1601979456\ldots\ldots 5984185143 \quad \text{(100 digits)}}{2^{119}}.
\end{aligned}
$$

Recall that although $S_{2,\mathbf{9_{12}}}$ is defined modulo an integer multiple of $1/24$, $S_{3,\mathbf{9_{12}}}$ is defined without ambiguity and the numerator N_3 is a prime number of 103 digits.

For a computation of the Reidemeister torsion τ^{R}_M of the discrete faithful representation of a cusped hyperbolic manifold M, we use a theorem of Yamaguchi [74] to identify it with

$$
\tau^{\mathrm{R}}_M = \frac{1}{c_M}\frac{d\tau^{\mathrm{R}}_M(t)}{dt}\Big|_{t=1}
$$

where c_M is the cusp shape of M and $\tau^{\mathrm{R}}_M(t) \in E_M[t^{\pm 1}]$ is the *torsion polynomial* of M using the adjoint representation of $\mathrm{SL}(2,\mathbb{C})$. Using the `hypertorsion` package of N Dunfield (see [21]), we can compute τ^{R}_M as follows:

```
cd Genus-Comp
sage:import snappy, hypertorsion

def torsion(manifold, precision=100):
    M = snappy.Manifold(manifold)
    p = hypertorsion.hyperbolic_adjoint_torsion(M, precision)
    q = p.derivative()
    rho = hypertorsion.polished_holonomy(M, precision)
    z = rho.cusp_shape()
    torsion = q(1)/z.conjugate()
    return [M.name(), torsion]
```

For the above example, we have

```
sage: torsion("9_12",500)
['L105002',  -3.13365780417462898 6\ldots
+ 14.0612395822080 47255\ldots*I]
```

numerically confirming Conjecture 1.8. Further computations gives a numerical confirmation of Conjecture 1.8 to 1000 digits for all 59924 hyperbolic knots with at most 14 crossings.

References

[1] **M Aganagic**, **V Bouchard**, **A Klemm**, *Topological strings and (almost) modular forms*, Comm. Math. Phys. 277 (2008) 771–819 MR2365453

[2] **J E Andersen**, **R Kashaev**, *A TQFT from quantum Teichmüller theory* arXiv:1109.6295

[3] **D Bar-Natan**, **E Witten**, *Perturbative expansion of Chern–Simons theory with noncompact gauge group*, Comm. Math. Phys. 141 (1991) 423–440 MR1133274

[4] **E W Barnes**, *The genesis of the double gamma functions*, Proc. London Math. Soc. 31 (1899) 358 MR1576719

[5] **R Benedetti**, **C Petronio**, *Branched standard spines of 3–manifolds*, Lecture Notes in Mathematics 1653, Springer, Berlin (1997) MR1470454

[6] **D Bessis**, **C Itzykson**, **J B Zuber**, *Quantum field theory techniques in graphical enumeration*, Adv. in Appl. Math. 1 (1980) 109–157 MR603127

[7] **D W Boyd**, **N M Dunfield**, **F Rodriguez-Villegas**, *Mahler's measure and the dilogarithm (II)* arXiv:0308041

[8] **B A Burton**, *Regina: Normal surface and 3–manifold topology software* Available at http://regina.sourceforge.net

[9] **D Calegari**, *Real places and torus bundles*, Geom. Dedicata 118 (2006) 209–227 MR2239457

[10] **B G Casler**, *An imbedding theorem for connected 3–manifolds with boundary*, Proc. Amer. Math. Soc. 16 (1965) 559–566 MR0178473

[11] **A A Champanerkar**, *A–polynomial and Bloch invariants of hyperbolic 3–manifolds*, PhD thesis, Columbia University (2003) MR2704573 Available at http://search.proquest.com/docview/305332823

[12] **D Cooper**, **M Culler**, **H Gillet**, **D D Long**, **P B Shalen**, *Plane curves associated to character varieties of 3–manifolds*, Invent. Math. 118 (1994) 47–84 MR1288467

[13] **M Culler**, **N M Dunfield**, **J R Weeks**, *SnapPy, a computer program for studying the geometry and topology of 3–manifolds* (2011) Available at http://snappy.computop.org

[14] **J A De Loera**, **J Rambau**, **F Santos**, *Triangulations*, Algorithms and Computation in Mathematics 25, Springer, Berlin (2010) MR2743368

[15] **T Dimofte**, *Quantum Riemann surfaces in Chern–Simons theory* arXiv:1102.4847

[16] **T Dimofte**, **D Gaiotto**, **S Gukov**, *Gauge theories labelled by three-manifolds* arXiv:1108.4389

[17] **T Dimofte**, **S Gukov**, **J Lenells**, **D Zagier**, *Exact results for perturbative Chern–Simons theory with complex gauge group*, Commun. Number Theory Phys. 3 (2009) 363–443 MR2551896

[18] **J Dubois**, *Non abelian twisted Reidemeister torsion for fibered knots*, Canad. Math. Bull. 49 (2006) 55–71 MR2198719

[19] **J Dubois**, **S Garoufalidis**, *Rationality of the* $\mathrm{SL}(2,\mathbb{C})$*–Reidemeister torsion in dimension* 3 arXiv:0908.1690

[20] **N M Dunfield**, *Cyclic surgery, degrees of maps of character curves, and volume rigidity for hyperbolic manifolds*, Invent. Math. 136 (1999) 623–657 MR1695208

[21] **N M Dunfield**, **S Friedl**, **N Jackson**, *Twisted Alexander polynomials of hyperbolic knots*, Exp. Math. 21 (2012) 329–352 MR3004250

[22] **N M Dunfield**, **S Garoufalidis**, *Incompressibility criteria for spun-normal surfaces*, Trans. Amer. Math. Soc. 364 (2012) 6109–6137 MR2946944

[23] **J L Dupont**, **C H Sah**, *Scissors congruences, II*, J. Pure Appl. Algebra 25 (1982) 159–195 MR662760

[24] **J L Dupont**, **C K Zickert**, *A dilogarithmic formula for the Cheeger–Chern–Simons class*, Geom. Topol. 10 (2006) 1347–1372 MR2255500

[25] **D B A Epstein**, **R C Penner**, *Euclidean decompositions of noncompact hyperbolic manifolds*, J. Differential Geom. 27 (1988) 67–80 MR918457

[26] **B Eynard**, *Invariants of spectral curves and intersection theory of moduli spaces of complex curves* arXiv:1110.2949

[27] **L D Faddeev**, *Discrete Heisenberg–Weyl group and modular group*, Lett. Math. Phys. 34 (1995) 249–254 MR1345554

[28] **L D Faddeev**, **R M Kashaev**, *Quantum dilogarithm*, Modern Phys. Lett. A 9 (1994) 427–434 MR1264393

[29] **L D Faddeev**, **R M Kashaev**, **A Y Volkov**, *Strongly coupled quantum discrete Liouville theory. I. Algebraic approach and duality*, Comm. Math. Phys. 219 (2001) 199–219 MR1828812

[30] **V V Fok**, **L O Chekhov**, *Quantum Teichmüller spaces*, Teoret. Mat. Fiz. 120 (1999) 511–528 MR1737362

[31] **S Garoufalides**, **D Zagier**, *Empirical relations betwen q–series and Kashaev's invariant of knots*, Preprint (2013)

[32] **S Garoufalidis**, *Quantum knot invariants*, Mathematische Arbeitstagung talk 2011 arXiv:1201.3314

[33] **S Garoufalidis**, *On the characteristic and deformation varieties of a knot*, from: "Proceedings of the Casson Fest", Geom. Topol. Monogr. 7 (2004) 291–309 MR2172488

[34] **S Garoufalidis**, *Chern–Simons theory, analytic continuation and arithmetic*, Acta Math. Vietnam. 33 (2008) 335–362 MR2501849

[35] **S Garoufalidis**, **C D Hodgson**, **J H Rubinstein**, **H Segerman**, *1–efficient triangulations and the index of a cusped hyperbolic 3–manifold* arXiv:1303.5278

[36] **S Garoufalidis**, **T T Q Lê**, *Asymptotics of the colored Jones function of a knot*, Geom. Topol. 15 (2011) 2135–2180 MR2860990

[37] **S Garoufalidis**, **D Zagier**, *Asymptotics of quantum knot invariants*, Preprint (2013)

[38] **S Goette**, **C K Zickert**, *The extended Bloch group and the Cheeger–Chern–Simons class*, Geom. Topol. 11 (2007) 1623–1635 MR2350461

[39] **S Gukov**, *Three-dimensional quantum gravity, Chern–Simons theory, and the A–polynomial*, Comm. Math. Phys. 255 (2005) 577–627 MR2134725

[40] **S Gukov**, **H Murakami**, *$\mathrm{SL}(2,\mathbb{C})$ Chern–Simons theory and the asymptotic behavior of the colored Jones polynomial*, Lett. Math. Phys. 86 (2008) 79–98 MR2465747

[41] **S Gukov**, **P Sułkowski**, *A–polynomial, B–model, and quantization*, J. High Energy Phys. (2012) 070, front matter+56 MR2996110

[42] **K Hikami**, *Hyperbolic structure arising from a knot invariant*, Internat. J. Modern Phys. A 16 (2001) 3309–3333 MR1848458

[43] **K Hikami**, *Generalized volume conjecture and the A–polynomials: the Neumann–Zagier potential function as a classical limit of the partition function*, J. Geom. Phys. 57 (2007) 1895–1940 MR2330673

[44] **C D Hodgson**, **J H Rubinstein**, **H Segerman**, *Triangulations of hyperbolic 3–manifolds admitting strict angle structures* arXiv:1111.3168

[45] **K Hori**, **S Katz**, **A Klemm**, **R Pandharipande**, **R Thomas**, **C Vafa**, **R Vakil**, **E Zaslow**, *Mirror symmetry*, Clay Mathematics Monographs 1, Amer. Math. Soc. (2003) MR2003030

[46] **V F R Jones**, *Hecke algebra representations of braid groups and link polynomials*, Ann. of Math. 126 (1987) 335–388 MR908150

[47] **E Kang**, *Normal surfaces in non-compact 3–manifolds*, J. Aust. Math. Soc. 78 (2005) 305–321 MR2142159

[48] **E Kang**, **J H Rubinstein**, *Ideal triangulations of 3–manifolds, I: Spun normal surface theory*, from: "Proceedings of the Casson Fest", (C Gordon, Y Rieck, editors), Geom. Topol. Monogr. 7 (2004) 235–265 MR2172486

[49] **R M Kashaev**, *Quantum dilogarithm as a 6j–symbol*, Modern Phys. Lett. A 9 (1994) 3757–3768 MR1317945

[50] **R M Kashaev**, *A link invariant from quantum dilogarithm*, Modern Phys. Lett. A 10 (1995) 1409–1418 MR1341338

[51] **R M Kashaev**, *Quantization of Teichmüller spaces and the quantum dilogarithm*, Lett. Math. Phys. 43 (1998) 105–115 MR1607296

[52] **S V Matveev**, *Transformations of special spines, and the Zeeman conjecture*, Izv. Akad. Nauk SSSR Ser. Mat. 51 (1987) 1104–1116, 1119 MR925096

[53] **S Matveev**, *Algorithmic topology and classification of 3–manifolds*, 2nd edition, Algorithms and Computation in Mathematics 9, Springer, Berlin (2007) MR2341532

[54] **W Müller**, *Analytic torsion and R–torsion for unimodular representations*, J. Amer. Math. Soc. 6 (1993) 721–753 MR1189689

[55] **H Murakami**, **J Murakami**, *The colored Jones polynomials and the simplicial volume of a knot*, Acta Math. 186 (2001) 85–104 MR1828373

[56] **W D Neumann**, *Combinatorics of triangulations and the Chern–Simons invariant for hyperbolic 3–manifolds*, from: "Topology '90", (B Apanasov, W D Neumann, A W Reid, L Siebenmann, editors), Ohio State Univ. Math. Res. Inst. Publ. 1, de Gruyter, Berlin (1992) 243–271 MR1184415

[57] **W D Neumann**, *Extended Bloch group and the Cheeger–Chern–Simons class*, Geom. Topol. 8 (2004) 413–474 MR2033484

[58] **W D Neumann**, **D Zagier**, *Volumes of hyperbolic three-manifolds*, Topology 24 (1985) 307–332 MR815482

[59] **R Piergallini**, *Standard moves for standard polyhedra and spines*, Rend. Circ. Mat. Palermo Suppl. (1988) 391–414 MR958750

[60] **M Polyak**, *Feynman diagrams for pedestrians and mathematicians*, from: "Graphs and patterns in mathematics and theoretical physics", (M Lyubich, L Takhtajan, editors), Proc. Sympos. Pure Math. 73, Amer. Math. Soc. (2005) 15–42 MR2131010

[61] **B Ponsot**, **J Teschner**, *Clebsch-Gordan and Racah–Wigner coefficients for a continuous series of representations of $\mathcal{U}_q(\mathrm{sl}(2,\mathbb{R}))$*, Comm. Math. Phys. 224 (2001) 613–655 MR1871903

[62] **J Porti**, *Torsion de Reidemeister pour les variétés hyperboliques*, Mem. Amer. Math. Soc. 128 (1997) x+139 MR1396960

[63] **D B Ray**, **I M Singer**, *R–torsion and the Laplacian on Riemannian manifolds*, Advances in Math. 7 (1971) 145–210 MR0295381

[64] **D Shale**, *Linear symmetries of free boson fields*, Trans. Amer. Math. Soc. 103 (1962) 149–167 MR0137504

[65] **V P Spiridonov**, **G S Vartanov**, *Elliptic hypergeometry of supersymmetric dualities II: orthogonal groups, knots, and vortices* `arXiv:1107.5788`

[66] **W P Thurston**, *The geometry and topology of three-manifolds*, Princeton Univ. Math. Dept. Lecture Notes (1979) Available at `http://msri.org/publications/books/gt3m/`

[67] **W P Thurston**, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. 6 (1982) 357–381 MR648524

[68] **S Tillmann**, *Normal surfaces in topologically finite 3–manifolds*, Enseign. Math. 54 (2008) 329–380 MR2478091

[69] **S Tillmann**, *Degenerations of ideal hyperbolic triangulations*, Math. Z. 272 (2012) 793–823 MR2995140

[70] **V G Turaev**, *The Yang–Baxter equation and invariants of links*, Invent. Math. 92 (1988) 527–553 MR939474

[71] **A Weil**, *Sur certains groupes d'opérateurs unitaires*, Acta Math. 111 (1964) 143–211 MR0165033

[72] **E Witten**, *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. 121 (1989) 351–399 MR990772

[73] **E Witten**, *Quantization of Chern–Simons gauge theory with complex gauge group*, Comm. Math. Phys. 137 (1991) 29–66 MR1099255

[74] **Y Yamaguchi**, *A relationship between the non-acyclic Reidemeister torsion and a zero of the acyclic Reidemeister torsion*, Ann. Inst. Fourier (Grenoble) 58 (2008) 337–362 MR2401224

[75] **C K Zickert**, *The volume and Chern–Simons invariant of a representation*, Duke Math. J. 150 (2009) 489–532 MR2582103

School of Natural Sciences, Institute for Advanced Study
Einstein Drive, Princeton, NJ 08540, USA

School of Mathematics, Georgia Institute of Technology
686 Cherry Street, Atlanta, GA 30332-0160, USA

`tdd@ias.edu`, `stavros@math.gatech.edu`

`http://www.math.gatech.edu/~stavros`

Proposed: Walter Neumann Received: 6 March 2012
Seconded: Danny Calegari, Jean-Pierre Otal Revised: 25 October 2012

THE COMPLEX VOLUME OF $\mathrm{SL}(n,\mathbb{C})$-REPRESENTATIONS OF 3-MANIFOLDS

STAVROS GAROUFALIDIS, DYLAN P. THURSTON,
and CHRISTIAN K. ZICKERT

Abstract

For a compact 3*-manifold M with arbitrary (possibly empty) boundary, we give a parameterization of the set of conjugacy classes of boundary-unipotent representations of $\pi_1(M)$ into* $\mathrm{SL}(n,\mathbb{C})$*. Our parameterization uses Ptolemy coordinates, which are inspired by coordinates on higher Teichmüller spaces due to Fock and Goncharov. We show that a boundary-unipotent representation determines an element in Neumann's extended Bloch group $\widehat{\mathcal{B}}(\mathbb{C})$, and we use this to obtain an efficient formula for the Cheeger–Chern–Simons invariant, and, in particular, for the volume. Computations for the census manifolds show that boundary-unipotent representations are abundant, and numerical comparisons with census volumes suggest that the volume of a representation is an integral linear combination of volumes of hyperbolic* 3*-manifolds. This is in agreement with a conjecture of Walter Neumann, stating that the Bloch group is generated by hyperbolic manifolds.*

Contents

DUKE MATHEMATICAL JOURNAL
Vol. 164, No. 11, © 2015 DOI 10.1215/00127094-3121185
Received 10 September 2013. Revision received 4 September 2014.
2010 *Mathematics Subject Classification*. Primary 57N10, 57M27, 57M50, 58J28; Secondary 11R70, 19F27, 11G55.

1. Introduction

For a closed 3-manifold M, the Cheeger–Chern–Simons invariant (see [5], [6]) of a representation ρ of $\pi_1(M)$ in $\mathrm{SL}(n,\mathbb{C})$ is given by the Chern–Simons integral

$$\widehat{c}(\rho)=\frac{1}{2}\int_M s^*\Big(\mathrm{Tr}\Big(A\wedge dA+\frac{2}{3}A\wedge A\wedge A\Big)\Big)\in \mathbb{C}/4\pi^2\mathbb{Z}, \tag{1.1}$$

where A is the flat connection in the flat $\mathrm{SL}(n,\mathbb{C})$-bundle E_ρ with holonomy ρ, and $s: M\to E_\rho$ is a section of E_ρ. Since $\mathrm{SL}(n,\mathbb{C})$ is 2-connected, a section always exists, and a different choice of section changes the value of the integral by a multiple of $4\pi^2$.

When $n=2$, the imaginary part of the Cheeger–Chern–Simons invariant equals the hyperbolic volume of ρ. More precisely, if $D:\widetilde{M}\to\mathbb{H}^3$ is a developing map for ρ and $\nu_{\mathbb{H}^3}$ is the hyperbolic volume form, $\mathrm{Im}(\widehat{c}(\rho))$ equals the integral of $D^*(\nu_\rho)$ over a fundamental domain for M. In particular, if $M=\mathbb{H}^3/\Gamma$ is a hyperbolic manifold, and ρ is a lift to $\mathrm{SL}(2,\mathbb{C})$ of the geometric representation $\rho_{\mathrm{geo}}:\pi_1(M)\to\mathrm{PSL}(2,\mathbb{C})$, then the imaginary part equals the volume of M. In fact, in this case we have

$$\widehat{c}(\rho)=i\big(\mathrm{Vol}(M)+i\,\mathrm{CS}(M)\big), \tag{1.2}$$

where $\mathrm{CS}(M)$ is the Chern–Simons invariant of M (with the Riemannian connection). The invariant $\mathrm{Vol}(M)+i\,\mathrm{CS}(M)$ is often referred to as *complex volume*. Motivated by this, we define the complex volume $\mathrm{Vol}_{\mathbb{C}}$ of a representation $\rho:\pi_1(M)\to\mathrm{SL}(n,\mathbb{C})$ by

$$\widehat{c}(\rho)=i\,\mathrm{Vol}_{\mathbb{C}}(\rho), \tag{1.3}$$

and define the *volume* of ρ to be the real part of the complex volume, that is, the imaginary part of the Cheeger–Chern–Simons invariant. Surprisingly, as we shall see, the relationship to hyperbolic volume seems to persist even when $n>2$.

The set of $\mathrm{SL}(n,\mathbb{C})$-representations is a complex variety with finitely many components, and the complex volume is constant on components. This follows from the fact that representations in the same component have cohomologous Chern–Simons forms. Hence, for any M, the set of complex volumes is a finite set.

We show that the definition of the Cheeger–Chern–Simons invariant naturally extends to compact manifolds with boundary, and representations $\rho:\pi_1(M)\to\mathrm{SL}(n,\mathbb{C})$ that are *boundary-unipotent*, that is, take peripheral subgroups to a conjugate of the unipotent group N of upper triangular matrices with 1's on the diagonal. We formulate all our results in this more general setup.

The main result of the paper is a concrete algorithm for computing the set of complex volumes. The idea is that the set of (conjugacy classes of) boundary-unipotent representations can be parameterized by a variety, called the *Ptolemy variety*, which is defined by homogeneous polynomials of degree 2. The Ptolemy variety depends on a choice of triangulation, but if the triangulation is sufficiently fine, then every representation is detected by the Ptolemy variety. We show that a point c in the Ptolemy variety naturally determines an element $\lambda(c)$ in Neumann's extended Bloch group $\widehat{\mathcal{B}}(\mathbb{C})$, such that if ρ is the representation corresponding to c, we have

$$R\big(\lambda(c)\big) = i\,\mathrm{Vol}_{\mathbb{C}}(\rho), \tag{1.4}$$

where $R\colon \widehat{\mathcal{B}}(\mathbb{C}) \to \mathbb{C}/4\pi^2\mathbb{Z}$ is a Rogers dilogarithm.

There is a canonical group homomorphism

$$\phi_n\colon\ \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(n,\mathbb{C}) \tag{1.5}$$

defined by taking a matrix A to its $(n-1)$th symmetric power (see Section 11). The map ϕ_n preserves unipotent elements, and we show that composing a boundary-unipotent representation in $\mathrm{SL}(2,\mathbb{C})$ with ϕ_n multiplies the complex volume by $\binom{n+1}{3}$. If $M = \mathbb{H}^3/\Gamma$ is a hyperbolic 3-manifold, then the geometric representation ρ_{geo} always lifts to a representation in $\mathrm{SL}(2,\mathbb{C})$; but if M has cusps, then lifts are not necessarily boundary-unipotent. In fact, by a result of Calegari [4], if M has a single cusp, then any lift of the geometric representation takes a longitude to an element with trace -2. When n is even, we shall thus, more generally, be interested in boundary-unipotent representations in

$$p\,\mathrm{SL}(n,\mathbb{C}) = \mathrm{SL}(n,\mathbb{C})/\langle \pm I\rangle. \tag{1.6}$$

Such representations have a complex volume defined modulo $\pi^2 i$, and our algorithm computes these as well. By studying representations in $p\,\mathrm{SL}(n,\mathbb{C})$, we make sure that when M is hyperbolic, there is always at least one representation with nontrivial complex volume, namely, $\phi_n \circ \rho_{\mathrm{geo}}$.

Walter Neumann has conjectured that every element in the Bloch group $\mathcal{B}(\mathbb{C})$ is an integral linear combination of Bloch group elements of hyperbolic 3-manifolds. Since the extended Bloch group equals the Bloch group up to torsion, Neumann's conjecture would imply that all complex volumes are, up to rational multiples of $i\pi^2$, integral linear combinations of complex volumes of hyperbolic 3-manifolds. In particular, the volumes should all be integral linear combinations of volumes of hyperbolic manifolds.

Our algorithm has been implemented by Matthias Goerner. The algorithm uses Magma [3] to compute a primary decomposition of the Ptolemy variety and then uses (1.4) to compute the complex volumes. For $n = 2$, we have computed primary decompositions of the Ptolemy varieties for all census manifolds with at most 8 simplices (these usually finish within a fraction of a second) and all link complements with at most 16 simplices in the SnapPy census [7] of knots with up to 11 crossings and links with up to 10 crossings. When there are more than 16 simplices, some of the computations do not terminate. For $n = 3$, computations are feasible for many manifolds with up to four simplices, but, for $n = 4$, the computations run out of memory for all manifolds with more than two simplices. It would be interesting to perform numerical calculations for $n \geq 4$. Our computations have revealed numerous (numerical) examples of linear combinations as predicted by Neumann's conjecture. To the best of our knowledge, our examples are the first concrete computations (the first of which were carried out in 2009) of the Cheeger–Chern–Simons invariant (complex volume) for $n > 2$.

1.1. Statement of our results

This section gives a brief summary of our main results. More details can be found in the paper.

1.1.1. The Ptolemy variety

Let M be a compact, oriented 3-manifold with (possibly empty) boundary, and let K be a closed 3-cycle (triangulated complex; see Definition 4.1) homeomorphic to the space obtained from M by collapsing each boundary component to a point. We identify each of the simplices of K with a standard simplex:

$$\Delta_n^3 = \left\{(x_0, x_1, x_2, x_3) \in \mathbb{R}^4 \;\middle|\; 0 \leq x_i \leq n, x_0 + x_1 + x_2 + x_3 = n\right\}. \tag{1.7}$$

Let $\Delta_n^3(\mathbb{Z})$ be the set of points in Δ_n^3 with integral coordinates, and let $\dot{\Delta}_n^3(\mathbb{Z})$ be $\Delta_n^3(\mathbb{Z})$ with the four vertex points removed.

Definition 1.1

A *Ptolemy assignment* on Δ_n^3 is an assignment $\dot{\Delta}_n^3(\mathbb{Z}) \to \mathbb{C}^*$, $t \mapsto c_t$, of a nonzero complex number c_t to each (nonvertex) integral point t of Δ_n^3 such that for each $\alpha \in \Delta_{n-2}^3(\mathbb{Z})$, the *Ptolemy relation*

$$c_{\alpha_{03}} c_{\alpha_{12}} + c_{\alpha_{01}} c_{\alpha_{23}} = c_{\alpha_{02}} c_{\alpha_{13}} \tag{1.8}$$

is satisfied (see Figure 2). Here, α_{ij} denotes the integral point $\alpha + e_i + e_j$. A Ptolemy assignment on K is a Ptolemy assignment c^i on each simplex Δ_i of K such that the Ptolemy coordinates agree on identified faces.

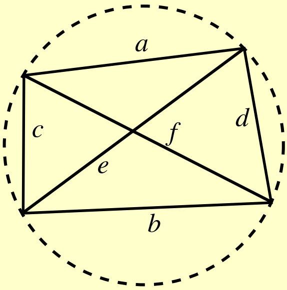


Figure 1. A quadrilateral is inscribed in a circle if and only if $ab+cd=ef$.

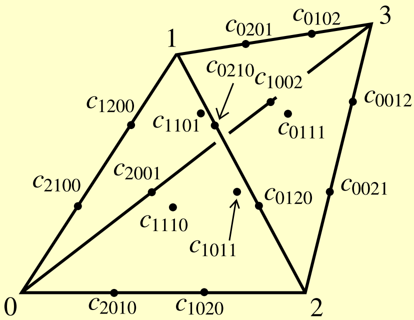


Figure 2. Ptolemy assignment for $n=3$. The Ptolemy relation for $\alpha=1000$ is $c_{2001}c_{1110}+c_{2100}c_{1011}=c_{2010}c_{1101}$.

Remark 1.2

The name is inspired by the resemblance of (1.8) with the Ptolemy relation between the lengths of the sides and diagonals of an inscribed quadrilateral (see Figure 1). In the work of Fock and Goncharov [13], the Ptolemy relations appear as relations between coordinates on the higher Teichmüller space when the triangulation of a surface is changed by a flip.

It follows immediately from the definition that the set of Ptolemy assignments on K is an algebraic set $P_n(K)$, which we shall refer to as the *Ptolemy variety*.

The extended pre-Bloch group $\widehat{\mathcal{P}}(\mathbb{C})$ is generated by tuples $(u,v)\in\mathbb{C}^2$ with $e^u+e^v=1$, and the extended Bloch group $\widehat{\mathcal{B}}(\mathbb{C})\subset\widehat{\mathcal{P}}(\mathbb{C})$ is the kernel of the map $\widehat{\mathcal{P}}(\mathbb{C})\to\wedge^2(\mathbb{C})$ taking (u,v) to $u\wedge v$. We refer to Section 3 for a review. Using (1.8), we obtain that a Ptolemy assignment c on Δ_n^3 gives rise to an element

$$\lambda(c) = \sum_{\alpha\in\Delta^3(n-2)} (\widetilde{c}_{\alpha_{03}} + \widetilde{c}_{\alpha_{12}} - \widetilde{c}_{\alpha_{02}} - \widetilde{c}_{\alpha_{13}}, \widetilde{c}_{\alpha_{01}} + \widetilde{c}_{\alpha_{23}} - \widetilde{c}_{\alpha_{02}} - \widetilde{c}_{\alpha_{13}})$$
$$\in \widehat{\mathcal{P}}(\mathbb{C}), \tag{1.9}$$

where the tilde denotes a branch of logarithm (the particular choice is inessential). We thus have a map

$$\lambda : P_n(K) \to \widehat{\mathcal{P}}(\mathbb{C}), \qquad c \mapsto \sum_i \epsilon_i \lambda(c^i), \tag{1.10}$$

where the sum is over the simplices of K. Let $R_{\mathrm{SL}(n,\mathbb{C}),N}(M)$ denote the set of conjugacy classes of boundary-unipotent representations $\pi_1(M) \to \mathrm{SL}(n,\mathbb{C})$. The following theorem (as well as Theorem 1.12 below) gives an efficient algorithm for computing complex volumes. See Section 10 for examples.

THEOREM 1.3 (Proof in Section 9.5)
A Ptolemy assignment c uniquely determines a boundary-unipotent representation $\mathcal{R}(c) \in R_{\mathrm{SL}(n,\mathbb{C}),N}(M)$. The map λ has image in $\widehat{\mathcal{B}}(\mathbb{C})$, and we have a commutative diagram

$$\begin{tikzcd} P_n(K) \arrow[r, "\lambda"] \arrow[d, "\mathcal{R}"] & \widehat{\mathcal{B}}(\mathbb{C}) \arrow[d, "R"] \\ R_{\mathrm{SL}(n,\mathbb{C}),N}(M) \arrow[r, "i\,\mathrm{Vol}_{\mathbb{C}}"] & \mathbb{C}/4\pi^2\mathbb{Z} \end{tikzcd} \tag{1.11}$$

Moreover, if the triangulation is sufficiently fine (a single barycentric subdivision suffices), then the map $\mathcal{R}$ is surjective.

Remark 1.4
We show in Section 9 that there is a one-to-one correspondence between points in $P_n(K)$ and generically decorated (see Section 5) boundary-unipotent $\mathrm{SL}(n,\mathbb{C})$-representations. Under this correspondence, the map $\mathcal{R}$ is just the forgetful map ignoring the decoration. Note that $P_n(K)$ depends on the triangulation and may be empty.

Let $H \subset \mathrm{SL}(n,\mathbb{C})$ denote the group of diagonal matrices, and let h denote the number of boundary components of M. In Section 4.1, we define an action of H^h on $P_n(K)$. We denote the quotient by $P_n(K)_{\mathrm{red}}$. The action only changes the decoration, and so $\mathcal{R}$ factors through $P_n(K)_{\mathrm{red}}$.

Definition 1.5
A boundary-unipotent representation $\rho\colon \pi_1(M) \to \mathrm{SL}(n,\mathbb{C})$ is *peripherally well behaved* if the image of each peripheral subgroup is either trivial or contains an element with a maximal Jordan block. If the latter condition holds for each peripheral subgroup, then we say that ρ is *peripherally nondegenerate*.

Remark 1.6
When $n = 2$, all representations are peripherally well behaved.

THEOREM 1.7 (Proof in Section 9.5)
The image of $\mathcal{R}\colon\ P_n(K)_{\mathrm{red}} \to R_{\mathrm{SL}(n,\mathbb{C}),N}(M)$ consists of the set of representations admitting a generic decoration (see Definition 5.2). If such a representation ρ is peripherally nondegenerate, then the preimage in $P_n(K)_{\mathrm{red}}$ is a single point; that is, any two decorations of ρ differ by the diagonal action. If ρ is peripherally well behaved, then any two preimages of $\mathcal{R}$ have the same image in $\widehat{\mathcal{B}}(\mathbb{C})$.

COROLLARY 1.8
A peripherally well-behaved boundary-unipotent representation ρ in $\mathrm{SL}(n,\mathbb{C})$ determines an element $[\rho] \in \widehat{\mathcal{B}}(\mathbb{C})$ such that $R([\rho]) = i\,\mathrm{Vol}_{\mathbb{C}}(\rho)$.

Remark 1.9
In general, the preimage of a representation under $\mathcal{R}$ can have large dimension.

1.1.2. Hyperbolic manifolds and $p\,\mathrm{SL}(n,\mathbb{C})$-representations
Let $\phi_n\colon\ \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(n,\mathbb{C})$ denote the canonical irreducible representation. Note that when n is odd, ϕ_n factors through $\mathrm{PSL}(2,\mathbb{C})$. If a representation ρ is in the image of $P_n(K) \to R_{\mathrm{SL}(n,\mathbb{C}),N}(M)$, then we say that $P_n(K)$ *detects* ρ.

THEOREM 1.10 (Proof in Section 11.1)
Suppose that $M = \mathbb{H}^3/\Gamma$ is an oriented, hyperbolic manifold with finite volume and geometric representation $\rho_{\mathrm{geo}}\colon \pi_1(M) \to \mathrm{PSL}(2,\mathbb{C})$. If the triangulation of K has no nonessential edges, and if n is odd, then $P_n(K)$ is nonempty and detects $\phi_n \circ \rho_{\mathrm{geo}}$.

When n is even, $\phi_n \circ \rho_{\mathrm{geo}}$ is only a representation in $p\,\mathrm{SL}(n,\mathbb{C}) = \mathrm{SL}(n,\mathbb{C})/\langle \pm I\rangle$.

Definition 1.11
Let $\sigma \in Z^2(\Delta_n^3;\mathbb{Z}/2\mathbb{Z})$ be a cocycle. A $p\,\mathrm{SL}(n,\mathbb{C})$-*Ptolemy assignment* on Δ_n^3 with *obstruction cocycle* σ is an assignment of Ptolemy coordinates to the integral points

of Δ_n^3 such that

$$\sigma_2\sigma_3 c_{\alpha_{03}} c_{\alpha_{12}} + \sigma_0\sigma_3 c_{\alpha_{01}} c_{\alpha_{23}} = c_{\alpha_{02}} c_{\alpha_{13}}. \tag{1.12}$$

Here, $\sigma_i \in \mathbb{Z}/2\mathbb{Z} = \langle \pm 1 \rangle$ is the value of σ on the face opposite the ith vertex of Δ_n^3. A $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignment on K with obstruction cocycle $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$ is a collection of $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignments c^i on Δ_i with obstruction class σ_{Δ_i} such that the Ptolemy coordinates agree on common faces.

The set of $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignments on K with obstruction cocycle σ is an algebraic set $P_n^\sigma(K)$, which, up to canonical isomorphism, only depends on the cohomology class of σ. The obstruction class to lifting a boundary-unipotent representation in $p\,\mathrm{SL}(n,\mathbb{C})$ to a boundary-unipotent representation in $\mathrm{SL}(n,\mathbb{C})$ is a class in $H^2(M,\partial M;\mathbb{Z}/2\mathbb{Z}) = H^2(K;\mathbb{Z}/2\mathbb{Z})$. For $\sigma \in H^2(K;\mathbb{Z}/2\mathbb{Z})$, let $R^\sigma_{p\,\mathrm{SL}(n,\mathbb{C}),N}(M)$ denote the set of (conjugacy classes of) boundary-unipotent representations in $p\,\mathrm{SL}(n,\mathbb{C})$ with obstruction class σ. If M is hyperbolic, we let $\sigma_{\mathrm{geo}} \in H^2(K;\mathbb{Z}/2\mathbb{Z})$ denote the obstruction class of the geometric representation.

THEOREM 1.12 (Proof in Section 9.5)
Let n be even. For each $\sigma \in H^2(K;\mathbb{Z}/2\mathbb{Z})$, we have a commutative diagram ($\widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}$ is defined in Section 3.2)

$$\begin{tikzcd} P_n^\sigma(K) \arrow[r, "\lambda"] \arrow[d, "\mathcal{R}"] & \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}} \arrow[d, "R"] \\ R^\sigma_{p\,\mathrm{SL}(n,\mathbb{C}),N}(M) \arrow[r, "i\,\mathrm{Vol}_{\mathbb{C}}"] & \mathbb{C}/\pi^2\mathbb{Z} \end{tikzcd} \tag{1.13}$$

If the triangulation of K is sufficiently fine, then $\mathcal{R}$ is surjective. If $M = \mathbb{H}^3/\Gamma$ is hyperbolic, and if K has no nonessential edges, then $P_n^{\sigma_{\mathrm{geo}}}(K)$ detects $\phi_n \circ \rho_{\mathrm{geo}}$.

Remark 1.13
The analogue of Theorem 1.7 also holds, except that the preimage of a peripherally well-behaved representation is now parameterized by $Z^1(K;\mathbb{Z}/2\mathbb{Z})$ (see Section 9.4).

Remark 1.14
If the triangulation has a nonessential edge, all Ptolemy varieties are empty. Hence, if $P_2^\sigma(K)$ is nonempty for some σ, and if M is hyperbolic, then the Ptolemy variety $P^{\sigma_{\mathrm{geo}}}(K)$ will detect the geometric representation.

THEOREM 1.15 (Proof in Section 11)
Let ρ be a peripherally well-behaved representation in $\mathrm{SL}(2,\mathbb{C})$ *or* $\mathrm{PSL}(2,\mathbb{C})$. *The extended Bloch group element of $\phi_n \circ \rho$ is $\binom{n+1}{3}$ times that of ρ. In particular, composition with ϕ_n multiplies complex volume by $\binom{n+1}{3}$.*

1.1.3. The Cheeger–Chern–Simons class

The Cheeger–Chern–Simons invariant can be viewed as a characteristic class $H_3(\mathrm{SL}(n,\mathbb{C})) \to \mathbb{C}/4\pi^2\mathbb{Z}$, and the result underlying the proof of commutativity of (1.11) is Theorem 1.16 below, giving an explicit cocycle formula for the Cheeger–Chern–Simons class. The formula generalizes the formula in [16] for $n = 2$. Recall that a homology class can be represented by a formal sum of tuples $(g_0, \dots, g_3)$. To such a tuple, we can assign a Ptolemy assignment $c(g_0, \dots, g_3)$ defined by

$$c(g_0, \dots, g_3)_t = \det\bigl(\{g_0\}_{t_0} \cup \dots \cup \{g_3\}_{t_3}\bigr), \quad t = (t_0, \dots, t_3), \tag{1.14}$$

where $\{g_i\}_{t_i}$ denotes the ordered set consisting of the first t_i column vectors of g_i. One can always represent a homology class by tuples, such that all the determinants in (1.14) are nonzero.

THEOREM 1.16 (Proof in Section 8)
The Cheeger–Chern–Simons class $\widehat{c}$ factors as

$$H_3\bigl(\mathrm{SL}(n,\mathbb{C})\bigr) \xrightarrow{\lambda} \widehat{\mathcal{B}}(\mathbb{C}) \xrightarrow{R} \mathbb{C}/4\pi^2\mathbb{Z}, \tag{1.15}$$

where λ is induced by the map taking a tuple $(g_0, \dots, g_3)$ to $\lambda(c(g_0, \dots, g_3)) \in \widehat{\mathcal{P}}(\mathbb{C})$.

1.1.4. Thurston's gluing equations

When $n = 2$, Thurston's gluing equation variety $V(K)$ is another variety, which is often used to compute volume. It is given by an equation for each edge of K and an equation for each generator of the fundamental groups of the boundary components of M (see Section 12).

THEOREM 1.17 (Proof in Section 12)
Suppose that M has h boundary components. There is a surjective regular map

$$\coprod_{\sigma \in H^2(K;\mathbb{Z}/2\mathbb{Z})} P_2^\sigma(K) \to V(K) \tag{1.16}$$

with fibers that are disjoint copies of $(\mathbb{C}^*)^h$.

Remark 1.18
The Ptolemy variety seems to offer significant computational advantage over the gluing equations, but according to Fabrice Rouillier (private communications) one can manipulate the gluing equations to mitigate this.

1.2. Neumann's conjecture
The fact that (1.10) has an image in $\widehat{\mathcal{B}}(\mathbb{C})$ as opposed to $\widehat{\mathcal{P}}(\mathbb{C})$ has very interesting conjectural consequences. It is well known (see, e.g., [25]) that the Bloch group $\mathcal{B}(\mathbb{C})$ is a $\mathbb{Q}$-vector space, and Walter Neumann has conjectured that it is generated by Bloch invariants of hyperbolic manifolds. More generally, Neumann has proposed the following stronger conjecture (see [20, Question 2.8]).

CONJECTURE 1.19
Let $F \subset \mathbb{C}$ be a concrete number field which is not in $\mathbb{R}$. The Bloch group $\mathcal{B}(F)$ is generated (integrally) modulo torsion by hyperbolic manifolds with an invariant trace field contained in F.

Using Theorems 1.3 and 1.12, Conjecture 1.19 implies the following (see Section 10, for example).

CONJECTURE 1.20
Let ρ be a boundary-unipotent representation of $\pi_1(M)$ in $\mathrm{SL}(n,\mathbb{C})$ or $p\,\mathrm{SL}(n,\mathbb{C})$. There exist hyperbolic 3-manifolds $M_1,\dots,M_k$ and integers $r_1,\dots,r_k$ such that

$$\mathrm{Vol}_{\mathbb{C}}(\rho) = \sum r_i \,\mathrm{Vol}_{\mathbb{C}}(M_i) \in \mathbb{C}/i\pi^2\mathbb{Q}. \tag{1.17}$$

In particular, $\mathrm{Vol}(\rho) = \sum r_i \,\mathrm{Vol}(M_i) \in \mathbb{R}$.

Remark 1.21
The Ptolemy coordinates may be considered as a 3-dimensional analogue of Fock and Goncharov's $\mathcal{A}$-coordinates (see [13]). They were defined for 3-manifolds in [29] (under the name *ideal cochain*) and have subsequently been studied by several other authors. These include Bergeron, Falbel, and Guilloux [2]; Garoufalidis, Goerner, and Zickert [14]; and Dimofte, Gabella, and Goncharov [8]. Shape coordinates for $\mathrm{PGL}(3,\mathbb{C})$-representations have also been used by Falbel [11] and Falbel–Wang [12] in connection with spherical CR-structures.

1.3. Overview of the paper
Section 2 reviews the Cheeger–Chern–Simons classes for flat bundles. Section 3 gives a brief review of the two variants of the extended Bloch group, and Section 4 reviews

the theory, introduced in Zickert [27], of decorated representations and relative fundamental classes. In Section 5, we introduce the notion of *generic* decorations and define the Ptolemy variety $P_n(K)$. In Section 6, we construct a chain complex of Ptolemy assignments and use it to construct a map from $H_3(\mathrm{SL}(n,\mathbb{C}),N)$ to $\widehat{\mathcal{B}}(\mathbb{C})$ commuting with stabilization. This shows that a decorated boundary-unipotent representation determines an element in the extended Bloch group, which is given explicitly in terms of the Ptolemy coordinates. In Section 7, we show that the extended Bloch group element of a decorated, peripherally well-behaved representation is independent of the decoration, and, in Section 8, we show that the Cheeger–Chern–Simons class is given as in Theorem 1.16. In Section 9, we show that the Ptolemy variety parameterizes generically decorated representations, and we give an explicit formula for recovering a representation from its Ptolemy coordinates. In Section 10, we give some examples of computations, and we list some interesting findings. Section 11 discusses the irreducible representations of SL$(2,\mathbb{C})$, and Section 12 discusses the relationship to Thurston's gluing equations when $n=2$.

2. The Cheeger–Chern–Simons classes

The Cheeger–Chern–Simons classes (see [5], [6]) are characteristic classes of principal bundles with connection. For general bundles, the characteristic classes are differential characters (see [5]), but for flat bundles they reduce to ordinary (singular) cohomology classes. In this paper we will focus exclusively on flat bundles. Let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$, and let Λ be a proper subring of $\mathbb{F}$. Let G be a Lie group over $\mathbb{F}$ with finitely many components. There is a characteristic class $S_{P,u}$ for each pair (P,u) consisting of an invariant polynomial $P\in I^k(G;\mathbb{F})$ and a class $u\in H^{2k}(BG;\Lambda)$, whose image in $H^{2k}(BG;\mathbb{F})$ equals $W(P)$, where W is the Chern–Weil homomorphism

$$W: I^k(G;\mathbb{F})\to H^{2k}(BG;\mathbb{F}). \tag{2.1}$$

The characteristic class $S_{P,u}$ associates to each flat G-bundle $E\to M$ a cohomology class $S_{P,u}(E)\in H^{2k-1}(M;\mathbb{F}/\Lambda)$.

2.1. *Simply connected, simple Lie groups*

If G is simply connected and simple, $H^1(G;\mathbb{Z})$ and $H^2(G;\mathbb{Z})$ are trivial, and $H^3(G;\mathbb{Z})\cong\mathbb{Z}$. Hence, by the Serre spectral sequence for the universal bundle, we have an isomorphism

$$S: H^4(BG;\mathbb{Z})\cong H^3(G;\mathbb{Z})\cong\mathbb{Z} \tag{2.2}$$

called the *suspension*. The Killing form on G defines an invariant polynomial $B\in I^2(G;\mathbb{F})$, and since B is real on the maximal compact subgroup K of G, $W(B)$ is a real class. Hence, there exists a unique positive real number α such that $W(\alpha B)$ is a generator of $H^4(BG; 4\pi^2\mathbb{Z})$.

Definition 2.1
The *Cheeger–Chern–Simons class* for G is the characteristic class of flat G bundles defined by $S_{\alpha B, W(\alpha B)}$. We denote it by $\widehat{c}$.

2.2. Complex groups and volume
Recall that there is a one-to-one correspondence between flat G-bundles over M and representations $\pi_1(M) \to G$ up to conjugation. This correspondence takes a flat bundle to its holonomy representation. If $\rho: \pi_1(M) \to G$ is a representation, then we let E_ρ denote the corresponding flat bundle. In the following, G denotes a simply connected, simple, complex Lie group, and M denotes a *closed*, oriented 3-manifold. The following definition is motivated by (1.2).

Definition 2.2
The *complex volume* $\mathrm{Vol}_{\mathbb{C}}(\rho)$ of a representation $\rho: \pi_1(M) \to G$ is defined by

$$\widehat{c}(E_\rho)\big([M]\big) = i \,\mathrm{Vol}_{\mathbb{C}}(\rho) \in \mathbb{C}/4\pi^2\mathbb{Z}. \tag{2.3}$$

The *volume* $\mathrm{Vol}(\rho)$ of ρ is the real part of $\mathrm{Vol}_{\mathbb{C}}(\rho)$.

2.3. The universal classes and group cohomology
The Cheeger–Chern–Simons classes are also defined for the universal flat bundle $EG^\delta \to BG^\delta$. For an explicit construction, we refer to [10] or [9]. In particular, we have a class $\widehat{c} \in H^3(BG^\delta; \mathbb{C}/4\pi^2\mathbb{Z})$. If $\rho: \pi_1(M) \to G$ is a representation, with classifying map $B\rho: M \to BG^\delta$, then we thus have

$$\widehat{c}\big(B\rho_*\big([M]\big)\big) = i \,\mathrm{Vol}_{\mathbb{C}}(\rho). \tag{2.4}$$

It is well known that the homology of BG^δ is the homology of the chain complex $C_* \otimes_{\mathbb{Z}[G]} \mathbb{Z}$, where C_* is any free $\mathbb{Z}[G]$-resolution of $\mathbb{Z}$. A convenient choice of free resolution is the complex C_*, generated in degree n by tuples $(g_0, \dots, g_n)$, and with boundary map given by

$$\partial(g_0, \dots, g_n) = \sum (-1)^i (g_0, \dots, \widehat{g}_i, \dots, g_n). \tag{2.5}$$

The homology of $C_* \otimes_{\mathbb{Z}[G]} \mathbb{Z}$ is denoted by $H_*(G)$, and so $H_*(G) = H_*(BG^\delta)$. Theorem 1.16 gives a concrete cocycle formula for $\widehat{c}: H_3(\mathrm{SL}(n, \mathbb{C})) \to \mathbb{C}/4\pi^2\mathbb{Z}$.

2.4. *Compact manifolds with boundary*

In Section 6.1 below, we construct a natural extension of $\widehat{c}\colon H_3(\mathrm{SL}(n,\mathbb{C}))\to\mathbb{C}/4\pi^2\mathbb{Z}$ to a homomorphism

$$\widehat{c}\colon H_3\big(\mathrm{SL}(n,\mathbb{C}),N\big)\to\mathbb{C}/4\pi^2\mathbb{Z}, \tag{2.6}$$

where N is the subgroup of upper triangular matrices with 1's on the diagonal.

Definition 2.3

Let $\rho:\pi_1(M)\to\mathrm{SL}(n,\mathbb{C})$ be a boundary-unipotent representation. The *complex volume* of ρ is defined by

$$\widehat{c}\big(B\rho_*\big([M,\partial M]\big)\big)=i\,\mathrm{Vol}_{\mathbb{C}}(\rho), \tag{2.7}$$

where $B\rho:(M,\partial M)\to(B\,\mathrm{SL}(n,\mathbb{C})^\delta,BN^\delta)$ is a classifying map for ρ.

Remark 2.4

Unlike when M is closed, the classifying map is not uniquely determined by ρ; it depends on a choice of decoration (see Section 4). The complex volume, however, is independent of this choice (see Remark 8.5).

2.5. *Central elements of order* 2

For any simple complex Lie group G, there is a canonical homomorphism (defined up to conjugation)

$$\phi_G:\ \mathrm{SL}(2,\mathbb{C})\to G. \tag{2.8}$$

The element $s_G=\phi_G(-I)$ is a central element of G of order dividing 2 and equals $(-I)^{n+1}$ if $G=\mathrm{SL}(n,\mathbb{C})$ (see, e.g., [13, Corollary 2.1]). Let

$$pG=G/\langle s_G\rangle. \tag{2.9}$$

Note that ϕ_G descends to a homomorphism $\mathrm{PSL}(2,\mathbb{C})\to pG$. The next proposition and its corollary follow easily from the Serre spectral sequence.

PROPOSITION 2.5

Suppose that s_G has order 2. The canonical map $p^*:H^4(BpG;\mathbb{Z})\to H^4(BG;\mathbb{Z})$ is surjective with kernel of order dividing 4.

COROLLARY 2.6

There is a canonical characteristic class $\widehat{c}:H_3(pG)\to\mathbb{C}/\pi^2\mathbb{Z}$.

Proof
By Proposition 2.5, there exists a canonical class $u \in H^4(BpG;\pi^2\mathbb{Z})$ such that $p^*(u) = W(P) \in H^4(BG;\pi^2\mathbb{Z})$. Define $\widehat{c} = S_{P,u}$. □

In Section 6.3, we construct a homomorphism

$$\widehat{c}\colon H_3\big(p\,\mathrm{SL}(n,\mathbb{C}), N\big) \to \mathbb{C}/\pi^2\mathbb{Z}, \tag{2.10}$$

which extends $\widehat{c}$ to a characteristic class of bundles with boundary-unipotent holonomy. The complex volume of a representation in $p\,\mathrm{SL}(n,\mathbb{C})$ is defined as in Definition 2.3.

3. The extended Bloch group

We use the conventions of [28]; the original reference is [19].

Definition 3.1
The *pre-Bloch group* $\mathcal{P}(\mathbb{C})$ is the free abelian group on $\mathbb{C}\setminus\{0,1\}$ modulo the *five-term relation*

$$x - y + \frac{y}{x} - \frac{1-x^{-1}}{1-y^{-1}} + \frac{1-x}{1-y} = 0, \quad \text{for } x \neq y \in \mathbb{C}\setminus\{0,1\}. \tag{3.1}$$

The *Bloch group* is the kernel of the map $\nu : \mathcal{P}(\mathbb{C}) \to \wedge^2(\mathbb{C}^*)$ taking z to $z \wedge (1-z)$.

Definition 3.2
The *extended pre-Bloch group* $\widehat{\mathcal{P}}(\mathbb{C})$ is the free abelian group on the set

$$\widehat{\mathbb{C}} = \big\{(e,f) \in \mathbb{C}^2 \;\big|\; \exp(e) + \exp(f) = 1\big\} \tag{3.2}$$

modulo the *lifted five-term relation*

$$(e_0,f_0) - (e_1,f_1) + (e_2,f_2) - (e_3,f_3) + (e_4,f_4) = 0 \tag{3.3}$$

if the equations

$$\begin{aligned} &e_2 = e_1 - e_0, \qquad e_3 = e_1 - e_0 - f_1 + f_0, \qquad f_3 = f_2 - f_1,\\ &e_4 = f_0 - f_1, \qquad f_4 = f_2 - f_1 + e_0 \end{aligned} \tag{3.4}$$

are satisfied. The *extended Bloch group* is the kernel of the map $\widehat{\nu} : \widehat{\mathcal{P}}(\mathbb{C}) \to \wedge^2(\mathbb{C})$ taking (e,f) to $e \wedge f$.

An element $(e,f) \in \widehat{\mathbb{C}}$ with $\exp(e) = z$ is called a *flattening* with *cross-ratio* z. Letting $\mu_{\mathbb{C}}$ denote the roots of unity in $\mathbb{C}^*$, we have a commutative diagram:

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & & \\
0 \arrow[r] & \mu_{\mathbb{C}} \arrow[r, "2\log"] \arrow[d, "\chi"] & \mathbb{C}/4\pi i\mathbb{Z} \arrow[r] \arrow[d, "\chi"] & \mathbb{C}^*/\mu_{\mathbb{C}} \arrow[r] \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \widehat{\mathcal{B}}(\mathbb{C}) \arrow[r] \arrow[d, "\pi"] & \widehat{\mathcal{P}}(\mathbb{C}) \arrow[r, "\widehat{\nu}"] \arrow[d, "\pi"] & \wedge^2(\mathbb{C}) \arrow[r] \arrow[d] & K_2(\mathbb{C}) \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{B}(\mathbb{C}) \arrow[r] \arrow[d] & \mathcal{P}(\mathbb{C}) \arrow[r, "\nu"] \arrow[d] & \wedge^2(\mathbb{C}^*) \arrow[r] \arrow[d] & K_2(\mathbb{C}) \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 & 0 &
\end{tikzcd} \tag{3.5}$$

The map π is induced by the map taking a flattening to its cross-ratio, and χ is the map taking $e \in \mathbb{C}/4\pi i\mathbb{Z}$ to $(e, f + 2\pi i) - (e, f)$, where $f \in \mathbb{C}$ is any element such that $(e, f) \in \widehat{\mathbb{C}}$.

3.1. *The regulator*

By fixing a branch of logarithm, we may write a flattening with cross-ratio z as $[z; p, q] = (\log(z) + p\pi i, \log(1-z) + q\pi i)$, where $p, q \in \mathbb{Z}$ are *even* integers. There is a well-defined *regulator map*

$$\begin{aligned} R :\ & \widehat{\mathcal{P}}(\mathbb{C}) \to \mathbb{C}/4\pi^2\mathbb{Z}, \\ & [z; p, q] \mapsto \mathrm{Li}_2(z) + \frac{1}{2}\big(\log(z) + p\pi i\big)\big(\log(1-z) - q\pi i\big) - \pi^2/6. \end{aligned} \tag{3.6}$$

3.2. *The* PSL$(2,\mathbb{C})$*-variant of the extended Bloch group*

There is another variant of the extended Bloch group using flattenings $[z; p, q]$, where p and q are allowed to be odd. This group is defined as above using the set

$$\widehat{\mathbb{C}}_{\mathrm{odd}} = \Big\{(e, f) \in \mathbb{C}^2 \;\Big|\; \pm\exp(e) \pm \exp(f) = 1\Big\} \tag{3.7}$$

and fits in a diagram similar to (3.5). We use the subscript PSL to denote the variant allowing odd flattenings. We have an exact sequence

$$0 \to \mathbb{Z}/4\mathbb{Z} \to \widehat{\mathcal{B}}(\mathbb{C}) \to \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}} \to 0. \tag{3.8}$$

For odd flattenings, the regulator (3.6) is well defined modulo $\pi^2\mathbb{Z}$.

THEOREM 3.3 (see [16], [19])
There are natural isomorphisms

$$H_3\big(\mathrm{PSL}(2,\mathbb{C})\big) \cong \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}, \qquad H_3\big(\mathrm{SL}(2,\mathbb{C})\big) \cong \widehat{\mathcal{B}}(\mathbb{C}) \tag{3.9}$$

such that the Cheeger–Chern–Simons classes agree with the regulators.

The following result is needed in Section 7. The first part is proved in [28, Lemma 3.16], and the second has a similar proof, which we leave to the reader.

LEMMA 3.4
For $(e,f) \in \widehat{\mathbb{C}}$ and $p,q \in \mathbb{Z}$, we have

$$(e+2\pi i p, f+2\pi i q)-(e,f) = \chi(qe-pf+2pq\pi i) \in \widehat{\mathcal{P}}(\mathbb{C}), \tag{3.10}$$

$$(e+\pi i p, f+\pi i q)-(e,f) = \chi(qe-pf+pq\pi i) \in \widehat{\mathcal{P}}(\mathbb{C})_{\mathrm{PSL}}. \tag{3.11}$$

4. Decorations of representations

In this section, we review the notion of decorated representations introduced in [27]. Throughout the section, G denotes an arbitrary group, not necessarily a Lie group. Let H be subgroup of G. An *ordered simplex* is a simplex with a fixed vertex ordering.

Definition 4.1
A *closed* 3*-cycle* is a cell complex K obtained from a finite collection of ordered 3-simplices Δ_i by gluing together pairs of faces using order-preserving simplicial attaching maps. We assume that all faces have been glued and that the space $M(K)$, obtained by truncating the Δ_i's before gluing, is an oriented 3-manifold with boundary. Let ϵ_i be a sign indicating whether or not the orientation of Δ_i given by the vertex ordering agrees with the orientation of $M(K)$.

Note that up to removing disjoint balls (which does not effect the fundamental group), the manifold $M(K)$ depends only on the underlying topological space of K and not on the choice of 3-cycle structure. Also note that, for any compact, oriented 3-manifold M with (possibly empty) boundary, the space $\widehat{M}$ obtained from M by collapsing each boundary component to a point has a structure of a closed 3-cycle K such that $M = M(K)$.

Let K be a closed 3-cycle, and let $M = M(K)$. Let L denote the space obtained from the universal cover $\widetilde{M}$ of M by collapsing each boundary component to a point. The 3-cycle structure of K induces a triangulation of L and also a triangulation of M by truncated simplices. The covering map extends to a map $L \to K$, and the action of $\pi_1(M)$ on $\widetilde{M}$ by deck transformations extends to an action on L, which is determined by fixing, once and for all, a base point in M together with one of its lifts. Note that the stabilizer of each 0-cell is a *peripheral* subgroup of $\pi_1(M)$, that is, a subgroup induced by inclusion of a boundary component.

Definition 4.2
Let H be a subgroup of G. A representation $\rho\colon \pi_1(M) \to G$ is a (G,H)-*representation* if the image of each peripheral subgroup lies in a conjugate of H.

Definition 4.3
Let ρ be a (G,H)-representation. A *decoration* (on K) of ρ is a ρ-equivariant map

$$D : L^{(0)} \to G/H, \tag{4.1}$$

where $L^{(0)}$ is the 0-skeleton of L.

Note that if $D(e) = gH$, then we have $g^{-1}\rho(\mathrm{Stab}(e))g \subset H$, where $\mathrm{Stab}(e)$ is the stabilizer of e. Since D is ρ-equivariant, it follows that D determines subgroup of H for each boundary component which is well defined up to conjugation in H.

Definition 4.4
Two decorations of ρ are *equivalent* for each boundary component of M, and the corresponding subgroups of H are conjugate (in H).

Remark 4.5
If D is a decoration of ρ, then gD is a decoration of $g\rho g^{-1}$. Since we are only interested in representations up to conjugation, we consider such two decorations to be equal.

PROPOSITION 4.6
Let E be a flat G-bundle over M whose holonomy representation is a (G,H)-representation ρ. There is a one-to-one correspondence between decorations of ρ up to equivalence and reductions of $E_{\partial M}$ to an H-bundle over ∂M.

Proof
For each boundary component S_i of M, choose a base point in S_i and a path to the base point of M. This determines a lift e_i in L of the vertex of K corresponding to S_i and an identification of $\pi_1(S_i)$ with $\mathrm{Stab}(e_i) \subset \pi_1(M)$. If F is a reduction of $E_{\partial M}$, then the holonomy representations $\rho_i : \pi_1(S_i) \to H$ of F_{S_i} are conjugate to ρ, and so there exist $g_i \in G$ such that $g_i^{-1}\rho g_i = \rho_i$. Assigning the coset $g_i H$ to e_i yields a decoration, which, up to equivalence, is independent of the choice of g_i's. On the

other hand, a decoration assigns cosets $g_i H$ to e_i such that $g_i^{-1}\rho(\mathrm{Stab}(e_i))g_i \subset H$. Hence, g_i defines an isomorphism of E_{S_i} with an H-bundle, which, up to isomorphism, depends only on the equivalence class of the decoration. □

4.1. *The diagonal action*

Let $N_G(H)$ denote the normalizer of H in G, and let h denote the number of boundary components of M. There is an action of $(N_G(H)/H)^h$ on the set of equivalence classes of decorations given by right multiplication. More precisely, $(x_1,\dots,x_h)$ acts by taking a decoration D to the decoration D' defined as follows: If D takes a lift v of the ith boundary component to gH, then D' takes v to gx_iH. If $H=N$ and $G=\mathrm{SL}(n,\mathbb{C})$, then $N_G(H)/H$ is the group of diagonal matrices. We thus refer to the action as the *diagonal action*.

PROPOSITION 4.7

If a boundary-unipotent representation ρ is peripherally well behaved, then the diagonal action on the set of equivalence classes of decorations of ρ is transitive.

Proof

It is enough to prove this is the case where there is only one boundary component. In this case, the image of the peripheral subgroup is either trivial or contains an element with a maximal Jordan block. In the first case, all decorations are equivalent; and in the second case, the result follows from the fact that, if a subgroup A of N contains an element with a maximal Jordan form, then the normalizer of A in $\mathrm{SL}(n,\mathbb{C})$ equals the normalizer of N. □

4.2. *The fundamental class of a decorated representation*

A flat G-bundle over M determines a classifying map $M \to BG^\delta$, where the δ indicates that G is regarded as a discrete group. It thus follows from Proposition 4.6 that a decorated representation $\rho\colon \pi_1(M)\to G$ determines a map

$$B\rho\colon (M,\partial M)\to (BG^\delta, BH^\delta). \tag{4.2}$$

In particular, ρ gives rise to a fundamental class

$$[\rho] = B\rho_*\big([M,\partial M]\big)\in H_3(G,H), \tag{4.3}$$

where, by definition, $H_*(G,H)=H_*(BG^\delta,BH^\delta)$. Note that the fundamental class is independent of the particular 3-cycle structure on K.

Recall that M is triangulated by truncated simplices. By restriction, a (G,H)-cocycle on M determines a (G,H)-cocycle on each truncated simplex $\overline{\Delta_i}$. Let

$\overline{B}_*(G,H)$ denote the chain complex generated in degree n by (G,H)-cocycles on a truncated n-simplex. As proved in [27, Section 3], $\overline{B}_*(G,H)$ computes the homology groups $H_3(G,H)$. Note that a (G,H)-cocycle on M determines (up to conjugation) a decorated (G,H)-representation.

PROPOSITION 4.8 ([27, Proposition 5.10])
Let τ be a (G,H)-cocycle on M representing a decorated (G,H)-representation ρ. The cycle

$$\sum \epsilon_i \tau_{\overline{\Delta}_i} \in \overline{B}_3(G,H) \tag{4.4}$$

represents the fundamental class of ρ.

5. Generic decorations and Ptolemy coordinates

In all of the following, $G = \mathrm{SL}(n,\mathbb{C})$, and N is the subgroup of upper triangular matrices with 1's on the diagonal. A (G,N)-representation $\rho\colon \pi_1(M) \to G$ is called *boundary-unipotent*. For a matrix $g \in G$ and a positive integer $i \leq n \in \mathbb{N}$, let $\{g\}_i$ be the ordered set consisting of the first i column vectors of g.

Definition 5.1
A tuple $(g_0N,\dots,g_kN)$ of N-cosets is *generic* if, for each tuple $t=(t_0,\dots,t_k)$ of nonnegative integers with sum n, we have

$$c_t := \det\Big(\bigcup_{i=0}^{k}\{g_i\}_{t_i}\Big) \neq 0, \tag{5.1}$$

where the determinant is viewed as a function on ordered sets of n vectors in $\mathbb{C}^n$. The numbers c_t are called *Ptolemy coordinates*.

Definition 5.2
A decoration of a boundary-unipotent representation is *generic* if, for each simplex Δ of L, the tuple of cosets assigned to the vertices of Δ is generic.

For a set X, let $C_*(X)$ be the acyclic chain complex generated in degree k by tuples $(x_0,\dots,x_k)$. If X is a G-set, then the diagonal G-action makes $C_*(X)$ into a complex of $\mathbb{Z}[G]$-modules. Let $C_*^{\mathrm{gen}}(G/N)$ be the subcomplex of $C_*(G/N)$ generated by generic tuples.

PROPOSITION 5.3
The complex $C_*^{\mathrm{gen}}(G/N)\otimes_{\mathbb{Z}[G]}\mathbb{Z}$ computes the relative homology. If $\rho\colon \pi_1(M)\to G$

is a generically decorated representation, then the fundamental class of ρ is represented by

$$\sum \epsilon_i (g_0^i N, g_1^i N, g_2^i N, g_3^i N) \in C_3^{\mathrm{gen}}(G/N), \tag{5.2}$$

where $(g_0^i N, \dots, g_3^i N)$ are the cosets assigned to lifts $\widetilde{\Delta}_i$ of the Δ_i's.

Proposition 5.3 is proved in Section 9. The idea is that a generic tuple canonically determines a (G, N)-cocycle on a truncated simplex. Hence, $C_*^{\mathrm{gen}}(G/N) \otimes_{\mathbb{Z}[G]} \mathbb{Z}$ is isomorphic to a subcomplex of $\overline{B}_3(G, N)$, and the representation (5.2) of the fundamental class is then an immediate consequence of (4.4).

PROPOSITION 5.4
After a single barycentric subdivision of K, every decoration of a boundary-unipotent representation $\rho : \pi_1(M) \to G$ is equivalent to a generic one.

Proof
After a barycentric subdivision of K, every simplex Δ of K has distinct vertices and at least three vertices of Δ are interior (link is a sphere). Fix lifts $e_i \in L$ of each interior vertex of K. Since the stabilizer of a lift of an interior vertex is trivial, assigning any coset $g_i H$ to e_i yields an equivalent decoration. Since the g_i's can be chosen arbitrarily, the result follows. □

5.1. The geometry of the Ptolemy coordinates
We canonically identify each ordered k-simplex with a standard simplex

$$\Delta_n^k = \left\{ (x_0, \dots, x_k) \in \mathbb{R}^{k+1} \,\middle|\, 0 \le x_i \le n, \sum_{i=0}^{k} x_i = n \right\}. \tag{5.3}$$

Recall that a tuple $(g_0 N, \dots, g_k N)$ has a Ptolemy coordinate for each tuple of $k+1$ nonnegative integers summing to n. In other words, there is a Ptolemy coordinate for each integral point of Δ_n^k. We denote the set of integral points in Δ_n^k by $\Delta_n^k(\mathbb{Z})$.

Definition 5.5
A *Ptolemy assignment* on Δ_n^k is an assignment of a nonzero complex number c_t to each integral point t of Δ_n^k such that the c_t's are the Ptolemy coordinates of some tuple $(g_0 N, \dots, g_k N) \in C_k^{\mathrm{gen}}(G/N)$. A Ptolemy assignment on K is a Ptolemy assignment on each simplex Δ_i of K such that the Ptolemy coordinates agree on identified faces.

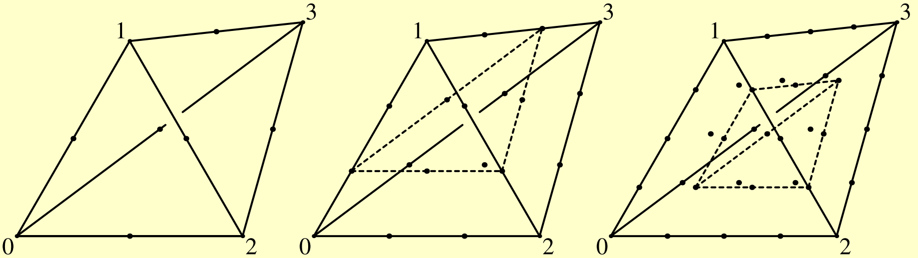


Figure 3. The integral points on Δ_n^3 for $n = 2$, 3, and 4. The indicated subsimplices correspond to $\alpha = (0,1,0,0)$ and $\alpha = (0,1,1,0)$.

Note that a generically decorated boundary-unipotent representation determines a Ptolemy assignment on K. In Section 9, we show that every Ptolemy assignment is induced by a unique decorated representation.

LEMMA 5.6
The number of elements in $\Delta_l^k(\mathbb{Z})$ is $\binom{l+k}{k}$.

Proof
The map $(a_0,\dots,a_k) \mapsto \{a_0+1, a_0+a_1+2,\dots,a_0+\cdots+a_{k-1}+k\}$ gives a bijection between $\Delta_l^k(\mathbb{Z})$ and subsets of $\{1,\dots,l+k\}$ with k elements. $\square$

Let e_i, $0 \le i \le k$, be the ith standard basis vector of $\mathbb{Z}^{k+1}$. For each $\alpha \in \Delta_{n-2}^k(\mathbb{Z})$, the points $\alpha + 2e_i$ in Δ_n^k span a simplex $\Delta^k(\alpha)$, whose integral points are the points $\alpha_{ij} := \alpha + e_i + e_j$ (see Figure 3). We refer to $\Delta^k(\alpha)$ as a *subsimplex* of Δ_n^k. By Lemma 5.6, Δ_n^3 has $\binom{n+3}{3}$ integral points and $\binom{n+1}{3}$ subsimplices.

PROPOSITION 5.7 ([13, Lemma 10.3])
The Ptolemy coordinates of a generic tuple (g_0N, g_1N, g_2N, g_3N) satisfy the Ptolemy relations

$$c_{\alpha_{03}}c_{\alpha_{12}} + c_{\alpha_{01}}c_{\alpha_{23}} = c_{\alpha_{02}}c_{\alpha_{13}}, \quad \alpha \in \Delta_{n-2}^3(\mathbb{Z}). \tag{5.4}$$

Proof
Let $\alpha = (a_0,a_1,a_2,a_3) \in \Delta_{n-2}^3(\mathbb{Z})$. By performing row operations, we may assume that the first $n-2$ rows of the $n \times (n-2)$ matrix

$$\big(\{g_0\}_{a_0}, \{g_1\}_{a_1}, \{g_2\}_{a_2}, \{g_3\}_{a_3}\big) \tag{5.5}$$

are the standard basis vectors. Letting x_i and y_i denote the last two entries of $(g_i)_{a_i+1}$, the Ptolemy relation for α is then equivalent to the (Plücker) relation

$$\det\begin{pmatrix}x_0 & x_3\\ y_0 & y_3\end{pmatrix}\det\begin{pmatrix}x_1 & x_2\\ y_1 & y_2\end{pmatrix}+\det\begin{pmatrix}x_0 & x_1\\ y_0 & y_1\end{pmatrix}\det\begin{pmatrix}x_2 & x_3\\ y_2 & y_3\end{pmatrix}$$
$$=\det\begin{pmatrix}x_0 & x_2\\ y_0 & y_2\end{pmatrix}\det\begin{pmatrix}x_1 & x_3\\ y_1 & y_3\end{pmatrix}, \tag{5.6}$$

which is easily verified. □

Note that the Ptolemy coordinate assigned to the ith vertex of Δ_n^k is $\det(\{g_i\}_n) = \det(g_i) = 1$. We shall thus often ignore the vertex points. Let $\dot{\Delta}_n^k(\mathbb{Z})$ denote the nonvertex integral points of Δ_n^k. The following is proved in Section 9.

PROPOSITION 5.8
For every assignment $c : \dot{\Delta}_n^3(\mathbb{Z}) \to \mathbb{C}^*$, $t \mapsto c_t$ satisfying the Ptolemy relations (5.4), there is a unique Ptolemy assignment on Δ_n^3 whose Ptolemy coordinates are c_t.

COROLLARY 5.9
The set of Ptolemy assignments on K is an algebraic set $P_n(K)$ called the Ptolemy variety. *Its ideal is generated by the Ptolemy relations* (5.4) *(together with an extra equation, making sure that all Ptolemy coordinates are nonzero).*

Remark 5.10
It thus follows that Definition 5.5 agrees with Definition 1.1 when $k = 3$. When $k > 3$ and $n > 2$ there are further relations among the Ptolemy coordinates. We shall not need these here.

5.2. *The diagonal action and the reduced Ptolemy variety*
If $d_0, \dots, d_3$ are diagonal matrices with $d_i = \operatorname{diag}(d_{i0}, \dots, d_{i,n-1})$, then it follows from (5.1) that if the Ptolemy coordinates of a tuple $(g_0N, \dots, g_3N)$ are c_t, then the Ptolemy coordinates c'_t of the tuple $(g_0d_0N, \dots, g_3d_3N)$ are given by

$$c'_t = c_t \prod_{k=0}^{t_0} d_{0k} \prod_{k=0}^{t_1} d_{1k} \prod_{k=0}^{t_2} d_{2k} \prod_{k=0}^{t_3} d_{3k}. \tag{5.7}$$

We therefore have an action of H^h on $P_n(K)$, which agrees with the action in Section 4.1. The quotient $P_n(K)_{\mathrm{red}}$ is called the *reduced Ptolemy variety*.

5.3. $p\,\mathrm{SL}(n,\mathbb{C})$*-Ptolemy coordinates*
When n is even, a $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignment on Δ_n^k may be defined as in Definition 5.5. Note, however, that the Ptolemy coordinates are now defined only up to

a sign. Since we are mostly interested in 3-cycles, the following definition is more useful.

Definition 5.11
Let $\Delta = \Delta_n^3$, and let $\sigma \in Z^2(\Delta;\mathbb{Z}/2\mathbb{Z})$ be a cellular 2-cocycle. A $p\,\mathrm{SL}(n,\mathbb{C})$-*Ptolemy assignment* on Δ with *obstruction cocycle* σ is an assignment $c : \dot{\Delta}_n^3(\mathbb{Z}) \to \mathbb{C}^*$ satisfying the $p\,\mathrm{SL}(n,\mathbb{C})$-*Ptolemy relations*

$$\sigma_2\sigma_3 c_{\alpha_{03}} c_{\alpha_{12}} + \sigma_0\sigma_3 c_{\alpha_{01}} c_{\alpha_{23}} = c_{\alpha_{02}} c_{\alpha_{13}}. \tag{5.8}$$

Here, $\sigma_i \in \mathbb{Z}/2\mathbb{Z} = \langle \pm 1\rangle$ is the value of σ on the face opposite the ith vertex of Δ. A $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignment on K with obstruction cocycle $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$ is a $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignment c^i on each simplex Δ_i of K such that the Ptolemy coordinates agree on identified faces, and such that the obstruction cocycle of c^i is σ_{Δ_i}.

Note that, for each $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$, the set of $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignments on K form a variety $P_n^\sigma(K)$. We show in Section 9 that this variety depends only on the cohomology class of σ in $H^2(K;\mathbb{Z}/2\mathbb{Z}) = H^2(M,\partial M;\mathbb{Z}/2\mathbb{Z})$ and that the Ptolemy variety parameterizes generically decorated boundary-unipotent $p\,\mathrm{SL}(n,\mathbb{C})$-representations whose obstruction class to lifting to a boundary-unipotent $\mathrm{SL}(n,\mathbb{C})$-representation is σ. The diagonal action (5.7) is defined on $P_n^\sigma(K)$ as well, and the quotient is denoted by $P_n^\sigma(K)_{\mathrm{red}}$. Note that when σ is the trivial cocycle taking all 2-cells to 1, $P^\sigma(K) = P(K)$.

5.4. *Cross-ratios and flattenings*

For $x \in \mathbb{C}\setminus\{0\}$, let $\widetilde{x} = \log(x)$, where log is some fixed (set theoretic) section of the exponential map.

Given a Ptolemy assignment c on $\Delta_{n=2}^3$, we endow $\Delta_{n=2}^3$ with the shape of an ideal simplex with cross-ratio $z = \frac{c_{03}c_{12}}{c_{02}c_{13}}$ and a flattening

$$\lambda(c) = (\widetilde{c}_{03} + \widetilde{c}_{12} - \widetilde{c}_{02} - \widetilde{c}_{13}, \widetilde{c}_{01} + \widetilde{c}_{23} - \widetilde{c}_{02} - \widetilde{c}_{13}) \in \widehat{\mathcal{P}}(\mathbb{C}). \tag{5.9}$$

By Propositions 5.7 and 5.8, a Ptolemy assignment on Δ_n^3 induces a Ptolemy assignment c_α on each subsimplex $\Delta^3(\alpha)$. We thus have a map

$$\lambda : P_n(K) \to \widehat{\mathcal{P}}(\mathbb{C}), \qquad c \mapsto \sum_i \epsilon_i \sum_{\alpha\in\Delta_{n-2}^3(\mathbb{Z})} \lambda(c_\alpha^i). \tag{5.10}$$

Similarly, we have a map $P_n^\sigma(K) \to \widehat{\mathcal{P}}(\mathbb{C})_{\mathrm{PSL}}$ defined by the same formula. We next prove that these maps have image in the respective extended Bloch groups.

Remark 5.12
The shapes associated to a Ptolemy assignment satisfy equations resembling Thurston's gluing equations. This is studied in [14].

6. A chain complex of Ptolemy assignments

Let Pt_k^n be the free abelian group on Ptolemy assignments on Δ_n^k. The usual boundary map induces a boundary map $Pt_k^n \to Pt_{k-1}^n$, and the natural map $C_*^{\mathrm{gen}}(G/N) \to Pt_*^n$ taking a tuple $(g_0N, \dots, g_kN)$ to its Ptolemy assignment is a chain map. The result below is proved in Section 9.

PROPOSITION 6.1
A generic tuple is determined up to the diagonal G-action by its Ptolemy coordinates.

COROLLARY 6.2
The natural map induces an isomorphism

$$C_*^{\mathrm{gen}}(G/N) \otimes_{\mathbb{Z}[G]} \mathbb{Z} \cong Pt_*^n. \tag{6.1}$$

In particular, $H_*(G, N) = H_*(Pt_*^n)$.

LEMMA 6.3
Let $c \in Pt_k^n$ be a Ptolemy assignment, and let $\alpha \in \Delta_{n-2}^k(\mathbb{Z})$. The Ptolemy coordinates $c_{\alpha_{ij}}$, $i \neq j$ are the Ptolemy coordinates of a unique Ptolemy assignment c_α on the subsimplex $\Delta^k(\alpha)$.

Proof
For $1 \leq k \leq 3$, this follows from Proposition 5.8. For $k > 3$, the result follows by induction, using the fact that 5 Ptolemy coordinates on Δ_2^3 determine the last. □

A Ptolemy assignment c on Δ_n^k thus induces a Ptolemy assignment c_α on each subsimplex. We thus have maps

$$J_k^n : Pt_k^n \to Pt_k^2, \qquad c \mapsto \sum_{\alpha \in \Delta_{n-2}^k(\mathbb{Z})} c_\alpha. \tag{6.2}$$

For a Ptolemy assignment $c \in Pt_k^n$, let $c_{\underline{i}} \in Pt_{k-1}^n$ be the induced Ptolemy assignment on the ith face of Δ_n^k; that is, we have $\partial(c) = \sum_{i=0}^k (-1)^i c_{\underline{i}}$. Note that

$$(c_{\underline{i}})_{(a_0,\dots,a_{k-1})} = c_{(a_0,\dots,a_{i-1},0,a_i,\dots,a_{k-1})_{\underline{i}}} \in Pt_{k-1}^2. \tag{6.3}$$

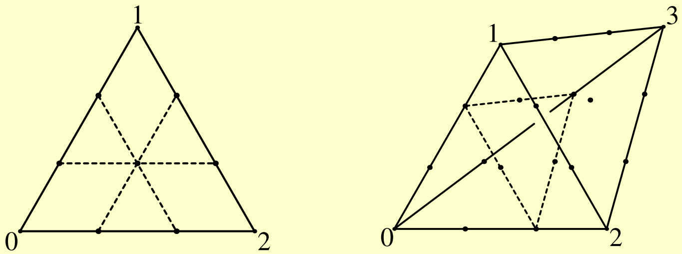


Figure 4. The dotted lines in the left figure indicate c_{β^0}, c_{β^1}, and c_{β^2} for $k=2$. The triangle in the right figure indicates c_{β^0} for $k=3$. Here, $n=3$ and $\beta=0$.

For $\beta\in\Delta^k_{n-3}(\mathbb{Z})$, let $c_{\beta^i}=c_{(\beta+e_i)_{\underline{i}}}\in Pt^2_{k-1}$, and define $\partial_\beta(c)\in Pt^2_{k-1}$ by

$$\partial_\beta(c)=\sum_{i=0}^{k}(-1)^i c_{\beta^i}\in Pt^2_{k-1}. \tag{6.4}$$

The geometry is explained in Figure 4.

PROPOSITION 6.4
Let $c\in Pt^n_k$. We have

$$\partial\big(J^n_k(c)\big)-J^n_{k-1}\big(\partial(c)\big)=\sum_{\beta\in\Delta^k_{n-3}(\mathbb{Z})}\partial_\beta(c)\in Pt^2_{k-1}. \tag{6.5}$$

Proof
By (6.3), we have

$$\begin{aligned}\partial\big(J^n_k(c)\big)-J^n_{k-1}\big(\partial(c)\big)&=\sum_{i=0}^{k}(-1)^i\sum_{\alpha\in\Delta^k_{n-2}(\mathbb{Z})}c_{\alpha_{\underline{i}}}-\sum_{i=0}^{k}(-1)^i\sum_{\substack{\alpha\in\Delta^k_{n-2}(\mathbb{Z})\\ a_i=0}}c_{\alpha_{\underline{i}}}\\ &=\sum_{i=0}^{k}(-1)^i\sum_{\substack{\alpha\in\Delta^k_{n-2}(\mathbb{Z})\\ a_i>0}}c_{\alpha_{\underline{i}}}\\ &=\sum_{\beta\in\Delta^k_{n-3}(\mathbb{Z})}\sum_{i=0}^{k}(-1)^i c_{(\beta+e_i)_{\underline{i}}}\\ &=\sum_{\beta\in\Delta^k_{n-3}(\mathbb{Z})}\partial_\beta(c)\end{aligned} \tag{6.6}$$

as desired. $\square$

6.1. The map to the extended Bloch group

We wish to define a map

$$\lambda : H_3\big(\mathrm{SL}(n,\mathbb{C}), N\big) \to \widehat{\mathcal{B}}(\mathbb{C}). \tag{6.7}$$

Letting $\widetilde{x}$ denote a logarithm of x, we consider the maps

$$\lambda : Pt_3^2 \to \mathbb{Z}[\widehat{\mathbb{C}}], \qquad c \mapsto (\widetilde{c}_{03} + \widetilde{c}_{12} - \widetilde{c}_{02} - \widetilde{c}_{13}, \widetilde{c}_{01} + \widetilde{c}_{23} - \widetilde{c}_{02} - \widetilde{c}_{13}) \tag{6.8}$$

$$\mu : Pt_2^2 \to \wedge^2(\mathbb{C}), \qquad c \mapsto -\widetilde{c}_{01} \wedge \widetilde{c}_{02} + \widetilde{c}_{01} \wedge \widetilde{c}_{12} - \widetilde{c}_{02} \wedge \widetilde{c}_{12} + \widetilde{c}_{02} \wedge \widetilde{c}_{02}. \tag{6.9}$$

Remark 6.5

The term $\widetilde{c}_{02} \wedge \widetilde{c}_{02}$ vanishes in $\wedge^2(\mathbb{C})$, but over general fields this term is needed.

LEMMA 6.6 ([28, Lemma 6.9])

Let $\mathbb{Z}[\widehat{\mathrm{FT}}]$ be the subgroup of $\mathbb{Z}[\widehat{\mathbb{C}}]$ generated by the lifted five-term relations. There is a commutative diagram

$$\begin{tikzcd} Pt_4^2 \arrow[r, "\partial"] \arrow[d, "\lambda\circ\partial"] & Pt_3^2 \arrow[r, "\partial"] \arrow[d, "\lambda"] & Pt_2^2 \arrow[d, "\mu"] \\ \mathbb{Z}[\widehat{\mathrm{FT}}] \arrow[r, hook] & \mathbb{Z}[\widehat{\mathbb{C}}] \arrow[r, "\widehat{\nu}"] & \wedge^2(\mathbb{C}) \end{tikzcd} \tag{6.10}$$

It follows that λ induces a map $\lambda : H_3(\mathrm{SL}(2,\mathbb{C}), N) \to \widehat{\mathcal{B}}(\mathbb{C})$. This map equals the map defined in [27, Section 7]. The fact that λ is independent of the choice of logarithm is proved in [27, Remark 6.11] and also follows from Proposition 7.7 below.

LEMMA 6.7

For each $c \in Pt_4^n$ and each $\beta \in \Delta_{n-3}^4(\mathbb{Z})$, we have

$$\lambda\big(\partial_\beta(c)\big) = 0 \in \widehat{\mathcal{P}}(\mathbb{C}). \tag{6.11}$$

Proof

Let $(e_i, f_i) = \lambda(c_{\beta^i})$ be the flattening associated to c_{β^i}. We prove that the flattenings satisfy the five-term relation by proving that the equations (3.4) are satisfied. We have

$$\begin{aligned} e_0 &= \widetilde{c}_{\beta+(1,1,0,0,1)} + \widetilde{c}_{\beta+(1,0,1,1,0)} - \widetilde{c}_{\beta+(1,1,0,1,0)} - \widetilde{c}_{\beta+(1,0,1,0,1)}, \\ e_1 &= \widetilde{c}_{\beta+(1,1,0,0,1)} + \widetilde{c}_{\beta+(0,1,1,1,0)} - \widetilde{c}_{\beta+(1,1,0,1,0)} - \widetilde{c}_{\beta+(0,1,1,0,1)}, \\ e_2 &= \widetilde{c}_{\beta+(1,0,1,0,1)} + \widetilde{c}_{\beta+(0,1,1,1,0)} - \widetilde{c}_{\beta+(1,0,1,1,0)} - \widetilde{c}_{\beta+(0,1,1,0,1)}, \end{aligned} \tag{6.12}$$

and it follows that $e_2 = e_1 - e_0$ as desired. The other four equations are proved similarly. $\square$

LEMMA 6.8
For each $c \in Pt_3^n$ and each $\beta \in \Delta^3_{n-3}(\mathbb{Z})$, $\mu(\partial_\beta(c)) = 0 \in \wedge^2(\mathbb{C})$.

Proof
We have

$$\begin{aligned}\mu(c_{\beta^0}) = &-\widetilde{c}_{\beta+(1,1,1,0)} \wedge \widetilde{c}_{\beta+(1,1,0,1)} + \widetilde{c}_{\beta+(1,1,1,0)} \wedge \widetilde{c}_{\beta+(1,0,1,1)} \\ &- \widetilde{c}_{\beta+(1,1,0,1)} \wedge \widetilde{c}_{\beta+(1,0,1,1)} + \widetilde{c}_{\beta+(1,1,0,1)} \wedge \widetilde{c}_{\beta+(1,1,0,1)}. \end{aligned} \tag{6.13}$$

Using this together with the similar formulas for $\mu(c_{\beta^i})$, we obtain that

$$\sum (-1)^i \mu(c_{\beta^i}) = 0 \in \wedge^2(\mathbb{C}),$$

proving the result. □

COROLLARY 6.9
The map $\lambda \circ J_3^n$ induces a map

$$\lambda : H_3\big(\mathrm{SL}(n, \mathbb{C}), N\big) \to \widehat{\mathcal{B}}(\mathbb{C}). \tag{6.14}$$

Proof
Using Proposition 6.4, this follows from Lemma 6.7 and Lemma 6.8. □

Remark 6.10
For $n = 3$, this map agrees with the map considered in [29, Section 7.1].

Definition 6.11
The *extended Bloch group element* of a decorated (G, N)-representation ρ is defined by $\lambda([\rho])$, where $[\rho] \in H_3(\mathrm{SL}(n, \mathbb{C}), N)$ is the fundamental class of ρ.

Note that, if the decoration of ρ is generic and c is the corresponding Ptolemy assignment, then the extended Bloch group element is given by $\lambda(c)$, where $\lambda : P_n(K) \to \widehat{\mathcal{P}}(\mathbb{C})$ is given by (5.10).

PROPOSITION 6.12
The map $\lambda : P_n(K) \to \widehat{\mathcal{P}}(\mathbb{C})$ has an image in $\widehat{\mathcal{B}}(\mathbb{C})$.

Proof
If $c \in P_n(K)$ is a Ptolemy assignment on K, then we have a cycle $\alpha = \sum_i \epsilon_i c^i \in Pt_3^n$, and one easily checks that $\lambda(c)$ as defined in (5.10) equals $\lambda([\alpha])$. This proves the result. □

6.2. Stabilization

We now prove that the map $\lambda : H_3(\mathrm{SL}(n,\mathbb{C}),N) \to \widehat{\mathcal{B}}(\mathbb{C})$ respects stabilization. We regard $\mathrm{SL}(n-1,\mathbb{C})$ as a subgroup of $\mathrm{SL}(n,\mathbb{C})$ via the standard inclusion adding a 1 as the upper-left entry.

Let $\pi : M(n,\mathbb{C}) \to M(n-1,\mathbb{C})$ be the map sending a matrix to the submatrix obtained by removing the first row and last column. The subgroup $D_k(\mathrm{SL}(n,\mathbb{C})/N)$ of $C_k^{\mathrm{gen}}(\mathrm{SL}(n,\mathbb{C})/N)$ generated by tuples $(g_0N,\dots,g_kN)$ such that the upper-left entry of each g_i is 1 and such that

$$\big(\pi(g_0)N,\dots,\pi(g_k)N\big) \in C_k^{\mathrm{gen}}\big(\mathrm{SL}(n-1,\mathbb{C})/N\big) \tag{6.15}$$

form an $\mathrm{SL}(n-1,\mathbb{C})$-complex. Consider the $\mathrm{SL}(n-1,\mathbb{C})$-invariant chain maps

$$\pi : D_*\big(\mathrm{SL}(n,\mathbb{C})/N\big) \to Pt_*^{n-1}, \qquad i : D_*\big(\mathrm{SL}(n,\mathbb{C})/N\big) \to Pt_*^{n}, \tag{6.16}$$

where the first map is induced by π and the second is induced by the inclusion $D_*(\mathrm{SL}(n,\mathbb{C})/N) \to C_*^{\mathrm{gen}}(\mathrm{SL}(n,\mathbb{C})/N)$. Let $D_k = D_k(\mathrm{SL}(n,\mathbb{C})/N) \otimes_{\mathbb{Z}[\mathrm{SL}(n-1,\mathbb{C})]} \mathbb{Z}$.

LEMMA 6.13

The maps $\lambda \circ \pi$ and $\lambda \circ i$ from D_3 to $\widehat{\mathcal{P}}(\mathbb{C})$ agree on cycles.

Proof

Let $c \in D_k$ be induced by a tuple $(g_0N,\dots,g_kN) \in D_k(\mathrm{SL}(n,\mathbb{C})/N)$, and let c^I be the collection of Ptolemy coordinates associated to $(N,g_0N,\dots,g_kN)$. Since some Ptolemy coordinates may be zero, c^I is not necessarily a Ptolemy assignment. Note, however, that c^I_α *is* a Ptolemy assignment for each $(a_0,\dots,a_{k+1}) \in \Delta^{k+1}_{n-2}(\mathbb{Z})$ with $a_0 = 0$. Note also that $c^I_\alpha \in Pt^2_{k+1}$ depends only on c. Hence, there is a map

$$P : D_k \to Pt^2_{k+1}, \qquad c \mapsto \sum_{\substack{\alpha \in \Delta^{k+1}_{n-2}(\mathbb{Z})\\ a_0=0}} c^I_\alpha. \tag{6.17}$$

We wish to prove the following:

$$\partial P(c) + P\partial(c) = J_k^n\big(i(c)\big) - J_k^{n-1}\big(\pi(c)\big) + \sum_{\substack{\beta \in \Delta^{k+1}_{n-3}(\mathbb{Z})\\ b_0=0}} \partial_\beta(c^I) \in Pt^2_{k+1}. \tag{6.18}$$

Given this, the result follows immediately from Lemma 6.7.

One easily verifies that

$$c^I_{(\underline{1},b_0,\dots,b_k)} = \pi(c)_{(b_0,\dots,b_k)} \in Pt_k^{n-1}, \quad (b_0,\dots,b_k) \in \Delta^k_{n-3}(\mathbb{Z}), \tag{6.19}$$

$$c^I_{(\underline{0},a_0,\dots,a_k)} = i(c)_{(a_0,\dots,a_k)}, \quad (a_0,\dots,a_k) \in \Delta^k_{n-2}(\mathbb{Z}). \tag{6.20}$$

Using this, one has

$$\begin{aligned}
\partial P(c) + P\partial(c) &= \sum_{\alpha\in\Delta^k_{n-2}(\mathbb{Z})} i(c)_\alpha + \sum_{i=1}^{k+1}(-1)^i \sum_{\substack{\alpha\in\Delta^{k+1}_{n-2}(\mathbb{Z})\\ a_0=0}} c^I_{\alpha_{\underline{i}}} \\
&\quad + \sum_{i=0}^{k}(-1)^i \sum_{\substack{\alpha\in\Delta^{k+1}_{n-2}(\mathbb{Z})\\ a_0=0, a_{i+1}=0}} c^I_{\alpha_{\underline{i+1}}} \\
&= \sum_{\alpha\in\Delta^k_{n-2}(\mathbb{Z})} i(c)_\alpha + \sum_{i=1}^{k+1}(-1)^i \sum_{\substack{\alpha\in\Delta^{k+1}_{n-2}(\mathbb{Z})\\ a_0=0, a_i>0}} c^I_{\alpha_{\underline{i}}} \\
&= \sum_{\alpha\in\Delta^k_{n-2}(\mathbb{Z})} i(c)_\alpha + \sum_{\substack{\beta\in\Delta^{k+1}_{n-3}(\mathbb{Z})\\ b_0=0}} \sum_{i=1}^{k+1}(-1)^i c^I_{\beta^i} \\
&= \sum_{\alpha\in\Delta^k_{n-2}(\mathbb{Z})} i(c)_\alpha - \sum_{\substack{\beta\in\Delta^{k+1}_{n-3}(\mathbb{Z})\\ b_0=0}} c^I_{\beta^0} + \sum_{\substack{\beta\in\Delta^{k+1}_{n-3}(\mathbb{Z})\\ b_0=0}} \partial_\beta(c^I) \\
&= J^n_k\big(i(c)\big) - J^{n-1}_k\big(\pi(c)\big) + \sum_{\substack{\beta\in\Delta^{k+1}_{n-3}(\mathbb{Z})\\ b_0=0}} \partial_\beta(c^I).
\end{aligned} \tag{6.21}$$

This proves (6.18), and hence the result. □

PROPOSITION 6.14
The map $\lambda : H_3(\mathrm{SL}(n,\mathbb{C}), N) \to \widehat{\mathcal{B}}(\mathbb{C})$ *respects stabilization.*

Proof
First, note that π induces an isomorphism $D^0(\mathrm{SL}(n,\mathbb{C})/N) \cong C^0(\mathrm{SL}(n-1)/N)$. Using a standard cone argument, one easily checks that $D_*(\mathrm{SL}(n,\mathbb{C})/N)$ is a free $\mathrm{SL}(n-1,\mathbb{C})$-resolution of $\mathrm{Ker}(D^0(\mathrm{SL}(n,\mathbb{C})/N) \to \mathbb{Z})$. Hence, D_* computes $H_*(\mathrm{SL}(n-1,\mathbb{C}), N)$, and the result follows from Lemma 6.13. □

6.3. $p\,\mathrm{SL}(n,\mathbb{C})$-*Ptolemy assignments*

When n is even, define pPt^n_* to be the complex of Ptolemy coordinates of generic tuples in $p\,\mathrm{SL}(n,\mathbb{C})/N$. The Ptolemy coordinates are defined as in (5.1) and take values in $\mathbb{C}^*/\langle\pm 1\rangle$. As in (6.1), we have an isomorphism $C^{\mathrm{gen}}_*(p\,\mathrm{SL}(n,\mathbb{C})/$

$N)_{p\,\mathrm{SL}(n,\mathbb{C})} \cong pPt_*^n$. For $c \in \mathbb{C}^*/\langle\pm 1\rangle$ let $\widetilde{c} \in \mathbb{C}$ be the image of some fixed set-theoretic section of $\mathbb{C} \xrightarrow{\exp} \mathbb{C}^* \to \mathbb{C}^*/\langle\pm 1\rangle$, for example, $\frac{1}{2}\log(x^2)$ (the particular choice is inessential). The map

$$\lambda : pPt_3^2 \to \mathbb{Z}[\widehat{\mathbb{C}}_{\mathrm{odd}}], \qquad c \mapsto (\widetilde{c}_{03}+\widetilde{c}_{12}-\widetilde{c}_{02}-\widetilde{c}_{13}, \widetilde{c}_{01}+\widetilde{c}_{23}-\widetilde{c}_{02}-\widetilde{c}_{13}) \tag{6.22}$$

induces a map $H_3(\mathrm{PSL}(2,\mathbb{C}), N) \to \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}$, which agrees with the map constructed in [27, Section 3]. By precomposing λ with the map $pJ_3^n : pPt_3^n \to pPt_3^2$ defined as in (6.2), we obtain a map

$$\lambda : H_3\big(p\,\mathrm{SL}(n,\mathbb{C}), N\big) \to \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}, \tag{6.23}$$

which commutes with stabilization. This proves that a decorated boundary-unipotent representation in $p\,\mathrm{SL}(n,\mathbb{C})$ determines an element in $\widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}$. The proofs of the above assertions are identical to their $\mathrm{SL}(n,\mathbb{C})$-analogues.

7. Invariance under the diagonal action

We now show that the extended Bloch group element of a decorated representation is invariant under the diagonal action. We first prove that we can choose logarithms of the Ptolemy coordinates independently, without affecting the extended Bloch group element.

Definition 7.1
Let $c : \dot{\Delta}_n^k(\mathbb{Z}) \to \mathbb{C}^*$ be a Ptolemy assignment. A *lift* of c is an assignment $\widetilde{c} : \dot{\Delta}_n^k(\mathbb{Z}) \to \mathbb{C}$ such that $\exp(\widetilde{c}) = c$.

For any lift $\widetilde{c}$ of a Ptolemy assignment c on Δ_2^3, we have a flattening

$$\lambda(\widetilde{c}) = (\widetilde{c}_{03}+\widetilde{c}_{12}-\widetilde{c}_{02}-\widetilde{c}_{13}, \widetilde{c}_{01}+\widetilde{c}_{23}-\widetilde{c}_{02}-\widetilde{c}_{13}) \in \widehat{\mathbb{C}}. \tag{7.1}$$

Definition 7.2
The *log-parameters* of a flattening $(e, f) \in \widehat{\mathbb{C}}$ are defined by

$$w_{ij} = \begin{cases} e & \text{if } ij = 01 \text{ or } ij = 23, \\ -f & \text{if } ij = 12 \text{ or } ij = 03, \\ -e+f & \text{if } ij = 02 \text{ or } ij = 13. \end{cases} \tag{7.2}$$

LEMMA 7.3
Let $\widetilde{c} : \dot{\Delta}_2^3(\mathbb{Z}) \to \mathbb{C}$ be a lifted Ptolemy assignment, and let w_{ij} be the log-parameters of $\lambda(\widetilde{c})$. Fix $i < j \in \{0, \dots, 3\}$, and let $\widetilde{c}'$ be the lifted Ptolemy assignment obtained from $\widetilde{c}$ by adding $2\pi\sqrt{-1}$ to $\widetilde{c}_{ij}$. Then

$$\lambda(\widetilde{c}') - \lambda(\widetilde{c}) = \chi(w_{ij} + 2\pi\sqrt{-1}\delta_{ij}), \tag{7.3}$$

where δ_{ij} is 1 *if* $ij = 02$ *or* 13 *and* 0 *otherwise.*

Proof
Denote the flattening $\lambda(\widetilde{c})$ by (e, f). If $ij = 03$ or 12, it follows from (7.1) that $\lambda(\widetilde{c}') = (e + 2\pi\sqrt{-1}, f)$. Similarly, $\lambda(\widetilde{c}') = (e, f + 2\pi\sqrt{-1})$ if $ij = 01$ or 23, and $\lambda(\widetilde{c}') = (e - 2\pi\sqrt{-1}, f - 2\pi\sqrt{-1})$ if $ij = 02$ or 13. By Lemma 3.4,

$$\begin{aligned} (e + 2\pi\sqrt{-1}, f) - (e, f) &= \chi(-f), \\ (e, f + 2\pi\sqrt{-1}) - (e, f) &= \chi(e), \\ (e - 2\pi\sqrt{-1}, f - 2\pi\sqrt{-1}) &= \chi(-e + f + 2\pi\sqrt{-1}). \end{aligned} \tag{7.4}$$

This proves the result. □

Let $\widetilde{c}$ be a lift of a Ptolemy assignment c. For each $\alpha \in \Delta^3_{n-2}(\mathbb{Z})$, $\widetilde{c}$ induces a lift $\widetilde{c}_\alpha$ of c_α. Consider the element

$$\tau = \sum_{\alpha \in \Delta^k_{n-2}(\mathbb{Z})} \lambda(\widetilde{c}_\alpha) \in \widehat{\mathcal{P}}(\mathbb{C}). \tag{7.5}$$

Fix a point $t_0 \in \dot{\Delta}^k_n(\mathbb{Z})$. We wish to understand the effect on τ of adding $2\pi\sqrt{-1}$ to $\widetilde{c}_{t_0}$. This changes τ into an element $\tau' \in \widehat{\mathcal{P}}(\mathbb{C})$. Let $w_{ij}(\alpha)$ denote the log-parameters of $\lambda(\widetilde{c}_\alpha)$. Note that t_0 either lies on an edge, on a face, or in the interior of Δ^3_n.

LEMMA 7.4
Suppose that t_0 is on the edge ij of Δ^3_n. Then

$$\tau' - \tau = \chi\big(w_{ij}(\alpha) + 2\pi\sqrt{-1}\delta_{ij}\big), \tag{7.6}$$

where $\alpha = t - e_i - e_j$ (i.e., α is such that t_0 is an edge point of $\Delta^3(\alpha)$).

Proof
This follows immediately from Lemma 7.3. □

LEMMA 7.5
Suppose that t_0 is on a face opposite vertex i. Then $\tau' - \tau = (-1)^i \chi(\kappa + 2\pi\sqrt{-1})$, where κ is given by

$$\begin{aligned} \kappa = {} & \widetilde{c}_{\eta_i(0,-1,1)} - \widetilde{c}_{\eta_i(0,1,-1)} - \big(\widetilde{c}_{\eta_i(-1,0,1)} - \widetilde{c}_{\eta_i(1,0,-1)}\big) \\ & + \widetilde{c}_{\eta_i(-1,1,0)} - \widetilde{c}_{\eta_i(1,-1,0)}, \end{aligned} \tag{7.7}$$

where η_i inserts a zero as the ith vertex.

Proof
For simplicity, assume $i=0$. The other cases are proved similarly. There are exactly three α's for which t_0 is an edge point of $\Delta^3(\alpha)$. These are

$$\alpha_0 = t_0 - (0,0,1,1), \qquad \alpha_1 = t_0 - (0,1,0,1), \qquad \alpha_2 = t_0 - (0,1,1,0). \tag{7.8}$$

Note that $\widetilde{c}_t = (\widetilde{c}_{\alpha_0})_{23} = (\widetilde{c}_{\alpha_1})_{13} = (\widetilde{c}_{\alpha_2})_{12}$. Since adding $2\pi\sqrt{-1}$ to $\widetilde{c}_{t_0}$ leaves $\widetilde{c}_\alpha$ unchanged unless $\alpha \in \{\alpha_0, \alpha_1, \alpha_2\}$, Lemma 7.3 implies that

$$\tau' - \tau = \chi\big(w_{23}(\alpha_0)\big) + \chi\big(w_{13}(\alpha_1) + 2\pi\sqrt{-1}\big) + \chi\big(w_{12}(\alpha_2)\big). \tag{7.9}$$

One easily checks that

$$\begin{aligned}
w_{23}(\alpha_0) &= \widetilde{c}_{(1,0,-1,0)} + \widetilde{c}_{(0,1,0,-1)} - \widetilde{c}_{(1,0,0,-1)} - \widetilde{c}_{(0,1,-1,0)},\\
w_{13}(\alpha_1) &= \widetilde{c}_{(1,0,0,-1)} + \widetilde{c}_{(0,-1,1,0)} - \widetilde{c}_{(1,-1,0,0)} - \widetilde{c}_{(0,0,1,-1)},\\
w_{12}(\alpha_2) &= \widetilde{c}_{(1,-1,0,0)} + \widetilde{c}_{(0,0,-1,1)} - \widetilde{c}_{(1,0,-1,0)} - \widetilde{c}_{(0,-1,0,1)},
\end{aligned} \tag{7.10}$$

from which the result follows. ☐

LEMMA 7.6
If t_0 is an interior point, $\tau' = \tau$.

Proof
If t_0 is an interior point, then there are six α's for which t_0 is an edge point of $\Delta^3(\alpha)$. These are α_0, α_1, and α_2 as defined in (7.8), as well as

$$\alpha_3 = t_0 - (1,1,0,0), \qquad \alpha_4 = t_0 - (1,0,1,0), \qquad \alpha_5 = t_0 - (1,0,0,1). \tag{7.11}$$

Again, by Lemma 7.3,

$$\begin{aligned}
\tau' - \tau = {} & \chi\big(w_{23}(\alpha_0)\big) + \chi\big(w_{13}(\alpha_1) + 2\pi\sqrt{-1}\big) + \chi\big(w_{12}(\alpha_2)\big)\\
& + \chi\big(w_{01}(\alpha_3)\big) + \chi\big(w_{02}(\alpha_4) + 2\pi\sqrt{-1}\big) + \chi\big(w_{03}(\alpha_5)\big).
\end{aligned} \tag{7.12}$$

Using (7.10) (and similar formulas for $w_{01}(\alpha_3)$, $w_{02}(\alpha_4)$, and $w_{03}(\alpha_5)$), we see that all terms in (7.12) cancel out. Hence, $\tau' = \tau$. ☐

PROPOSITION 7.7
Let c be a Ptolemy assignment on K. For any lift $\widetilde{c}$ of c, the element

$$\lambda(\widetilde{c}) = \sum_i \sum_{\alpha \in \Delta^k_{n-2}(\mathbb{Z})} \epsilon_i \lambda(\widetilde{c}^i_\alpha) \in \widehat{\mathcal{P}}(\mathbb{C}) \tag{7.13}$$

is independent of the choice of lift. In particular, if c is the Ptolemy assignment of a decorated representation ρ, then $\lambda(\widetilde{c})$ is the extended Bloch group element of ρ.

Proof

Let $\widetilde{c}$ and $\widetilde{c}'$ be lifts of c. Let $t_0 \in \dot{\Delta}^3_n(\mathbb{Z})$. We wish to prove that $\lambda(\widetilde{c}) = \lambda(\widetilde{c}')$. It is enough to prove this when $\widetilde{c}'$ is obtained from $\widetilde{c}$ by adding $2\pi\sqrt{-1}$ to $\widetilde{c}_t$. If t_0 is an interior point, then the result follows immediately from Lemma 7.6. If t_0 is a face point, then t_0 lies in exactly two simplices of K, and it follows from Lemma 7.5 that the two contributions to the change in $\lambda(\widetilde{c})$ appear with opposite signs (by (3.5), $2\chi(2\pi\sqrt{-1}) = 0$). Suppose that t_0 is an edge point. Let C be the 3-cycle obtained by gluing together all the $\Delta^3(\alpha)$'s having t_0 as an edge point, using the face pairings induced from K. Let e be the (interior) 1-cell of C containing t_0. The argument in [27, Theorem 6.5] shows that the total log-parameter around e is zero. It thus follows from Lemma 7.4 that adding $2\pi\sqrt{-1}$ to $\widetilde{c}_{t_0}$ changes $\lambda(\widetilde{c})$ by 2-torsion, which is trivial if and only if the number n of simplices in C for which t is a 02 edge or a 13 edge is even. Consider a curve λ in C encircling e. The vertex ordering induces an orientation on each face of each simplex of C, such that when λ passes through two faces of a simplex in C, the two orientations agree unless e is a 02 edge or a 13 edge. Since M is orientable, it follows that n is even. The second statement follows by letting $\widetilde{c} = \log c$. □

PROPOSITION 7.8

The extended Bloch group element of a decorated boundary-unipotent representation is invariant under the diagonal action.

Proof

The argument is local. Let c be a Ptolemy assignment on Δ^3_n, and let c' be obtained from c by the diagonal action. By (5.7), we have

$$c'_t = c_t \prod_{k=0}^{t_0} d_{0k} \prod_{k=0}^{t_1} d_{1k} \prod_{k=0}^{t_2} d_{2k} \prod_{k=0}^{t_3} d_{3k} \tag{7.14}$$

for diagonal matrices $d_i = \mathrm{diag}(d_{i0}, \ldots, d_{i,n-1})$. Letting log denote a logarithm, and $\widetilde{c}$ a lift of c, define a lift $\widetilde{c}'$ of c' by

$$\widetilde{c}'_t = \widetilde{c}_t + \sum_{k=0}^{t_0} \log(d_{0k}) + \sum_{k=0}^{t_1} \log(d_{1k}) + \sum_{k=0}^{t_2} \log(d_{2k}) + \sum_{k=0}^{t_3} \log(d_{3k}). \tag{7.15}$$

Using this, one easily checks that $\lambda(c_\alpha) = \lambda(c'_\alpha)$ for each i and each $\alpha \in \Delta^3_{n-2}(\mathbb{Z})$. Applying this local argument to each simplex, the result follows from Proposition 7.7. □

COROLLARY 7.9
The extended Bloch group element of a peripherally well-behaved boundary-unipotent representation ρ is independent of the decoration.

Proof
By performing a barycentric subdivision if necessary, we may assume that any decoration is generic. Since ρ is peripherally well behaved, the diagonal action is transitive on equivalence classes of decorations. Since equivalent decorations have the same fundamental class, the result follows. □

7.1. $p\,\mathrm{SL}(n,\mathbb{C})$-decorations
Let n be even. All results in this section have natural analogs for $p\,\mathrm{SL}(n,\mathbb{C})$. The proofs of these are obtained by replacing $2\pi\sqrt{-1}$ by $\pi\sqrt{-1}$, and logarithms by lifts of $\mathbb{C} \xrightarrow{\exp} \mathbb{C}^*/\langle\pm 1\rangle$.

8. A cocycle formula for $\widehat{c}$

Let $i_*: H_3(\mathrm{SL}(n,\mathbb{C})) \to H_3(\mathrm{SL}(n,\mathbb{C}),N)$ denote the natural map. We wish to prove that the composition

$$H_3\big(\mathrm{SL}(n,\mathbb{C})\big) \xrightarrow{i_*} H_3\big(\mathrm{SL}(n,\mathbb{C}),N\big) \xrightarrow{\lambda} \widehat{\mathcal{B}}(\mathbb{C}) \xrightarrow{R} \mathbb{C}/4\pi^2\mathbb{Z} \tag{8.1}$$

equals the Cheeger–Chern–Simons class $\widehat{c}$. Note that i_* is induced by the map $(g_0,\dots,g_3) \mapsto (g_0N,\dots,g_3N)$.

We shall make use of the canonical isomorphisms

$$H_3\big(\mathrm{SL}(n,\mathbb{C})\big) \cong H_3\big(\mathrm{SL}(3,\mathbb{C})\big) \cong H_3\big(\mathrm{SL}(2,\mathbb{C})\big) \oplus K_3^M(\mathbb{C}). \tag{8.2}$$

The first isomorphism is induced by stabilization (see [24]) and the second isomorphism is the $\pm$-eigenspace decomposition with respect to the transpose-inverse involution on $\mathrm{SL}(3,\mathbb{C})$ (see [22]).

LEMMA 8.1 (Suslin [24])
Let $H \subset \mathrm{SL}(3,\mathbb{C})$ be the subgroup of diagonal matrices. The $K_3^M(\mathbb{C})$ summand of $H_3(\mathrm{SL}(3,\mathbb{C}))$ is generated by the elements $B\rho_*([T])$, where $T = S^1\times S^1\times S^1$ is the 3-torus and $\rho: \pi_1(T)\to H$ is a representation.

LEMMA 8.2
Let $T = S^1\times S^1\times S^1$, and let $\rho: \pi_1(T)\to H$ be a representation. The extended Bloch group element $[\rho]\in\widehat{\mathcal{B}}(\mathbb{C})$ of ρ is trivial.

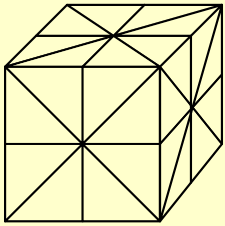


Figure 5. Triangulation of ∂C.

Proof

We regard T as a cube C with opposite faces identified and triangulate C as the cone on the triangulation on ∂C indicated in Figure 5 with cone point in the center. We order the vertices of each simplex by codimension; that is, the 0-vertex is the cone point, the 1-vertex is a face point, and so on. Let $\rho : \pi_1(T) \to H$ be a representation, and pick a decoration of ρ by cosets in general position (the triangulation is such that this is always possible). Note that, for every 3-simplex Δ of T, there is a unique opposite 3-simplex Δ^{opp}, such that the faces opposite the cone point are identified. We may assume that the cone point is decorated by the coset N. If a simplex Δ is decorated by the cosets (N, g_0N, g_1N, g_2N), then the simplex Δ^{opp} must be decorated by the cosets (N, dg_0N, dg_1N, dg_2N), where d is the image of the generator of $\pi_1(T)$ pairing the faces of Δ and Δ^{opp}. It thus follows from (5.2) that the fundamental class is represented by a sum of terms of the form

$$(N, dg_0N, dg_1N, dg_2N) - (N, g_0N, g_1N, g_2N) \in C_3^{\mathrm{gen}}\big(\mathrm{SL}(n,\mathbb{C})/N\big). \tag{8.3}$$

Let c and c' be the Ptolemy assignments associated to the tuples (N, g_0N, g_1N, g_2N) and (N, dg_0N, dg_1N, dg_2N). Letting $d = \mathrm{diag}(d_1, \dots, d_n)$, we have $c'_t = c_t \prod_{i=t_0}^{n} d_i$. Fix a lift $\widetilde{c}$ of c, and consider the lift

$$\widetilde{c}'_t = \widetilde{c}_t + \sum_{i=t_0}^{n} \log(d_i) \tag{8.4}$$

of c'. One now checks that $\lambda(\widetilde{c}'_\alpha) = \lambda(\widetilde{c}_\alpha)$ for all $\alpha \in \dot{\Delta}_n^k(\mathbb{Z})$, so $\lambda(\widetilde{c}) - \lambda(\widetilde{c}') = 0$. This proves the result. □

THEOREM 8.3

The composition $R \circ \lambda \circ i_*$ equals $\widehat{c}$.

Proof

Since λ commutes with stabilization, it follows from [16] that $R \circ \lambda \circ i_* = \widehat{c}$ on $H_3(\mathrm{SL}(2,\mathbb{C}))$. Since $\widehat{c}$ is zero on $K_3^M(\mathbb{C})$ (this follows from Lemma 8.1 and [5, Theorem 8.22]), the result follows from (8.2) and Lemma 8.2. ☐

Remark 8.4

By defining $\widehat{c} = R \circ \lambda : H_3(\mathrm{SL}(n,\mathbb{C}), N) \to \mathbb{C}/4\pi^2\mathbb{Z}$, we have a natural extension of the Cheeger–Chern–Simons class to bundles with boundary-unipotent holonomy, and we can define the complex volume as in Definition 2.3.

Remark 8.5

The fact that the complex volume is independent of the choice of decoration can be seen as follows: We can regard $\widehat{c}$ as a map $P_n(\Delta^3) \to \mathbb{C}/4\pi^2\mathbb{Z}$, and a simple computation shows that the holomorphic 1-form $d\widehat{c}$ involves only coordinates on the boundary of Δ^3. Hence, for a closed 3-cycle K, $\widehat{c} : P_n(K) \to \mathbb{C}/4\pi^2\mathbb{Z}$ is locally constant. The result now follows from the fact that the space of decorations of a representation is path-connected.

9. Recovering a representation from its Ptolemy coordinates

We now show that a Ptolemy assignment on K determines a generically decorated boundary-unipotent representation, which is given explicitly in terms of the Ptolemy coordinates. The idea is that a Ptolemy assignment canonically determines a (G,N)-cocycle on M.

9.1. *The generic (G,N)-cocycle of a tuple*

Definition 9.1

An $(n \times n)$-matrix A is *counterdiagonal* if the only nonzero entries of A are on the lower-left to upper-right diagonal; that is, $A_{ij} = 0$ unless $j = n-i+1$. If $A_{ij} = 0$ for $j > n-i+1$ (resp., $j < n-i+1$), then A is *upper* (resp., *lower*) *countertriangular*.

Given subsets I, J of $\{1,\dots,n\}$, let $A_{I,J}$ denote the submatrix of A whose rows and columns are indexed by I and J, respectively. If $|I| = |J|$, then let $|A|_{I,J}$ denote the minor $\det(A_{I,J})$. Let I^c denote $\{1,\dots,n\} \setminus I$.

Recall that the adjugate $\mathrm{Adj}(A)$ of a matrix A is the matrix whose ijth entry is $(-1)^{i+j}|A|_{\{j\}^c,\{i\}^c}$. It is well known that $\mathrm{Adj}(A) = \det(A)A^{-1}$. The following result by Jacobi (see, e.g., [1, Section 42]) expresses the minors of $\mathrm{Adj}(A)$ in terms of the minors of A.

LEMMA 9.2
Let I, J be subsets of $\{1,\dots,n\}$ with $|I| = |J| = r$. We have

$$\big|\mathrm{Adj}(A)\big|_{I,J} = (-1)^{\sum(I,J)} \det(A)^{r-1} |A|_{J^c,I^c}, \tag{9.1}$$

where $\sum(I,J)$ is the sum of the indices occurring in I and J.

Definition 9.3
A matrix $A \in \mathrm{GL}_n(\mathbb{C})$ is *generic* if $|A|_{\{k,\dots,n\},\{1,\dots,n-k+1\}} \neq 0$ for all $k \in \{1,\dots,n\}$.

Note that A is generic if and only if the Ptolemy coordinates of (N, AN) are nonzero. The following is a generalization of [27, Lemma 3.5].

PROPOSITION 9.4
Let $A \in \mathrm{GL}_n(\mathbb{C})$ be generic. There exist unique $x \in N$ and $y \in N$ such that $q = x^{-1}Ay$ is counterdiagonal. The entries of x, y, and q are given by

$$\begin{aligned} q_{n,1} &= A_{n,1}, \\ q_{n-j+1,j} &= (-1)^{j-1} \frac{|A|_{\{n-j+1,\dots,n\},\{1,\dots,j\}}}{|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}} \quad \textit{for } 1 < j \le n, \end{aligned} \tag{9.2}$$

$$x_{ij} = \frac{|A|_{\{i,j+1,\dots,n\},\{1,\dots,n-j+1\}}}{|A|_{\{j,\dots,n\},\{1,\dots,n-j+1\}}} \quad \textit{for } j > i, \tag{9.3}$$

$$y_{ij} = (-1)^{i+j} \frac{|A|_{\{n-j+2,\dots,n\},\{1,\dots,\widehat{i},\dots,j\}}}{|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}} \quad \textit{for } j > i. \tag{9.4}$$

Proof
It is enough to prove existence and uniqueness of x and y in N such that Ay and $x^{-1}A$ are upper and lower countertriangular, respectively. Suppose that Ay is upper countertriangular. Then the vector $y_{\{1,\dots,j-1\},\{j\}}$ consisting of the part above the counterdiagonal of the jth column vector of y must satisfy

$$A_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}} y_{\{1,\dots,j-1\},\{j\}} + A_{\{n-j+2,\dots,n\},\{j\}} = 0. \tag{9.5}$$

The existence and uniqueness of y, as well as the given formula for the entries, now follow from Cramer's rule. Since $x^{-1}A$ is lower countertriangular if and only if $A^{-1}x$ is upper countertriangular, the existence and uniqueness of x follows. The explicit formula for the entries follows from Jacobi's identity (9.1) and the formula for the entries of y. To obtain the formula for the entries of q, note that $q_{n-j+1,j} = (Ay)_{n-j+1,j}$. Hence, $q_{n,1} = A_{n,1}$, and, for $1 < j \le n$,

$$
\begin{aligned}
q_{n-j+1,j} &= \sum_{i=1}^{j-1} A_{n-j+1,i}\, y_{i,j} + A_{n-j+1,j} \\
&= \frac{\sum_{i=1}^{j} (-1)^{i+j} A_{n-j+1,i} |A|_{\{n-j+2,\dots,n\},\{1,\dots\widehat{i},\dots,j\}}}{|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}} \\
&= (-1)^{j-1} \frac{|A|_{\{n-j+1,\dots,n\},\{1,\dots,j\}}}{|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}},
\end{aligned}
$$

where the second equality follows from (9.4). □

For a generic matrix A, let x_A, y_A, and q_A be the unique matrices provided by Proposition 9.4. Given cosets $g_i N$, $g_j N$, $g_k N$, define

$$
q_{ij} = q_{g_i^{-1} g_j}, \qquad \alpha^i_{jk} = (x_{g_i^{-1} g_j})^{-1} x_{g_i^{-1} g_k}. \tag{9.6}
$$

Definition 9.5
The *generic cocycle* of a generic tuple $(g_0 N, \dots, g_k N)$ is the (G, N)-cocycle on a truncated simplex $\overline{\Delta}$ defined by labeling the long edges by q_{ij} and the short edges by α^i_{jk} (see Figure 6).

PROPOSITION 9.6
The diagonal left G-action on $C^{\mathrm{gen}}_k(G/N)$ is free when $k \geq 1$, and the chain complex $C^{\mathrm{gen}}_{*\geq 1}(G/N) \otimes_{\mathbb{Z}[G]} \mathbb{Z}$ computes relative homology.

Proof
By Proposition 9.4, every generic tuple $(g_0 N, \dots, g_k N)$ may be uniquely written as

$$
g_0 x_{g_0^{-1} g_1} (N, q_{01} N, \alpha^0_{12} q_{02} N, \dots, \alpha^0_{1k} q_{0k} N). \tag{9.7}
$$

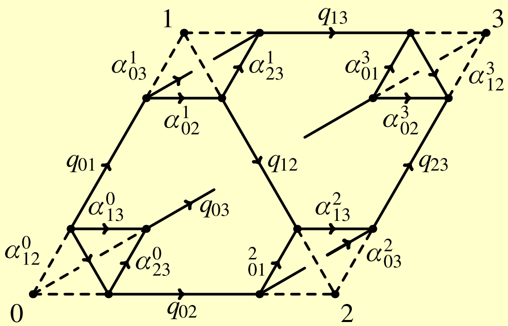


Figure 6. A (G, N)-cocycle on a truncated 3-simplex.

This proves that the G-action is free. Also note that, for each generic tuple $(g_0 N, \dots, g_k N)$, there exists a coset gN such that $(gN, g_0 N, \dots, g_k N)$ is generic. Hence, $C_{*\geq 1}^{\mathrm{gen}}(G/N)$ is acyclic and is thus a free resolution of $\operatorname{Ker}(C_0(G/N) \to \mathbb{Z})$. This proves the result (see, e.g., [27, Theorem 2.1]). □

A generically decorated representation ρ thus determines a (G, N)-cocycle representing ρ. Let $\overline{B}_*^{\mathrm{gen}}(G, N)$ be the subcomplex of $\overline{B}_*(G, N)$ generated by generic cocycles on a standard simplex.

COROLLARY 9.7
We have a canonical isomorphism

$$\overline{B}_*^{\mathrm{gen}}(G, N) = C_*^{\mathrm{gen}}(G/N) \otimes_{\mathbb{Z}[G]} \mathbb{Z}, \tag{9.8}$$

and the fundamental class of a decorated representation is represented as in (4.4).

Proof
The first statement follows from Proposition 9.6 and the second from Theorem 4.8. □

9.2. *Formulas for the long and short edges*
We wish to prove that a generic (G, N)-cocycle is uniquely determined by the Ptolemy coordinates.

Notation 9.8
Let $k \in \{1, \dots, n-1\}$.
(i) For $a_1, \dots, a_n \in \mathbb{C}^*$, let $q(a_1, \dots, a_n)$ be the counterdiagonal matrix whose entries on the counterdiagonal (from lower left to upper right) are $a_1, \dots, a_n$.
(ii) For $x \in \mathbb{C}$, let $x_k(x)$ be the elementary matrix whose $(k, k+1)$ entry is x.
(iii) For $b_1, \dots, b_k \in \mathbb{C}$, let $\pi_k(b_1, \dots, b_k) = x_1(b_1) x_2(b_2) \cdots x_k(b_k)$.

PROPOSITION 9.9
The long edges of a generic (G, N)-cocycle are determined by the Ptolemy coordinates as follows:

$$q_{ij} = q(a_1, \dots, a_n), \qquad a_k = (-1)^{k-1} \frac{c_{(n-k)e_i + k e_j}}{c_{(n-k+1)e_i + (k-1)e_j}}. \tag{9.9}$$

Proof
Let $(g_0 N, \dots, g_k N)$ be a generic tuple, and let $A = g_i^{-1} g_j$. Then $q_{ij} = q_A$. Since

$$|A|_{\{n-j+1,\dots,n\},\{1,j\}} = \det\bigl(\{g_i\}_{n-k}, \{g_j\}_k\bigr) = c_{(n-k)e_i + k e_j}, \tag{9.10}$$

the result follows from (9.2). □

The corresponding formula for the short edges requires considerably more work and is given in Proposition 9.14 below.

LEMMA 9.10
Let A be generic, and let $L = x_A^{-1}A$. The entries $L_{i,n-i+2}$ right below the counter-diagonal are given by

$$L_{i,n-i+2} = (-1)^{n-i} \frac{|A|_{\{i,\dots,n\},\{1,\dots,\widehat{n-i+1},n-i+2\}}}{|A|_{\{i+1,\dots,n\},\{1,\dots,n-i\}}}. \tag{9.11}$$

Proof
We proceed as in the proof of Proposition 9.4. Let $x = x_A^{-1}$. Since L is lower countertriangular, we must have

$$x_{\{i\},\{i+1,\dots,n\}} A_{\{i+1,\dots,n\},\{1,\dots,n-i\}} + A_{\{i\},\{1,\dots,n-i\}} = 0, \tag{9.12}$$

and so, by Cramer's rule,

$$x_{ij} = (-1)^{i+j} \frac{|A|_{\{i,\dots,\widehat{j},\dots,n\},\{1,\dots,n-i\}}}{|A|_{\{i+1,\dots,n\},\{1,\dots,n-i\}}} \quad \text{for } j > i. \tag{9.13}$$

We thus have

$$\begin{aligned}
|A|_{\{i+1,\dots,n\},\{1,\dots,n-i\}} L_{i,n-i+2} &= A_{i,n-i+2} |A|_{\{i+1,\dots,n\},\{1,\dots,n-i\}} \\
&\qquad + \sum_{k=i+i}^{n} (-1)^{i+k} |A|_{\{j,\dots,\widehat{k},\dots,n\},\{1,\dots,n-j\}} A_{k,n-i+2} \\
&= \sum_{k=j}^{n} (-1)^{i+k} |A|_{\{j,\dots,\widehat{k},\dots,n\},\{1,\dots,n-i\}} A_{k,n-i+2} \\
&= (-1)^{n-i} |A|_{\{i,\dots,n\},\{1,\dots,\widehat{n-i+1},\dots,n-i+2\}},
\end{aligned}$$

which proves the result. □

Definition 9.11
Let $A, B \in \mathrm{GL}(n,\mathbb{C})$.
(i) A and B are related by a *type* 0 *move* if all but the last column of A and B are equal.
(ii) A and B are related by a *type* 1 *move* if all but the second last column of A and B are equal.
(iii) A and B are related by a *type* 2 *move* if, for some $j < n-1$, B is obtained from A by switching columns j and $j+1$.

PROPOSITION 9.12

Let A and B be generic, and let A_i and B_i denote the ith column of A, respectively B.

(i) *If A and B are related by a type 0 move, then $x_B = x_A$.*

(ii) *If A and B are related by a type 1 move, then $x_B = x_A x_1(x)$, where*

$$x = -\frac{\det(A_1, \dots, A_{n-1}, B_{n-1}) \det(e_1, e_2, A_1, \dots, A_{n-2})}{\det(e_1, A_1, \dots, A_{n-1}) \det(e_1, A_1, \dots, A_{n-2}, B_{n-1})}. \tag{9.14}$$

(iii) *If A and B are related by a type 2 move switching columns j and $j+1$, $x_B = x_A x_{n-j}(x)$, where*

$$x = -\frac{\det(e_1, \dots, e_{n-j-1}, A_1, \dots, A_{j+1}) \det(e_1, \dots, e_{n-j+1}, A_1, \dots, A_{j-1})}{\det(e_1, \dots, e_{n-j}, A_1, \dots, A_j) \det(e_1, \dots, e_{n-j}, A_1, \dots, A_{j-1}, B_j)}. \tag{9.15}$$

Proof

The first statement follows from the fact that x_A is independent of the last column of A. Suppose that A and B are related by a type 1 move. Using (9.3), one sees that $(x_A)_{ij} = (x_B)_{ij}$ except when $i = 1$ and $j = 2$. It thus follows that $x_B = x_A x_1(x)$, where $x = (x_B)_{12} - (x_A)_{12}$. Letting C be the matrix obtained from A by replacing the nth column by the $(n-1)$th column of B, one has

$$|A|_{\{1,3,\dots,n\},\{1,\dots,n-1\}} = \mathrm{Adj}(C)_{n,2}, \qquad |B|_{\{1,3,\dots,n\},\{1,\dots,n-1\}} = \mathrm{Adj}(C)_{n-1,2},$$
$$|A|_{\{2,\dots,n\},\{1,\dots,n-1\}} = \mathrm{Adj}(C)_{n,1}, \qquad |B|_{\{2,\dots,n\},\{1,\dots,n-1\}} = \mathrm{Adj}(C)_{n-1,1},$$

and it follows from (9.3) that

$$x = (x_B)_{12} - (x_A)_{12} = \frac{\mathrm{Adj}(C)_{n-1,2}}{\mathrm{Adj}(C)_{n-1,1}} - \frac{\mathrm{Adj}(C)_{n,2}}{\mathrm{Adj}(C)_{n,1}}. \tag{9.16}$$

We then have

$$\begin{aligned} x \,\mathrm{Adj}(C)_{n,1} \,\mathrm{Adj}(C)_{n-1,1} &= \mathrm{Adj}(C)_{n-1,2} \,\mathrm{Adj}(C)_{n,1} - \mathrm{Adj}(C)_{n-1,1} \,\mathrm{Adj}(C)_{n,2} \\ &= -\big|\mathrm{Adj}(C)\big|_{\{n-1,n\},\{1,2\}} \\ &= -\det(C) |C|_{\{3,\dots,n\},\{1,\dots,n-2\}} \\ &= -\det(A_1, \dots, A_{n-1}, B_{n-1}) \det(e_1, e_2, A_1, \dots, A_{n-2}), \end{aligned}$$

where the third equality follows from Jacobi's identity (9.1). Since

$$\mathrm{Adj}(C)_{n,1} \,\mathrm{Adj}(C)_{n-1,1} = \det(e_1, A_1, \dots, A_{n-1}) \det(e_1, A_1, \dots, A_{n-2}, B_{n-1}),$$

this proves the second statement.

Now suppose that A and B are related by a type 2 move. Let $E_{j,j+1}$ be the elementary matrix obtained from the identity matrix by switching the jth and $(j+1)$th columns. Then $B = AE_{j,j+1}$. Since $L = x_A^{-1}A$ is lower countertriangular, $x_{n-j}(-\frac{L_{n-j,j+1}}{L_{n-j+1,j+1}})LE_{j,j+1}$ must also be lower countertriangular. We thus have

$$x_B = x_A x_{n-j}\Big(-\frac{L_{n-j,j+1}}{L_{n-j+1,j+1}}\Big)^{-1} = x_A x_{n-j}\Big(\frac{L_{n-j,j+1}}{L_{n-j+1,j+1}}\Big). \tag{9.17}$$

By (9.11) and (9.2), we have

$$\begin{aligned} L_{n-j+1,j+1} &= (-1)^{j-1}\frac{|A|_{\{n-j+1,\dots,n\},\{1,\dots,\widehat{j},j+1\}}}{|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}}, \\ L_{n-j,j+1} &= (-1)^{j}\frac{|A|_{\{n-j,\dots,n\},\{1,\dots,j+1\}}}{|A|_{\{n-j+1,\dots,n\},\{1,\dots,j\}}}. \end{aligned} \tag{9.18}$$

Hence

$$\begin{aligned} &\frac{L_{n-j,j+1}}{L_{n-j+1,j+1}} \\ &\quad = -\frac{|A|_{\{n-j,\dots,n\},\{1,\dots,j+1\}}|A|_{\{n-j+2,\dots,n\},\{1,\dots,j-1\}}}{|A|_{\{n-j+1,\dots,n\},\{1,\dots,j\}}|A|_{\{n-j+1,\dots,n\},\{1,\dots,\widehat{j},j+1\}}} \\ &\quad = -\frac{\det(e_1,\dots,e_{n-j-1},A_1,\dots,A_{j+1})\det(e_1,\dots,e_{n-j+1},A_1,\dots,A_{j-1})}{\det(e_1,\dots,e_{n-j},A_1,\dots,A_j)\det(e_1,\dots,e_{n-j},A_1,\dots,A_{j-1},B_j)}, \end{aligned}$$

proving the third statement. □

Note that any two matrices $A, B \in \mathrm{GL}(n,\mathbb{C})$ are related by a sequence of moves of type 1, 2, and 0 as follows:

$$\begin{aligned} A &\xrightarrow{1} [A_1,\dots,A_{n-2},B_1,A_n] \xrightarrow{2} [A_1,\dots,A_{n-3},B_1,A_{n-2},A_n] \xrightarrow{2} \cdots \\ &\xrightarrow{2} [B_1,A_1,\dots,A_{n-2},A_n] \xrightarrow{1} [B_1,A_1,\dots,A_{n-3},B_2,A_n] \xrightarrow{2} \cdots \\ &\xrightarrow{2} [B_1,B_2,A_1,\dots,A_{n-3},A_n] \xrightarrow{1,2} \cdots \xrightarrow{1,2} [B_1,\dots,B_{n-1},A_n] \xrightarrow{0} B. \end{aligned} \tag{9.19}$$

Consider the tilings of a face ijk, $i<j<k$, of Δ_n^2 by *diamonds* shown in Figure 7. We refer to the diamonds as being of type i, j, and k, respectively.

Definition 9.13
The *diamond coordinate* $d_{r,s}^k$ of a diamond (r,s) of type k is the alternating product of the Ptolemy coordinates assigned to its vertices (see Figure 7).

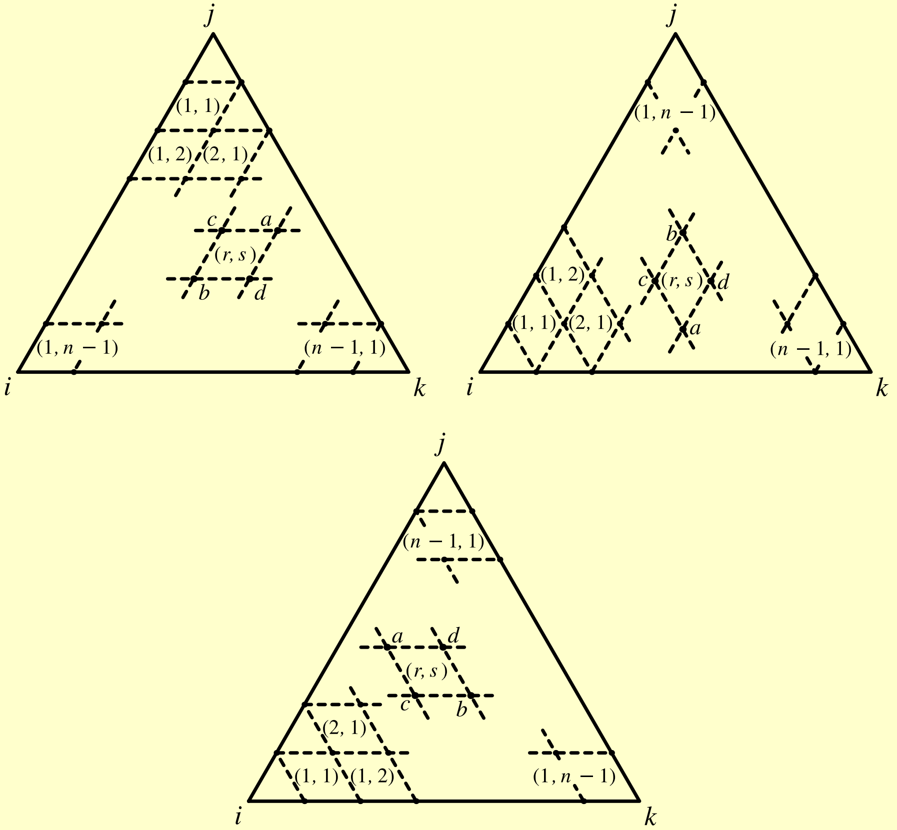


Figure 7. Diamonds of type i, j, and k. The diamond coordinates are $d^i_{r,s} = d^k_{r,s} = \frac{-ab}{cd}$, and $d^j_{r,s} = \frac{ab}{cd}$, where a, b, c, and d are Ptolemy coordinates.

PROPOSITION 9.14

The short edges α^i_{jk}, $j<k$, of a generic (G,N)-cocycle are determined by the Ptolemy coordinates as follows (π_* is defined in 9.8(iii)):

$$\alpha^i_{jk} = \pi_{n-1}(d^i_{1,1},\dots,d^i_{1,n-1})\pi_{n-2}(d^i_{2,1},\dots,d^i_{2,n-2})\cdots\pi_1(d^i_{n-1,1}), \qquad (9.20)$$

where the $d^i_{j,k}$'s are the type i diamond coordinates on the face ijk.

Proof

Let $(g_0N,\dots,g_lN)$ be a generic tuple, and let $A=g_i^{-1}g_j$ and $B=g_i^{-1}g_k$. We assume that $i<j<k$, the other cases being similar. Note that the Ptolemy coordinates on the ijk face are given by

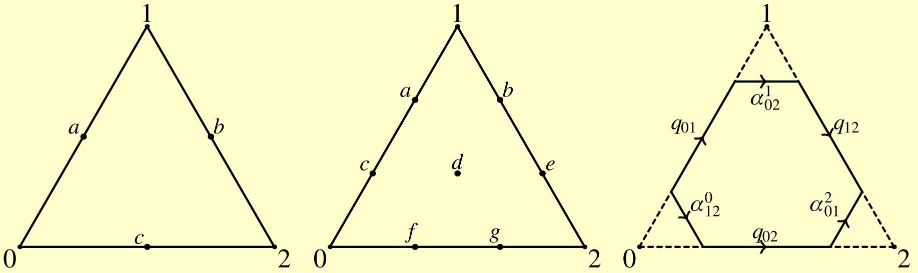


Figure 8. Ptolemy assignments and the corresponding cocycle for $n = 2$ and $n = 3$.

$$c_{t_i e_i + t_j e_k + t_k e_k} = \det(e_1, \dots, e_{t_i}, A_1, \dots, A_{t_j}, B_1, \dots, B_{t_k}). \tag{9.21}$$

Performing the sequence of moves in (9.19), the result follows from Proposition 9.12. □

COROLLARY 9.15
A generic tuple is determined up to the diagonal G-action by its Ptolemy coordinates.

Example 9.16
Suppose that Ptolemy assignments on Δ_n^2, $n \in \{2, 3\}$, are given as in Figure 8. Using (9.9) and (9.20), we obtain that the corresponding (G, N)-cocycle is given by

$$\begin{aligned} &q_{01} = q(a, -1/a), \qquad q_{12} = q(b, -1/b), \qquad q_{02} = q(c, -1/c), \\ &\alpha_{12}^0 = x_1\Big(\frac{-b}{ac}\Big), \qquad \alpha_{02}^1 = x_1\Big(\frac{c}{ab}\Big), \qquad \alpha_{01}^2 = x_1\Big(\frac{-a}{cb}\Big) \end{aligned} \tag{9.22}$$

when $n = 2$, and

$$\begin{aligned} &q_{01} = q(c, -a/c, 1/a), \qquad q_{12} = q(b, -e/b, 1/e), \qquad q_{02} = q(f, -g/f, 1/g), \\ &\alpha_{02}^1 = x_1\Big(\frac{fa}{cd}\Big) x_2\Big(\frac{d}{ab}\Big) x_1\Big(\frac{gb}{de}\Big), \\ &\alpha_{12}^0 = x_1\Big(\frac{-bc}{ad}\Big) x_2\Big(\frac{-d}{cf}\Big) x_1\Big(\frac{-ef}{dg}\Big), \qquad \alpha_{01}^2 = x_1\Big(\frac{-cg}{fd}\Big) x_2\Big(\frac{-d}{ge}\Big) x_1\Big(\frac{-ae}{db}\Big) \end{aligned} \tag{9.23}$$

when $n = 3$.

9.3. *From Ptolemy assignments to decorations*

Corollary 9.15 shows that there is at most one generic (G, N)-cocycle with a given collection of Ptolemy coordinates. We now prove that, when $k \le 3$, there is exactly one.

LEMMA 9.17

Let $a_{i,j}$ and $b_{i,j}$ be nonzero complex numbers. The equality

$$\begin{aligned}&\pi_{n-1}(a_{1,1},\dots,a_{1,n-1})\cdots\pi_1(a_{n-1,1})\\ &\quad=\pi_{n-1}(b_{1,1},\dots,b_{1,n-1})\cdots\pi_1(b_{n-1,1})\end{aligned}\tag{9.24}$$

holds if and only if $a_{i,j}=b_{i,j}$ for all i,j.

Proof

For any $c_{i,j}$, the nth column of $\pi_{n-1}(c_{1,1},\dots,c_{1,n-1})\cdots\pi_1(c_{n-1,1})$ is equal to the nth column of $\pi_{n-1}(c_{1,1},\dots,c_{1,n-1})$, which equals

$$\Big(\prod_{i=1}^{n-1}c_{1,i},\prod_{i=2}^{n-1}c_{1,i},\dots,c_{1,n-1}\Big).$$

This proves that $a_{1,j}=b_{1,j}$ for all j. The result now follows by induction. □

PROPOSITION 9.18

For any assignment $c:\dot{\Delta}_n^2(\mathbb{Z})\to\mathbb{C}^*$, there is a unique Ptolemy assignment $c\in Pt_2^n$ whose Ptolemy coordinates are c_t.

Proof

We prove that the Ptolemy coordinates c'_t of $(N,q_{01}N,\alpha_{12}^0q_{02}N)$ equal c_t, where q_{01}, q_{02}, and α_{12}^0 are given in terms of the c_t's by (9.9) and (9.20). First, note that $c_t=c'_t$ if either t_1 or t_2 is 0, that is, if t is on one of the edges of Δ_n^2 containing the 0th vertex. Each of the other integral points t is the upper-right vertex of a unique diamond (r,s) of type 0. Let τ_k be the upper-right vertex of the kth diamond D_k in the sequence

$$(1,n-1),(1,n-2),\dots,(1,1),(2,n-2),\dots,(2,1),\dots,(n-1,1).\tag{9.25}$$

By Lemma 9.17, $d_{r,s}^{0\prime}=d_{r,s}^0$ for all diamonds (r,s) of type 0. It thus follows that if $c_t=c'_t$ for all but one of the vertices of a diamond D, then $c_t=c'_t$ for all vertices of D. In particular, $c'_{\tau_1}=c_{\tau_1}$. Suppose by induction that $c'_{\tau_i}=c_{\tau_i}$ for all $i<k$. Then $c'_t=c_t$, for all vertices of D_k except τ_k. Hence, we also have $c'_{\tau_k}=c_{\tau_k}$, completing the induction. □

PROPOSITION 9.19

For any assignment $c:\dot{\Delta}_n^3(\mathbb{Z})\to\mathbb{C}^*$ satisfying the Ptolemy relations, there is a unique Ptolemy assignment $c\in Pt_3^n$ whose Ptolemy coordinates are c_t.

Proof
Let c'_t be the Ptolemy coordinates of the tuple $(N, q_{01}N, \alpha^0_{12}q_{02}N, \alpha^0_{13}q_{03}N)$ defined from the c_t's by (9.9) and (9.20). We wish to prove that $c'_t = c_t$ for all t. Note that if, for some subsimplex $\Delta^3(\alpha)$, $c'_{\alpha_{ij}} = c_{\alpha_{ij}}$ for all but one of the 6 α_{ij}'s, then $c'_{\alpha_{ij}} = c_{\alpha_{ij}}$ holds for all α_{ij}. This is a direct consequence of the Ptolemy relations. By Proposition 9.18, $c'_t = c_t$, when either t_2 or t_3 is zero. Hence, for each $\alpha = (a_0, a_1, a_2, a_3)$ with $a_2 = a_3 = 0$, $c'_{\alpha_{ij}} = c_{\alpha_{ij}}$ except possibly when $(i,j) = (2,3)$. As explained above, $c'_{\alpha_{23}} = c_{\alpha_{23}}$ as well. Now suppose by induction that $c'_{\alpha_{ij}} = c_{\alpha_{ij}}$ for all α with $a_2 + a_3 < k$. Then $c'_{\alpha_{ij}} = c_{\alpha_{ij}}$ holds except possibly when $(i,j) = (2,3)$. Again, $c'_{\alpha_{23}} = c_{\alpha_{23}}$ must also hold, completing the induction. $\square$

A (G,N)-cocycle on M obviously determines a decorated representation (up to conjugation). The main results of this section can thus be summarized by the diagram below:

$$\begin{aligned}\big\{\text{Points in } P_n(K)\big\} &\longleftrightarrow \big\{\text{Generic } (G,N)\text{-cocycles on } M\big\}\\ &\longleftrightarrow \big\{\text{Generically decorated } (G,N)\text{-representations}\big\}. \qquad (9.26)\end{aligned}$$

Remark 9.20
We stress that the Ptolemy variety parameterizes decorated representations and *not* decorated representations up to equivalence. In particular, the dimension of $P_n(K)$ depends on the triangulation and may be very large if K has many interior vertices.

9.4. Obstruction cocycles and the $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy varieties
Suppose that n is even. The projection $G \to pG$ maps N isomorphically onto its image (also denoted by N), and by elementary obstruction theory (see, e.g., [23]), the obstruction to lifting a (pG,N)-representation ρ to a (G,N)-representation is a class in

$$H^2(M, \partial M; \mathbb{Z}/2\mathbb{Z}) = H^2(K; \mathbb{Z}/2\mathbb{Z}). \qquad (9.27)$$

We can represent it by an explicit cocycle in $Z^2(K;\mathbb{Z}/2\mathbb{Z})$ as follows: Pick any $(p\,\mathrm{SL}(n,\mathbb{C}), N)$-cocycle $\bar{\tau}$ on M representing ρ and a lift τ of $\bar{\tau}$ to a (G,N)-cochain. Each 2-cell of K corresponds to a hexagonal 2-cell of M, and the 2-cocycle $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$ taking a 2-cell to the product of the τ-labelings along the corresponding hexagonal 2-cell of M represents the obstruction class.

PROPOSITION 9.21
Suppose that the interior of M is a 1-cusped hyperbolic 3-manifold with finite volume. The obstruction class in $H^2(K;\mathbb{Z}/2\mathbb{Z})$ to lifting the geometric representation is nontrivial.

Proof

By a result of Calegari [4, Corollary 2.4], any lift of the geometric representation takes a longitude to an element in $\mathrm{SL}(2,\mathbb{C})$ with trace -2. This shows that no lift is boundary-unipotent, and so the obstruction class must be nontrivial. □

Proposition 9.4 also holds in $p\,\mathrm{SL}(n,\mathbb{C})$, and we thus have a one-to-one correspondence between generically decorated representations and (pG,N)-cocycles on M.

Definition 9.22

Let $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$. A lifted (pG,N) cocycle on M with obstruction cocycle σ is a generic (G,N)-assignment on M lifting a (pG,N)-cocycle on M such that the 2-cocycle on K obtained by taking products along hexagonal faces of M equals σ.

A 1-cochain $\eta \in C^1(K;\mathbb{Z}/2\mathbb{Z})$ acts on a lifted (pG,N)-cocycle τ by multiplying a long edge e by $\eta(e)$. Note that if τ has obstruction cocycle σ, then $\eta\tau$ has obstruction cocycle $\delta(\eta)\sigma$, where δ is the standard coboundary operator. Recall that there is a one-to-one correspondence between generic (G,N)-cocycles on M and points in the Ptolemy variety. We shall prove a similar result for pG.

We wish to define a coboundary action on pG-Ptolemy assignments (see Definition 5.11). Let c be a pG-Ptolemy assignment on Δ, and let $\eta_{ij} \in C^1(\Delta;\mathbb{Z}/2\mathbb{Z})$ be the cochain taking the edge ij to -1 and all other edges to 1. Define

$$\eta_{ij}c : \dot{\Delta}^3_n(\mathbb{Z}) \to \mathbb{C}^*, \qquad (\eta_{ij}c)_t = (-1)^{t_i t_j} c_t, \tag{9.28}$$

and extend in the natural way to define ηc for a pG-Ptolemy assignment c on K and $\eta \in C^1(K;\mathbb{Z}/2\mathbb{Z})$. A priori ηc is only an assignment of complex numbers to the integral points of the simplices of K. However, we have the following lemma.

LEMMA 9.23

If c is a pG-Ptolemy assignment on K with obstruction cocycle σ, then ηc is a pG-Ptolemy assignment on K with obstruction cocycle $\delta(\eta)\sigma$.

Proof

It is enough to prove this for a simplex Δ and for $\eta = \eta_{ij}$. Let $c' = \eta_{ij}c$. We assume for simplicity that $ij = 01$; the other cases are proved similarly. For any $\alpha = (a_0, a_1, a_2, a_3) \in \Delta^k_{n-2}(\mathbb{Z})$, we then have

$$\begin{aligned} &c'_{\alpha_{03}}c'_{\alpha_{12}} + c'_{\alpha_{01}}c'_{\alpha_{23}} - c'_{\alpha_{02}}c'_{\alpha_{13}} \\ &\quad = (-1)^{a_0+a_1}(c_{\alpha_{03}}c_{\alpha_{12}} - c_{\alpha_{01}}c_{\alpha_{23}} - c_{\alpha_{02}}c_{\alpha_{13}}). \end{aligned} \tag{9.29}$$

Let $\tau = \delta(\eta_{01})$. Since $\delta(\eta_{01})_2 = \delta(\eta_{01})_3 = -1$ and $\delta(\eta_{01})_0 = 1$, (9.29) implies that

$$\tau_2\tau_3 c'_{\alpha_{03}} c'_{\alpha_{12}} + \tau_0\tau_3 c'_{\alpha_{03}} c'_{\alpha_{01}} c'_{\alpha_{23}} = c'_{\alpha_{02}} c'_{\alpha_{13}}, \tag{9.30}$$

as desired. □

Definition 9.24
The diamond coordinates of a $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignment with obstruction cocycle σ are defined as in Definition 9.13, but multiplied by the sign (provided by σ) of the face.

Note that, for $\eta \in C^1(K;\mathbb{Z}/2/\mathbb{Z})$, the diamond coordinates of c and ηc are identical.

PROPOSITION 9.25
For any $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$, there is a one-to-one correspondence between $p\,\mathrm{SL}(n,\mathbb{C})$-Ptolemy assignments on K with obstruction cocycle σ and lifted $(p\,\mathrm{SL}(n,\mathbb{C}), N)$-cocycles on M with obstruction cocycle σ. The correspondence preserves the coboundary actions.

Proof
It is enough to prove this for a simplex Δ. For a pG-Ptolemy assignment c on Δ with obstruction cocycle $\sigma \in Z^2(\Delta;\mathbb{Z}/2\mathbb{Z})$, define a cochain τ on $\overline{\Delta}$ by the formulas (9.9) and (9.20) using the σ-modified diamond coordinates (Definition 9.24). Let $\eta \in C^1(\Delta;\mathbb{Z}/2\mathbb{Z})$ be such that $\delta\eta = \sigma$, where δ is the standard coboundary map. By Lemma 9.23, ηc satisfies the $\mathrm{SL}(n,\mathbb{C})$-Ptolemy relations (5.4) and hence corresponds to an $(\mathrm{SL}(n,\mathbb{C}), N)$-cocycle τ'. Since the diamond coordinates of c and ηc are the same, the short edges of τ' agree with those of τ and the long edges differ from those of τ by η. This proves that τ is a lifted (pG, N)-cocycle with obstruction cocycle σ. The inductive arguments of Propositions 9.18 and 9.19 show that this is a one-to-one correspondence. The fact that the actions by coboundaries correspond is immediate from the construction. □

COROLLARY 9.26
Let $\sigma \in Z^2(K;\mathbb{Z}/2\mathbb{Z})$. There is an algebraic variety $P_n^\sigma(K)$ of generically decorated boundary-unipotent representations $\rho : \pi_1(M) \to p\,\mathrm{SL}(n,\mathbb{C})$ whose obstruction class to lifting to $\mathrm{SL}(n,\mathbb{C})$ is represented by σ. Up to canonical isomorphism, the variety $P_n^\sigma(K)$ depends only on the cohomology class of σ.

Proof
This follows immediately from Proposition 9.25. □

Note that the canonical isomorphisms in Corollary 9.26 respect the extended Bloch group element. This follows from the pG variant of Proposition 7.7. The analogue of (9.26) is

$$\begin{aligned}&\{\text{Points in } P_n^\sigma(K)\}\\ &\quad\longleftrightarrow \{\text{Lifted } (pG,N)\text{-cocycles on } M \text{ with obstruction cocycle } \sigma\}\\ &\quad\xrightarrow{k:1}\!\!\!\!\rightarrow \{\text{Generically decorated } (pG,N)\text{-representations with}\\ &\qquad\quad \text{obstruction cocycle } \sigma\},\end{aligned}\tag{9.31}$$

where k is the number of lifts, that is, that $k = |Z^1(K;\mathbb{Z}/2\mathbb{Z})|$.

9.5. Proof of Theorems 1.3, 1.12, and 1.7

Let $\mathcal{R} : P_n(K) \to R_{G,N}(M)$ be the composition of the map in (9.26) with the forgetful map ignoring the decoration. The fact that λ has image in $\widehat{\mathcal{B}}(\mathbb{C})$ follows from Proposition 6.12, and commutativity of (1.11) follows from Remark 8.4. The fact that $\mathcal{R}$ is surjective if K is sufficiently fine follows from Proposition 5.4. This concludes the proof of Theorem 1.3. The first part of Theorem 1.12 is proved similarly, and the last part follows from Theorem 11.7 below. The first statement of Theorem 1.7 follows from the definition of $\mathcal{R}$. The second statement follows from the fact that if ρ is boundary nondegenerate the only freedom in choosing a decoration is the diagonal action. Finally, the third statement is proved in Corollary 7.9.

10. Examples

In the examples below, all computations of Ptolemy varieties are exact, whereas the computations of complex volume are numerical with at least 50-digit precision.

Example 10.1 (The 5_2-knot complement)

Consider the 3-cycle K obtained from the simplices in Figure 9 by identifying the faces via the unique simplicial attaching maps preserving the arrows. The space obtained from K by removing the 0-cell is homeomorphic to the complement of the 5_2-knot, as can be verified by SnapPy [7].

Labeling the Ptolemy coordinates as in Figure 9, the Ptolemy variety for $n = 3$ is given by the equations

$$\begin{array}{lll} a_0x_3 + b_0x_1 = b_0x_2, & a_0y_3 + a_0x_0 = c_0y_2, & a_0x_2 + b_0y_2 = a_0x_1, \\ x_2c_0 + b_1x_0 = x_3a_0, & y_2b_0 + a_1x_3 = y_3b_0, & x_1a_0 + b_1y_3 = x_2c_0, \\ x_1c_1 + x_3c_0 = b_1x_0, & x_0b_1 + y_3c_0 = c_1x_3, & y_2a_1 + x_2b_0 = a_1y_3, \\ a_1x_0 + x_2c_1 = x_1a_1, & a_1x_3 + y_2c_1 = x_0b_1, & a_1y_3 + x_1b_1 = y_2c_1 \end{array}\tag{10.1}$$

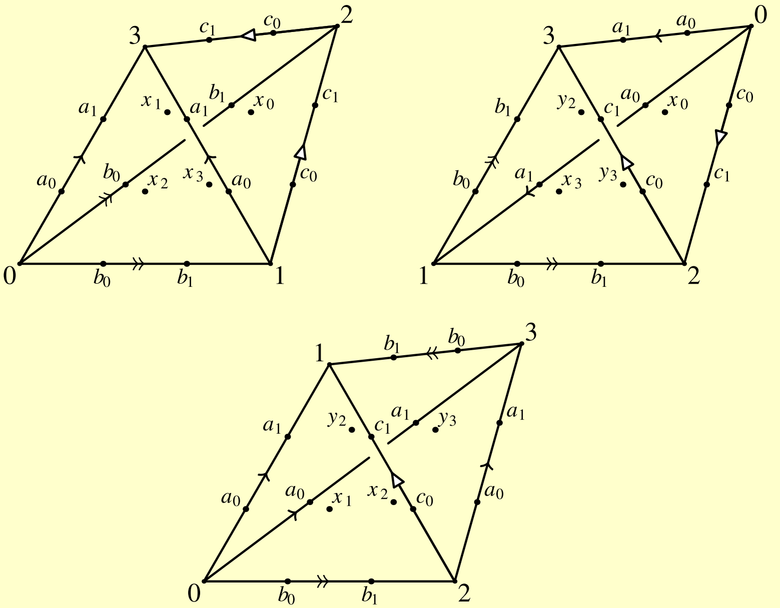


Figure 9. A 3-cycle structure on the 5_2 knot complement, and Ptolemy coordinates for $n = 3$.

together with an extra equation (involving an additional variable t)

$$a_0 a_1 b_0 b_1 c_0 c_1 x_0 x_1 x_2 x_3 y_2 y_3 t = 1, \tag{10.2}$$

making sure that all Ptolemy coordinates are nonzero. By (5.7), a diagonal matrix $\operatorname{diag}(x, y, z)$ acts by multiplying a Ptolemy coordinate on an edge by $x^2 y$ and a Ptolemy coordinate on a face by x^3. Since we are not interested in the particular decoration, we may thus assume, for example, that $a_0 = y_3 = 1$. Using Magma [3], one finds that the Ptolemy variety, after setting $a_0 = y_3 = 1$, has three 0-dimensional components with 3, 4, and 6 points, respectively. One of these is given by

$$\begin{gathered}
a_0 = a_1 = y_3 = 1, \qquad x_1 = -1, \qquad c_0 = c_1 = x_0^2 + 2x_0 + 1, \\
y_2 = x_0^2 + 2 = -x_2, \qquad x_3 = -x_0^2 - x_0 - 1, \\
x_0^3 + x_0^2 + 2x_0 + 1 = 0.
\end{gathered} \tag{10.3}$$

Thus, this component gives rise to three representations, one for each solution to $x_0^3 + x_0^2 + 2x_0 + 1 = 0$. Using the fact that $R(\lambda(c)) = i \operatorname{Vol}_{\mathbb{C}}(\rho)$, the complex volumes of these can be computed to be

$$\begin{gathered}
0.0 - 4.453818209\ldots i \in \mathbb{C}/4\pi^2 i\mathbb{Z}, \\
\pm 11.31248835\ldots + 12.09651350\ldots i \in \mathbb{C}/4\pi^2 i\mathbb{Z}
\end{gathered} \tag{10.4}$$

corresponding to the values $x_0 = -0.5698\dots$ and $x_0 = -0.2150 \mp 1.3071\dots i$, respectively.

In [27, Section 6], the complex volumes of the Galois conjugates of the geometric representation are computed to be

$$\begin{aligned} 0.0 - 1.113454552\dots i &\in \mathbb{C}/\pi^2 i\mathbb{Z}, \\ \pm 2.828122088\dots + 3.024128376\dots i &\in \mathbb{C}/\pi^2 i\mathbb{Z}. \end{aligned} \tag{10.5}$$

Notice that (10.4) is (approximately) 4 times (10.5). It thus follows from Theorem 1.10 that the representations given by (10.3) are ϕ_3 composed with the geometric component of PSL$(2, \mathbb{C})$-representations and that the factor of 4 is exact.

Another component is given by

$$\begin{aligned} & a_0 = a_1 = y_3 = 1, \qquad x_1 = -1, \qquad b_1 = -x_0, \\ & b_0 = 1/4x_0^3 - 1/4x_0^2 + 3/4x_0 - 1/2, \\ & c_0 = c_1 = 1/4x_0^3 - 1/4x_0^2 - 1/4x_0 + 1/2, \\ & y_2 = -x_2 = 1/4x_0^3 + 3/4x_0^2 + 7/4x_0 + 3/2, \qquad x_3 = -x_0^2 - x_0 - 1, \\ & \qquad\qquad x_0^4 + x_0^3 + x_0^2 - 4x_0 - 4 = 0. \end{aligned} \tag{10.6}$$

In this case, there are two distinct complex volumes given by

$$\begin{aligned} 0.0 + 2.631894506\dots i &= \frac{4}{15}\pi^2 i \in \mathbb{C}/4\pi^2 i\mathbb{Z}, \\ 0.0 + 10.527578027\dots i &= \frac{16}{15}\pi^2 i \in \mathbb{C}/4\pi^2 i\mathbb{Z}. \end{aligned} \tag{10.7}$$

The third component has somewhat larger coefficients, but after introducing a variable u with $u^6 + 5u^4 + 8u^2 - 2u + 1 = 0$, the defining equations simplify to

$$\begin{aligned} & a_0 = y_3 = 1, \qquad a_1 = 1/4u^5 + 1/4u^4 + 5/4u^3 + 1/2u^2 + 2u - 3/4, \\ & b_0 = b_1 = -1/4u^4 - 3/4u^2 - 1/4u - 3/4, \\ & c_1 = -1/4u^5 - 3/4u^3 - 1/4u^2 - 3/4u, \\ & c_0 = 1/2u^5 + 9/4u^3 + 1/4u^2 + 7/2u - 1/4, \\ & y_2 = -8/17u^5 - 1/34u^4 - 79/34u^3 - 3/17u^2 - 105/34u + 26/17, \\ & x_3 = 1/17u^5 - 1/17u^4 + 6/17u^3 - 6/17u^2 + 14/17u - 16/17, \\ & x_2 = 9/34u^5 + 4/17u^4 + 37/34u^3 + 31/34u^2 + 75/34u + 13/17, \\ & x_1 = 8/17u^5 + 1/34u^4 + 79/34u^3 + 3/17u^2 + 139/34u - 9/17, \end{aligned} \tag{10.8}$$

$$x_0 = 15/34u^5 + 1/17u^4 + 73/34u^3 + 29/34u^2 + 125/34u - 1/17,$$
$$u^6 + 5u^4 + 8u^2 - 2u + 1 = 0.$$

In this case, there are three distinct complex volumes:

$$0.0 + 1.241598704\ldots i, \qquad \pm 6.332666642\ldots + 1.024134714\ldots i. \tag{10.9}$$

According to Conjecture 1.20, $6.33\ldots + 1.02\ldots i$ should (up to rational multiples of $\pi^2 i$) be an integral linear combination of complex volumes of hyperbolic manifolds. Using, for example, Snap [17], one checks that the complex volume of the manifold $m034$ is given by

$$3.166333321\ldots + 2.157001424\ldots i, \tag{10.10}$$

and we have

$$6.332666642\ldots + 1.024134714\ldots i = 2\,\mathrm{Vol}_{\mathbb{C}}(m034) - \frac{1}{3}\pi^2 i \in \mathbb{C}/4\pi^2 i\mathbb{Z}. \tag{10.11}$$

Example 10.2 (The figure-eight knot complement)
Let K be the 3-cycle in Figure 10. Then $M = M(K)$ is the figure-eight knot complement, and $H^2(K;\mathbb{Z}/2\mathbb{Z}) = H^2(M,\partial M;\mathbb{Z}/2\mathbb{Z}) = \mathbb{Z}/2\mathbb{Z}$.

For the trivial obstruction class, the Ptolemy variety for $n = 2$ is given by

$$yx + y^2 = x^2, \qquad xy + x^2 = y^2, \tag{10.12}$$

and is thus empty since x and y are nonzero. In fact, the only boundary-unipotent representations in $\mathrm{SL}(2,\mathbb{C})$ are reducible, so this is not surprising. The nontrivial obstruction class can be represented by the cocycle indicated in Figure 10, and the Ptolemy variety is given by

$$yx - y^2 = x^2, \qquad xy - x^2 = y^2. \tag{10.13}$$

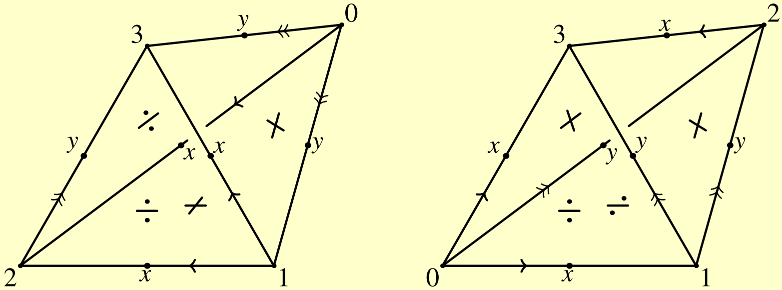


Figure 10. A 3-cycle structure on the figure-eight knot complement and Ptolemy coordinates for $n = 2$. The signs indicate the nontrivial second $\mathbb{Z}/2\mathbb{Z}$ cohomology class.

As in Example 10.1, we may assume $y = 1$. Hence, the Ptolemy variety detects two (complex conjugate) representations corresponding to the solutions to $x^2 - x + 1 = 0$. The extended Bloch group elements are

$$-(-\widetilde{x}, -2\widetilde{x}) + (\widetilde{x}, 2\widetilde{x}) \in \widehat{\mathcal{B}}(\mathbb{C})_{\mathrm{PSL}}, \tag{10.14}$$

with complex volume

$$\pm 2.029883212\ldots + 0.0i. \tag{10.15}$$

We thus recover the well-known complex volume of the figure-eight knot complement.

For $n = 3$, similar calculations as those in Example 10.1 show that the Ptolemy variety detects three 0-dimensional components, but the only one with nonzero volume is the one induced by the geometric representation. For $n = 4$, lots of new complex volumes emerge. For the trivial obstruction class, the nonzero complex volumes are

$$\pm 7.327724753\ldots + 0.0i = 2\,\mathrm{Vol}_{\mathbb{C}}(5_1^2) + \pi^2 i/4, \tag{10.16}$$

where the manifold 5_1^2 is the Whitehead link complement. For the nontrivial obstruction class, the complex volumes are

$$\begin{gathered}
\pm 20.29883212\ldots + 0.0i = 10\,\mathrm{Vol}_{\mathbb{C}}(4_1) \in \mathbb{C}/\pi^2 i\mathbb{Z}, \\
\pm 4.260549384\ldots \pm 0.136128165\ldots i, \\
\pm 3.230859569\ldots + 0.0i, \\
\pm 8.355502146\ldots + 2.428571615\ldots i = \mathrm{Vol}_{\mathbb{C}}(-9_{15}^3) + 2\pi^2 i/3, \\
\pm 3.276320849\ldots + 9.908433886\ldots i.
\end{gathered} \tag{10.17}$$

Remark 10.3

When $n = 2$, examples of Conjecture 1.20 are abundant. For example, for the 10_{155}-knot complement (10 simplices), the volumes of the representations detected by the Ptolemy variety are (numerically)

$$\begin{aligned}
&\mathrm{Vol}\big(m032(6,1)\big), && 2\,\mathrm{Vol}(4_1), \\
&3\,\mathrm{Vol}(10_{155}) - 4\,\mathrm{Vol}(v3461), && \mathrm{Vol}(10_{155}).
\end{aligned} \tag{10.18}$$

Remark 10.4

For the hyperbolic census manifolds, most of the components of the reduced Ptolemy varieties tend to be 0-dimensional. By a result of Menal-Ferrer and Porti [18], the composition of the geometric representation with ϕ_n is isolated among boundary-unipotent $p\,\mathrm{SL}(n,\mathbb{C})$-representations. Higher-dimensional components also occur (rarely for $n = 2$, but quite often for $n > 2$); but, as mentioned earlier, the complex volume is constant on components.

Remark 10.5

If the face pairings do not respect the vertex orderings, then one can still define a Ptolemy variety by introducing more signs. See [14] for details.

Remark 10.6

The fact that the reduced Ptolemy varieties $P_n(K)_{\mathrm{red}}$ are given by setting some of the variables (chosen appropriately) equal to 1 is proved in [15].

11. The irreducible representations of $\mathrm{SL}(2,\mathbb{C})$

Let $\phi_n : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(n,\mathbb{C})$ denote the canonical irreducible representation. It is induced by the Lie algebra homomorphism $\mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{sl}(n,\mathbb{C})$ given by

$$\begin{aligned}
\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} &\mapsto \mathrm{diag}^+(n-1,\dots,1), \\
\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &\mapsto \mathrm{diag}^-(1,\dots,n-1), \\
\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} &\mapsto \mathrm{diag}(n-1,n-3,\dots,-n+1),
\end{aligned} \tag{11.1}$$

where $\mathrm{diag}^+(v)$ and $\mathrm{diag}^-(v)$ denote matrices whose first upper (resp., lower) diagonal is v and all other entries are zero. One has

$$\phi_n\left(\begin{bmatrix} 0 & -a^{-1} \\ a & 0 \end{bmatrix}\right) = q\big(a^{n-1}, -a^{n-3}, \dots, (-1)^{n-1}a^{-(n-1)}\big), \tag{11.2}$$

$$\phi_n\left(\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}\right) = \pi_{n-1}(x,\dots,x)\pi_{n-2}(x,\dots,x)\cdots\pi_1(x). \tag{11.3}$$

PROPOSITION 11.1

Let c be a Ptolemy assignment on Δ_2^3, and let τ denote the corresponding cocycle. The Ptolemy assignment corresponding to $\phi_n(\tau)$ is given by

$$\phi_n(c) : \dot{\Delta}_n^3(\mathbb{Z}) \to \mathbb{C}^*, \qquad t \mapsto \phi_n(c)_t = \prod_{i<j} c_{ij}^{t_i t_j}. \tag{11.4}$$

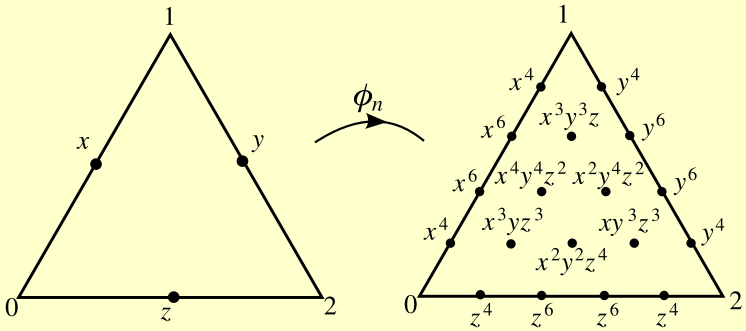


Figure 11. ϕ_n acting on Ptolemy assignments.

Proof
Let $\alpha = (a_0, \dots, a_3) \in \Delta^3_{n-2}(\mathbb{Z})$. Letting $k_\alpha = \prod_{i<j} c_{ij}^{a_i a_j}$ and $l_\alpha = \prod_{i<j} c_{ij}^{a_i + a_j}$, we have

$$\begin{aligned} \phi_n(c)_{\alpha_{03}} \phi_n(c)_{\alpha_{12}} &= k_\alpha^2 l_\alpha c_{03} c_{12}, \\ \phi_n(c)_{\alpha_{01}} \phi_n(c)_{\alpha_{23}} &= k_\alpha^2 l_\alpha c_{01} c_{23}, \\ \phi_n(c)_{\alpha_{02}} \phi_n(c)_{\alpha_{13}} &= k_\alpha^2 l_\alpha c_{02} c_{13}. \end{aligned} \tag{11.5}$$

Hence, the appropriate Ptolemy relations are satisfied. The long and short edges of the cocycle corresponding to $\phi_n(c)$ are given by (9.9) and (9.20), and we must prove that these agree with those of $\phi_n(\tau)$. For the long edges, this follows immediately from (11.2). For the short edges, an easy computation shows that all the diamond coordinates of a face are equal, and equal to the corresponding diamond coordinate of c. For example, the type 1 diamond coordinate on face 3 whose left vertex is $t = (t_0, t_1, t_2, 0)$ is given by

$$\begin{aligned} &\frac{\phi_n(c)_{t+(0,-1,1,0)} \phi_n(c)_{t+(-1,1,0,0)}}{\phi_n(c)_t \phi_n(c)_{t+(-1,0,1,0)}} \\ &\quad = \frac{c_{01}^{t_0(t_1-1)} c_{02}^{t_0(t_2+1)} c_{12}^{(t_1-1)(t_2+1)} c_{01}^{(t_0-1)(t_1+1)} c_{02}^{(t_0-1)t_2} c_{12}^{(t_1+1)t_2}}{c_{01}^{t_0 t_1} c_{02}^{t_0 t_2} c_{12}^{t_1 t_2} c_{01}^{(t_0-1)t_1} c_{02}^{(t_0-1)(t_2+1)} c_{12}^{t_1(t_2+1)}} \\ &\quad = \frac{c_{02}}{c_{01} c_{12}}, \end{aligned} \tag{11.6}$$

which is a diamond coordinate for c. By (11.3), the short edges thus agree with those of $\phi_n(\tau)$, proving the result. $\square$

COROLLARY 11.2
If a representation $\rho: \pi_1(M) \to \mathrm{PSL}(2,\mathbb{C})$ *is detected by* $P_2^\sigma(K)$*, then* $\phi_{2k+1} \circ \rho$ *is detected by* $P_{2k+1}(K)$ *and* $\phi_{2k} \circ \rho$ *is detected by* $P_{2k}^\sigma(K)$.

THEOREM 11.3
Let ρ *be a boundary-unipotent representation in* $\mathrm{SL}(2,\mathbb{C})$ *or* $\mathrm{PSL}(2,\mathbb{C})$. *The extended Bloch group element of* $\phi_n \circ \rho$ *is* $\binom{n+1}{3}$ *times that of* ρ. *In fact, the shapes of all subsimplices are equal.*

Proof
By refining the triangulation if necessary, we may represent ρ by a Ptolemy assignment c on K. Then $\phi = \phi_n(c)$ is a Ptolemy assignment representing $\phi_n \circ \rho$, and the extended Bloch group element of $\phi_n \circ \rho$ is given by

$$\left[\phi_n(\rho)\right] = \sum_i \epsilon_i \sum_{\alpha \in \Delta^3_{n-2}(\mathbb{Z})} (\widetilde{\phi}^i_{\alpha_{03}} + \widetilde{\phi}^i_{\alpha_{12}} - \widetilde{\phi}^i_{\alpha_{02}} - \widetilde{\phi}^i_{\alpha_{13}}, \widetilde{\phi}^i_{\alpha_{01}} + \widetilde{\phi}^i_{\alpha_{23}} - \widetilde{\phi}^i_{\alpha_{02}} - \widetilde{\phi}^i_{\alpha_{13}}). \tag{11.7}$$

By Proposition 7.7, we may choose the logarithms independently as long as we use the same logarithm for identified points. Defining $\widetilde{\phi}^i_t = \sum_{j<k} t_j t_k \widetilde{c}^{\,i}_{jk}$, we see that

$$\begin{aligned}(\widetilde{\phi}^i_{\alpha_{03}} + \widetilde{\phi}^i_{\alpha_{12}} - \widetilde{\phi}^i_{\alpha_{02}} - \widetilde{\phi}^i_{\alpha_{13}}, \widetilde{\phi}^i_{\alpha_{01}} + \widetilde{\phi}^i_{\alpha_{23}} - \widetilde{\phi}^i_{\alpha_{02}} - \widetilde{\phi}^i_{\alpha_{13}})\\ = (\widetilde{c}_{03} + \widetilde{c}_{12} - \widetilde{c}_{02} - \widetilde{c}_{13}, \widetilde{c}_{01} + \widetilde{c}_{23} - \widetilde{c}_{02} - \widetilde{c}_{13}),\end{aligned} \tag{11.8}$$

which means that the flattenings assigned to each subsimplex of Δ^i_n are equal. By Lemma 5.6, $|\Delta^3_{n-2}(\mathbb{Z})| = \binom{n+1}{3}$, and the result follows. □

11.1. Essential edges

Definition 11.4
An edge of K is *essential* if the lifts to L have distinct end points.

Note that an edge may be essential even though it is homotopically trivial in K. Let $L^{(0)}$ denote the 0-skeleton of L.

LEMMA 11.5
Let ρ *be a representation in* $\mathrm{SL}(2,\mathbb{C})$ *or* $\mathrm{PSL}(2,\mathbb{C})$. *A decoration of* ρ *determines a* ρ*-equivariant map*

$$D: L^{(0)} \to \partial\overline{\mathbb{H}}^3 = \mathbb{C} \cup \{\infty\}, \qquad e_i \mapsto g_i \infty. \tag{11.9}$$

Every such map comes from a decoration, and the decoration is generic if and only if the vertices of each simplex of L map to distinct points in $\mathbb{C} \cup \{\infty\}$.

Proof
Equivariance of (11.9) follows from the definition of a decoration. A ρ-equivariant map $D : L^{(0)} \to \mathbb{C} \cup \{\infty\}$ is uniquely determined by its image of lifts $\widetilde{e_i} \in L$ of the 0-cells e_i of K. Picking g_i such that $g_i \infty = D(\widetilde{e_i})$, we define a decoration by assigning the coset $g_i N$ to $\widetilde{e_i}$. The last statement follows from the fact that $\det(g_1 e_1, g_2 e_1) = 0$ if and only if $g_1 \infty = g_2 \infty$. □

In the following, we assume that the interior of M is a cusped hyperbolic 3-manifold $\mathbb{H}^3/\Gamma$ with finite volume.

PROPOSITION 11.6
If all edges of K are essential, then the geometric representation has a generic decoration.

Proof
We identify $\pi_1(M)$ with $\Gamma \subset \mathrm{PSL}(2,\mathbb{C})$. Each cusp of M determines a Γ-orbit of points in $\partial\overline{\mathbb{H}^3}$, and these orbits are distinct (if two orbits intersected, then they would be identical, and thus correspond to the same cusp). Each vertex of L corresponds to either a cusp of M or an interior point of M. Accordingly, we have $L^{(0)} = L^{(0)}_{\mathrm{cusp}} \cup L^{(0)}_{\mathrm{int}}$. The stabilizer of a point in $L^{(0)}_{\mathrm{cusp}}$ is a parabolic subgroup of $\mathrm{PSL}(2,\mathbb{C})$ and thus fixes a unique point in $\mathbb{C} \cup \{\infty\}$. We thus have an equivariant map $D : L^{(0)}_{\mathrm{cusp}} \to \mathbb{C} \cup \{\infty\}$ taking a vertex v to the fixed point in $\partial\mathbb{H}^3$ of $\mathrm{Stab}(v) \subset \mathrm{PSL}(2,\mathbb{C})$. Let e_1 and e_2 be points in $L^{(0)}_{\mathrm{cusp}}$ connected by an edge. Since all edges of K are essential, $e_1 \neq e_2$. Since the Γ-orbits of different cusps are distinct, it follows that $D(e_1) \neq D(e_2)$ if e_1 and e_2 correspond to different cusps. If e_1 and e_2 correspond to the same cusp, there exists an element in Γ taking e_1 to e_2. Since only peripheral elements (i.e., cusp stabilizers) have fixed points in $\mathbb{C} \cup \{\infty\}$, it follows that $D(e_1) \neq D(e_2)$. We extend D to $L^{(0)}$ by choosing any equivariant map $L^{(0)}_{\mathrm{int}} \to \mathbb{C} \cup \{\infty\}$. Since such a map is uniquely determined by finitely many values (which may be chosen freely), we can pick the extension so that the vertices of each simplex map to distinct points. This proves the result. □

THEOREM 11.7
Suppose that all edges of K are essential. The representation $\phi_n \circ \rho_{\mathrm{geo}}$ is detected by $P_n(K)$ if n is odd and by $P_n^{\sigma_{\mathrm{geo}}}(K)$ if n is even.

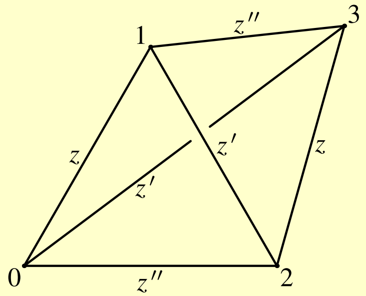


Figure 12. Assigning cross-ratio parameters to the edges of Δ_i. By definition, $z' = \frac{1}{1-z}$ and $z'' = 1 - \frac{1}{z}$.

Proof
By Proposition 11.6, $P_2^{\sigma_{\text{geo}}}(K)$ detects ρ_{geo}. The result now follows from Corollary 11.2. □

Remark 11.8
The census triangulations all have essential edges.

12. Gluing equations and Ptolemy assignments

In this section, we discuss the relation between Ptolemy assignments and solutions to the gluing equations. The latter were invented by Thurston [26] to explicitly compute the hyperbolic structure (and its deformations) of a triangulated hyperbolic manifold and used effectively in [17], [21], and [7]. The gluing equations make sense for any 3-cycle. They are defined by assigning a *cross-ratio* $z_i \in \mathbb{C}\backslash\{0,1\}$ to each simplex Δ_i of K. Given these, we assign cross-ratio parameters to the edges of Δ_i as in Figure 12.

There is a gluing equation for each edge E in K and each generator γ of the fundamental group of each boundary component of M. These are given by

$$\prod_{e \mapsto E} z(e)^{\epsilon_i(e)} = 1, \qquad \prod_{\gamma \text{ passes } e} z(e)^{\epsilon_i(e)} = 1. \tag{12.1}$$

Here $z(e)$ denotes the cross-ratio parameter assigned to e, and $\epsilon_i(e) = \epsilon_i$ if e is an edge of Δ_i. It follows that the set of assignments $\Delta_i \mapsto z_i \in \mathbb{C}\backslash\{0,1\}$ satisfying the gluing equations (12.1) is an algebraic set $V(K)$.

LEMMA 12.1
For every point $\{z_i\} \in V(K)$, there is a map $D : L^{(0)} \to \mathbb{C} \cup \{\infty\}$ such that if $\widetilde{\Delta}_i$ is a lift of Δ_i with vertices $e_1, \dots, e_3$ in L, the cross-ratio of the ideal simplex with vertices $D(e_1), \dots, D(e_3)$ is z_i. It is unique up to multiplication by an element in $\mathrm{PSL}(2,\mathbb{C})$.

Moreover, there is a unique (up to conjugation) boundary-unipotent representation $\pi_1(M) \to \mathrm{PSL}(2,\mathbb{C})$ *such that* D *is* ρ*-equivariant.*

Proof

Pick a fundamental domain F for K in L. Pick a simplex Δ in F and define D by mapping the first three vertices of Δ to 0, ∞ and 1. The map D is now uniquely determined by the cross-ratios. The fundamental group of M has a presentation with a generator for each face pairing of F. The second statement thus follows from the fact that $\mathrm{PSL}(2,\mathbb{C})$ is 3-transitive. We leave the details to the reader. □

Given a Ptolemy assignment on K, we assign the cross-ratio $z_i = \frac{c^i_{03}c^i_{12}}{c^i_{02}c^i_{13}}$ to Δ_i. Note that the Ptolemy relations imply that the cross-ratio parameters are given by

$$z_i = \frac{c^i_{03}c^i_{12}}{c^i_{02}c^i_{13}}, \qquad z'_i = \frac{c^i_{02}c^i_{13}}{c^i_{01}c^i_{23}}, \qquad z''_i = -\frac{c^i_{01}c^i_{23}}{c^i_{03}c^i_{12}}. \tag{12.2}$$

THEOREM 12.2

There is a surjective regular map

$$\coprod_{\sigma \in H^2(K;\mathbb{Z}/2\mathbb{Z})} P_2^\sigma(K) \to V(K), \qquad c \mapsto \left\{ z_i = \frac{c^i_{03}c^i_{12}}{c^i_{02}c^i_{13}} \right\}. \tag{12.3}$$

The fibers are disjoint copies of $(\mathbb{C}^*)^h$*, where* h *is the number of* 0*-cells of* K*.*

Proof

By a simple cancellation argument (as in the proof of Zickert [27, Theorem 6.5]), the gluing equations would be satisfied if the formula (12.2) for z''_i did not have the minus sign. The minus sign appears whenever the edge is 02 or 13. As explained in the proof of Proposition 7.7, any curve passes these an even number of times. It thus follows that the cross-ratios satisfy the gluing equations. Surjectivity follows from Lemma 11.5, and the fact that fibers are $(\mathbb{C}^*)^h$ follows from the fact that $g_1\infty = g_2\infty$ if and only if $g_1N = g_2dN$ for a unique diagonal matrix d. □

Remark 12.3

Gluing equation varieties for $n > 2$ are studied in [14].

Acknowledgments. The authors wish to thank Ian Agol, Johan Dupont, Matthias Goerner, and Walter Neumann for stimulating conversations and the referees for valuable comments and corrections. We are particularly grateful to Matthias Goerner for a computer implementation of our formulas and for supplying our theory with computational data for more than 20,000 manifolds. The software has been incorporated into

SnapPy [7], and computational data can be found at http://curve.unhyperbolic.org/database.html. Zickert's work was supported by National Science Foundation (NSF) grants DMS-1007054 and DMS-1309088. Garoufalidis's work was supported by NSF grant DMS-1105678. Thurston's work was supported by NSF grant DMS-1008049 and by a Sloan fellowship.

References

[1] A. C. AITKEN. *Determinants and Matrices*, Oliver and Boyd, Edinburgh, 1939. *(2134)*

[2] N. BERGERON, E. FALBEL, and A. GUILLOUX, *Tetrahedra of flags, volume and homology of* SL(3), Geom. Topol. **18** 1911–1971. MR 3268771. DOI 10.2140/gt.2014.18.1911. *(2108)*

[3] W. BOSMA, J. CANNON, and C. PLAYOUST, *The Magma algebra system, I: The user language*, J. Symbolic Comput. 24 (1997), 235–265. MR 1484478. DOI 10.1006/jsco.1996.0125. *(2102, 2148)*

[4] D. CALEGARI, *Real places and torus bundles*, Geom. Dedicata **118** (2006), 209–227. MR 2239457. DOI 10.1007/s10711-005-9037-9. *(2101, 2145)*

[5] J. CHEEGER and J. SIMONS, "Differential characters and geometric invariants" in *Geometry and Topology (College Park, Md., 1983/84)*, Lecture Notes in Math. **1167**, Springer, Berlin, 50–80, 1985. MR 0827262. DOI 10.1007/BFb0075216. *(2100, 2109, 2134)*

[6] S. S. CHERN and J. SIMONS, *Characteristic forms and geometric invariants*, Ann. of Math. (2) **99** (1974), 48–69. MR 0353327. *(2100, 2109)*

[7] M. CULLER, N. M. DUNFIELD, and J. R. WEEKS, *SnapPy, a computer program for studying the geometry and topology of 3-manifolds*, http://www.math.uic.edu/t3m/SnapPy. *(2102, 2147, 2156, 2158)*

[8] T. DIMOFTE, M. GABELLA, and A. B. GONCHAROV, *K-decompositions and 3d gauge theories*, preprint, arXiv:1301.0192v1 [hep-th]. *(2108)*

[9] J. DUPONT, R. HAIN, and S. ZUCKER, "Regulators and characteristic classes of flat bundles" in *The Arithmetic and Geometry of Algebraic Cycles (Banff, Alb., 1998)*, CRM Proc. Lecture Notes **24**, Amer. Math. Soc., Providence, 2000, 47–92. MR 1736876. *(2110)*

[10] J. L. DUPONT and F. W. KAMBER, *On a generalization of Cheeger–Chern–Simons classes*, Illinois J. Math. **34** (1990), 221–255. MR 1046564. *(2110)*

[11] E. FALBEL, *A spherical CR structure on the complement of the figure eight knot with discrete holonomy*, J. Differential Geom. **79** (2008), 69–110. MR 2401419. *(2108)*

[12] E. FALBEL and Q. WANG, *A combinatorial invariant for spherical CR structures*, Asian J. Math. **17** (2013), 391–422. MR 3119793. DOI 10.4310/AJM.2013.v17.n3.a1. *(2108)*

[13] V. FOCK and A. GONCHAROV, *Moduli spaces of local systems and higher Teichmüller theory*, Publ. Math. Inst. Hautes Études Sci. **103** (2006), 1–211. MR 2233852. DOI 10.1007/s10240-006-0039-4. *(2103, 2108, 2111, 2119)*

[14] S. GAROUFALIDIS, M. GOERNER, and C. K. ZICKERT, *Gluing equations for* PGL(n, $\mathbb{C}$)*-representations of 3-manifolds*, Algebr. Geom. Top. **15** (2015), 565–622. *(2108, 2122, 2152, 2157)*

[15] ———, *The Ptolemy field of 3-manifold representations*, Algebr. Geom. Top. **15** (2015), 371–397. *(2152)*

[16] S. GOETTE and C. ZICKERT, *The extended Bloch group and the Cheeger–Chern–Simons class*, Geom. Topol. **11** (2007), 1623–1635. MR 2350461. DOI 10.2140/gt.2007.11.1623. *(2107, 2114, 2134)*

[17] O. GOODMAN, *Snap*, version 1.11.3, http://www.ms.unimelb.edu.au/~snap. *(2150, 2156)*

[18] P. MENAL-FERRER and J. PORTI, *Local coordinates for* SL(n, **C**)*-character varieties of finite-volume hyperbolic 3-manifolds*, Ann. Math. Blaise Pascal **19** (2012), 107–122. MR 2978315. DOI 10.5802/ambp.306. *(2151)*

[19] W. D. NEUMANN, *Extended Bloch group and the Cheeger–Chern–Simons class*, Geom. Topol. **8** (2004), 413–474. MR 2033484. DOI 10.2140/gt.2004.8.413. *(2112, 2114)*

[20] W. NEUMANN, "Realizing arithmetic invariants of hyperbolic 3-manifolds" in *Interactions Between Hyperbolic Geometry, Quantum Topology and Number Theory)*, Contemp. Math. **541**, Amer. Math. Soc., Providence, 2011, 233–246. MR 2796636. DOI 10.1090/conm/541/10687. *(2108)*

[21] W. D. NEUMANN and D. ZAGIER, *Volumes of hyperbolic three-manifolds*, Topology **24** (1985), 307–332. MR 0815482. DOI 10.1016/0040-9383(85)90004-7. *(2156)*

[22] C.-H. SAH, *Homology of classical Lie groups made discrete, III*, J. Pure Appl. Algebra **56** (1989), 269–312. MR 0982639. DOI 10.1016/0022-4049(89)90061-3. *(2132)*

[23] N. STEENROD, *The Topology of Fibre Bundles*, Princeton Math. Ser. **14**, Princeton Univ. Press, Princeton, 1951. MR 0039258. *(2144)*

[24] A. A. SUSLIN, "Homology of GL_n, characteristic classes and Milnor K-theory" in *Algebraic K-theory, Number Theory, Geometry and Analysis (Bielefeld, 1982)*, Lecture Notes in Math. **1046**, Springer, Berlin, 1984, 357–375. MR 0750690. DOI 10.1007/BFb0072031. *(2132)*

[25] ———, *K_3 of a field, and the Bloch group* (in Russian), Trudy Mat. Inst. Steklov. **183** (1990), 180–199, 229; English translation in Proc. Steklov Inst. Math. **1991**, no. 4, 217–239. MR 1092031. *(2108)*

[26] W. P. THURSTON, *The geometry and topology of three-manifolds*, lecture notes, Princeton University, Princeton, 1980, http://library.msri.org/books/gt3m. *(2156)*

[27] C. K. ZICKERT, *The volume and Chern-Simons invariant of a representation*, Duke Math. J. **150** (2009), 489–532. MR 2582103. DOI 10.1215/00127094-2009-058. *(2109, 2114, 2117, 2124, 2128, 2131, 2135, 2137, 2149, 2157)*

[28] ———, *The extended Bloch group and algebraic K-theory*, J. Reine Angew. Math. **704** (2015), 21–54. MR 3365773. DOI 10.1515/crelle-2013-0055. *(2112, 2114, 2124)*

[29] ———, *The extended Bloch group and algebraic K-theory*, preprint, arXiv:0910.4005 [math.GT]. *(2108, 2125)*

Garoufalidis

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia, USA;
stavros@math.gatech.edu; http://www.math.gatech.edu/~stavros

Thurston

Department of Mathematics, Columbia University, New York, New York, USA;
dthurston@barnard.edu; http://www.math.columbia.edu/~dpt

Zickert

University of Maryland, Department of Mathematics, College Park, Maryland, USA;
zickert@umd.edu; http://www2.math.umd.edu/~zickert

BLOCH GROUPS, ALGEBRAIC K-THEORY, UNITS, AND NAHM'S CONJECTURE

FRANK CALEGARI, STAVROS GAROUFALIDIS, AND DON ZAGIER

Abstract. Given an element of the Bloch group of a number field F and a natural number n, we construct an explicit unit in the field $F_n = F(e^{2\pi i/n})$, well-defined up to n-th powers of nonzero elements of F_n. The construction uses the cyclic quantum dilogarithm, and under the identification of the Bloch group of F with the K-group $K_3(F)$ gives an explicit formula for a certain abstract Chern class from $K_3(F)$. The units we define are conjectured to coincide with numbers appearing in the quantum modularity conjecture for the Kashaev invariant of knots (which was the original motivation for our investigation), and also appear in the radial asymptotics of Nahm sums near roots of unity. This latter connection is used to prove one direction of Nahm's conjecture relating the modularity of certain q-hypergeometric series to the vanishing of the associated elements in the Bloch group of $\overline{\mathbf{Q}}$.

Contents

1. Introduction

The purpose of the paper is to associate to an element ξ of the Bloch group of a number field F and a primitive nth root of unity ζ an explicit unit or near unit $R_\zeta(\xi)$ in the field $F_n = F(\zeta)$, well-defined up to n-th powers of nonzero elements of F_n. Our construction uses the cyclic quantum dilogarithm and is shown to give an explicit formula for an abstract Chern class map on $K_3(F)$. The near unit is conjectured (and checked numerically in many cases) to coincide with a specific number that appears in the Quantum Modularity Conjecture of the Kashaev invariant of a knot. This was in fact the starting point of our investigation [13], [39].

As a surprising consequence of our main theorem we were able to prove one direction of Werner Nahm's famous conjecture, namely that the modularity of certain q-hypergeometric series ("Nahm sums") implies the vanishing of certain explicit elements in the Bloch group of $\overline{\mathbf{Q}}$. A precise statement will be given in Section 1.3 of this introduction.

1.1. Bloch groups and associated units.

We first recall the definition of the classical Bloch group, as introduced in [2]. Let $Z(F)$ denote the free abelian group on $\mathbf{P}^1(F) = F \cup \{\infty\}$, i.e. the group of formal finite combinations $\xi = \sum_i n_i[X_i]$ with $n_i \in \mathbf{Z}$ and $X_i \in \mathbf{P}^1(F)$.

Definition 1.1. The *Bloch group* of a field F is the quotient

$$B(F) = A(F)/C(F)\,, \tag{1}$$

where $A(F)$ is the kernel of the map

$$d:\ Z(F) \longrightarrow \textstyle\bigwedge^2 F^\times\,, \qquad [X] \mapsto (X) \wedge (1-X) \tag{2}$$

(and $[0]$, $[1]$, $[\infty] \mapsto 0$) and $C(F) \subseteq A(F)$ the group generated by the *five-term relation*

$$\xi_{X,Y} \;=\; [X] - [Y] + \left[\frac{Y}{X}\right] - \left[\frac{1-X^{-1}}{1-Y^{-1}}\right] + \left[\frac{1-X}{1-Y}\right] \tag{3}$$

with X and Y ranging over $\mathbf{P}^1(F)$ (but forbidding arguments $\frac{0}{0}$ or $\frac{\infty}{\infty}$ on the right-hand side).

In this paper, we will study an invariant of the Bloch group whose values are units in F_n modulo nth powers of units, where n is a natural number and F_n the field obtained by adjoining to F a primitive n-th root of unity $\zeta = \zeta_n$. The extension F_n/F is Galois with Galois group $G = \mathrm{Gal}(F_n/F)$, and G admits a canonical map

$$\chi:\ G \longrightarrow (\mathbf{Z}/n\mathbf{Z})^\times \tag{4}$$

determined by $\sigma\zeta = \zeta^{\chi(\sigma)}$. The powers χ^j ($j \in \mathbf{Z}/n\mathbf{Z}$) of this character define eigenspaces $\bigl(F_n^\times/F_n^{\times n}\bigr)^{\chi^j}$ in the obvious way as the set of $x \in F_n^\times/F_n^{\times n}$ such that $\sigma(x) = x^{\chi^j(\sigma)}$ for all $\sigma \in G$, and similarly for $(\mathcal{O}_n^\times/\mathcal{O}_n^{\times n})^{\chi^j}$ or $(\mathcal{O}_{S,n}^\times/\mathcal{O}_{S,n}^{\times n})^{\chi^j}$, where $\mathcal{O}_n$ (resp. $\mathcal{O}_{S,n}$) is the ring of integers (resp. S-integers) of F_n. Then our main result is the following theorem.

Theorem 1.2. *Suppose that F does not contain any non-trivial nth root of unity. Then there is a canonical map*

$$R_\zeta :\, B(F)/nB(F) \;\longrightarrow\; \bigl(\mathcal{O}_{S,n}^\times/\mathcal{O}_{S,n}^{\times n}\bigr)^{\chi^{-1}} \;\subset\; \bigl(F_n^\times/F_n^{\times n}\bigr)^{\chi^{-1}} \tag{5}$$

for some finite set S of primes depending only on F. If n is prime to a certain integer M_F depending on F, then the map R_ζ is injective and its image is contained in $\bigl(\mathcal{O}_n^\times/\mathcal{O}_n^{\times n}\bigr)^{\chi^{-1}}$, and equal to this if n is prime.

Remark 1.3. Note that the field F_n and the character χ of (4) do not depend on the primitive nth root of unity ζ. The map R_ζ from $B(F)$ to $F_n^\times/F_n^{\times n}$ does depend on ζ, but in a very simple way, described by either of the formulas

$$\sigma\bigl(R_\zeta(\xi)\bigr) \;=\; R_{\sigma(\zeta)}(\xi) \quad (\sigma \in G), \qquad R_\zeta(\xi) \;=\; R_{\zeta^k}(\xi)^k \quad (k \in (\mathbf{Z}/n\mathbf{Z})^\times)\,, \tag{6}$$

where the simultaneous validity of these two formulas explains why the image of each map R_ζ lies in the χ^{-1} eigenspace of $F_n^\times/F_n^{\times n}$.

Remark 1.4. The optimal definition of M_F is somewhat complicated to state. However, one may take it to be $6\,\Delta_F\,|K_2(\mathcal{O}_F)|$.

The detailed construction of the map R_ζ will be given in Section 2. A rough description is as follows. Let $\xi = \sum n_i[X_i]$ be an element of $Z(F)$ whose image in $\wedge^2(F^\times/F^{\times n})$ under the map induced by d vanishes. We define an algebraic number $P_\zeta(\xi)$ by the formula

$$P_\zeta(\xi) \;=\; \prod_i D_\zeta(x_i)^{n_i}\,, \tag{7}$$

where x_i is some nth root of X_i and $D_\zeta(x)$ is the *cyclic quantum dilogarithm function*

$$D_\zeta(x) \;=\; \prod_{k=1}^{n-1}(1 \,-\, \zeta^k x)^k\,. \tag{8}$$

The number $P_\zeta(\xi)$ belongs to the Kummer extension H_ξ of F defined by adjoining all of the x_i to F_n and is well-defined modulo H_ξ^n. We show that for n prime to some M_F it has the form ab^n with b in $H_\xi^\times$ and $a \in F_n^\times$ (or even $a \in \mathcal{O}_n^\times$ under a sufficiently strong coprimality assumption about n). Then $R_\zeta(\xi)$ is defined as the image of a modulo nth powers.

1.2. **Algebraic K-groups and associated units.** A second main theme of the paper concerns the relation to the algebraic K-theory of fields. The group $B(F)$ was introduced by Bloch as a concrete model for the abstract K-group $K_3(F)$. It was proved by Suslin [31] that, if F is a number field, then (up to 2-torsion) $K_3(F)$ is an extension of $B(F)$ by the roots of unity in F, and in this case one also knows by results of Borel and Suslin-Merkurjev [30], [21], [36] that $K_3(F)$ has the structure

$$K_3(F) \;\cong\; \mathbf{Z}^{r_2(F)} \;\oplus\; \begin{cases} \mathbf{Z}/w_2(F)\mathbf{Z} & \text{if } r_1(F)=0,\\ \mathbf{Z}/2w_2(F)\mathbf{Z} \,\oplus\, (\mathbf{Z}/2\mathbf{Z})^{r_1(F)-1} & \text{if } r_1(F)\ge 1,\end{cases} \tag{9}$$

where $(r_1(F), r_2(F))$ is the signature of F and $w_2(F)$ is the integer

$$w_2(F) \;=\; 2\prod_p p^{\nu_p}\,, \qquad \nu_p \;:=\; \max\bigl\{\nu\in\mathbf{Z} \mid \zeta_{p^\nu}+\zeta_{p^\nu}^{-1}\in F\bigr\}\,. \tag{10}$$

For a detailed introduction to the the algebraic K-theory of number fields, see [36].

Theorem 1.2 is then a companion of the following result concerning $K_3(F)$:

Theorem 1.5. *Let F be a number field. Then there is a canonical map*

$$c_\zeta:\; K_3(F)/nK_3(F) \;\longrightarrow\; \bigl(\mathcal{O}_{S,n}^\times/\mathcal{O}_{S,n}^{\times n}\bigr)^{\chi^{-1}} \;\subset\; \bigl(F_n^\times/F_n^{\times n}\bigr)^{\chi^{-1}} \tag{11}$$

defined using the theory of Chern classes for some finite set S of primes depending only on F. If n is prime to a certain integer M_F depending on F, then the map R_ζ is injective and its image is contained in $\bigl(\mathcal{O}_n^\times/\mathcal{O}_n^{\times n}\bigr)^{\chi^{-1}}$, and equal to this if n is prime.

We note that the proof of Theorem 1.2 relies upon the precise computation of $K_3(F)$ and the properties of c_ζ given above. Finally, in view of the near isomorphism between $B(F)$ and $K_3(F)$, one might guess that the two maps P_ζ and c_ζ are the same, at least up to a simple scalar. This is the content of our next theorem.

Theorem 1.6. *For n prime to M_F, the map R_ζ equals c_ζ^γ for some $\gamma\in(\mathbf{Z}/n\mathbf{Z})^\times$.*

The constant γ does not depend on the underlying field — both our construction and the Chern class map are well behaved in finite extensions, so we can compare the maps over any two fields with the maps in their compositum. We conjecture that the constant γ is, up to sign, a small power of 2 that is independent of n. To justify our conjecture, and to determine γ, it suffices to compute the image under both maps R_ζ and c_ζ of some element of $K_3(F)/nK_3(F)$ of exact order n. For each root of unity ζ of order n, there is a specific element η_ζ (eq. (23)) of the finite Bloch group $B(\mathbf{Q}(\zeta+\zeta^{-1}))$ that is of exact order n. Using the relation of the map R_ζ to the radial asymptotics of certain q-series called Nahm sums discussed in Section 8, we will prove

$$R_\zeta(\eta_\zeta)^4 \;=\; \zeta \tag{12}$$

(Theorem 8.5). On the other hand, certain expected functorial properties of the map c_ζ, discussed in Section 5.3 indicate that up to sign and a small power of 2, we have:

$$c_\zeta(\eta_\zeta) \;\overset{?}{=}\; \zeta\,, \tag{13}$$

and in combination with (12) this justifies our conjecture concerning γ.

The above-mentioned relation between our *mod n regulator map* on Bloch groups and the asyptotics of Nahm sums near roots of unity is also an ingredient of our proof of one direction of Nahm's conjecture (under some restrictions) relating the modularity of his sums to torsion in the Bloch group. The argument, described in Section 8.3, uses the full strength of Theorem 1.2 and gives a nice demonstration of the usefulness, despite its somewhat abstract statement, of that theorem.

Theorem 1.2 motivates a mod n (or *étale*) version of the Bloch group of a number field F, defined by

$$B(F;\mathbf{Z}/n\mathbf{Z}) \;=\; A(F;\mathbf{Z}/n\mathbf{Z})/(nZ(F)+C(F))\,, \tag{14}$$

where $A(F;\mathbf{Z}/n\mathbf{Z})$ is the kernel of the map $d : Z(F) \to \wedge^2(F^\times/F^{\times\, n})$ induced by d. This is studied in Section 6, where we establish the following relation to $K_2(F)$.

Theorem 1.7. *The étale Bloch group is related to the original Bloch group by an exact sequence*

$$0 \longrightarrow B(F)/nB(F) \longrightarrow B(F;\mathbf{Z}/n\mathbf{Z}) \longrightarrow K_2(F)[n] \longrightarrow 0\,, \tag{15}$$

where $K_2(F)[n]$ is the n-torsion in the K-group $K_2(F)$.

There is a corresponding exact sequence (equation (27)) with $B(F)/nB(F)$ replaced by $K_3(F)/nK_3(F)$ and $B(F;\mathbf{Z}/n\mathbf{Z})$ replaced by a Galois cohomology group.

A large part of the story that we have told here for the Bloch group $B(F)$ and the third K-group $K_3(F)$ can be generalized to higher Bloch groups $B_m(F)$ and $K_{2m-1}(F)$ with $m \ge 2$, and here the étale version really comes into its own, because the higher Bloch groups as originally introduced in [37] have several alternative definitions that are only conjecturally isomorphic and are difficult or impossible to compute rigorously, whereas their étale versions turn out to have a canonial definition and be amenable to rigorous computations. The study of the higher cases has many proofs in common with the $m = 2$ case studied here, but there are also many new aspects, and the discussion will therefore be given in a separate paper [3] which is work in progress.

1.3. **Nahm's Conjecture.** The near unit constructed in Section 1.1 also appears in connection with the asymptotics near roots of unity of certain q-hypergeometric series called Nahm sums. These series are defined by

$$f_{A,B,C}(q) \;=\; \sum_{m\in\mathbf{Z}^r_{\ge 0}} \frac{q^{\frac12 m^tAm+Bm+C}}{(q)_{m_1}\cdots(q)_{m_r}}\,,$$

where $A \in M_r(\mathbf{Q})$ is a positive definite symmetric matrix, B an element of $\mathbf{Q}^r$, and C a rational number. Based on ideas coming from characters of rational conformal field theories, Nahm conjectured a relation between the modularity of the associated holomorphic function $\widetilde{f}_{A,B,C}(\tau) = f_{A,B,C}(e^{2\pi i\tau})$ in the complex upper half-plane and the vanishing of a certain element or elements in the Bloch group of $\overline{\mathbf{Q}}$. (See [24], [38], and Section 8 for more details.) This relation conjecturally goes in both directions, but with the implication from the vanishing of the Bloch elements to the modularity of certain Nahm sums not yet having a sufficiently precise formulation to be studied. The conjectural implication from modularity to vanishing of Bloch elements, on the other hand, had a completely precise formulation, as follows. Let A be as above and $(X_1,\dots,X_r)$ the unique solution in $(0,1)^r$ of *Nahm's equation*

$$1 \,-\, X_i \;=\; \prod_{j=1}^r X_j^{a_{ij}} \qquad (i=1,\dots,r)\,.$$

Then Nahm shows that the element $\xi_A = \sum_{i=1}^{r}[X_i]$ belongs to $B(\mathbf{R}\cap\overline{\mathbf{Q}})$, and his assertion is the following theorem, which we will prove as a consequence of the injectivity statement in Theorem 1.2.

Theorem 1.8 (One direction of Nahm's Conjecture)**.** *If the function $\widetilde{f}_{A,B,C}(\tau)$ is modular for some A, B and C as above, then ξ_A vanishes in the Bloch group of $\overline{\mathbf{Q}}$.*

We remark that the vanishing condition can be (and often is) stated by saying that ξ_A is a torsion element in the Bloch group of the smallest real (but in general not totally real) number field containing all the X_i, but when we take the image of this Bloch group in the Bloch group of $\overline{\mathbf{Q}}$ or $\mathbf{C}$, then the torsion vanishes.

1.4. **Plan of the paper.** In Section 2 we recall the cyclic quantum dilogarithm and use it, together with some basic facts about Kummer extensions, to define the map R_ζ. The fact that the map R_ζ satisfies the 5-term relation follows from some state-sum identities of Kashaev-Mangazeev-Stroganov [17], reviewed in Section 2.4. The remaining statements of Theorem 1.2 are deduced from Theorems 1.5 and 1.6.

In Section 3 we recall the basic properties of Chern classes and use them to define the map c_ζ and prove Theorem 1.5. Its proof follows from Lemmas 3.1 and 3.5.

The comparison of the maps c_ζ and R_ζ is done via reduction to the case of finite fields. This reduction is discussed in Section 4, and the proof of Theorem 1.6 is given in Section 5.

In Section 6, we discuss the connection of our map R_ζ with Tate's results on $K_2(\mathcal{O}_F)$.

The units produced by our map R_ζ have also appeared in two related contexts, namely in the Quantum Modularity Conjecture concerning the asymptotics of the Kashaev invariant at roots of unity, and in the asymptotics of Nahm sums at roots of unity. In Section 7, we give examples of the units produced by our map R_ζ and compare them with those that appear in the Quantum Modularity Conjecture. In Section 8, we state the connection of our map R_ζ with the radial asymptotics of Nahm sums at roots of unity and give two applications: a proof equation (12) (as a consequence of a special modular Nahm sum, the Andrews-Gordon identity), and a proof of Theorem 1.8.

Remark. During the writing of this paper, we learned that Gangl and Kontsevich in unpublished work also proposed the map P_ζ as an explicit realization of the Chern class map. Although they did not check in general that the image of P_ζ could be lifted to a suitable element $R_\zeta \in (F_n^\times/F_n^{\times n})^{\chi^{-1}}$, they did propose an alternate proof of the 5-term identity using cyclic algebras. Goncharov also informs us that he was aware many years ago that the function P_ζ should be an explicit realization of the Chern class map.

2. The maps P_ζ and R_ζ

Let n be a positive integer, and let F be a field of characteristic prime to n. Let $F_n = F(\zeta_n)$, and let $\zeta = \zeta_n \in F_n$ denote a primitive nth root of unity, which we usually consider as fixed and omit from the notations. For convenience, we will always assume that n prime to 6.

2.1. **The map P_ζ.** Let $\mu = \langle\zeta\rangle$ denote the G_F-module of nth roots of unity. Recall that $F_n = F(\zeta)$. The *universal Kummer extension* is by definition the extension H/F_n obtained by adjoining nth roots of every element in F. Let $\Phi = \mathrm{Gal}(H/F_n)$. We have [20, Chpt.VI]:

Lemma 2.1. The extension H/F is Galois. There is a natural isomorphism

$$\phi : F^\times/F^{\times n} \;\simeq\; \mathrm{Hom}(\Phi,\mu) \;\simeq\; H^1(\Phi,\mu)$$

given by $X \mapsto (\sigma \in \Phi \mapsto \sigma x/x)$, where $x \in H^\times$ is any element that satisfies $x^n = X$.

Consider the function

$$P_\zeta(X) := D_\zeta(x) \;\in\; H^\times/H^{\times n} \qquad (X \in F^\times \smallsetminus \{0,1\},\; x^n = X)\,, \tag{16}$$

where $D_\zeta(x)$ is the cyclic quantum dilogarithm defined in (8). (We previously defined $P_\zeta(X)$, in equation (8) of the induction, as an element of $H^\times$, but only its image modulo nth powers was ever used, and it is more canonical to define it in the manner above.)

Lemma 2.2. The function $P_\zeta : F^\times \to H^\times/H^{\times n}$ has the following properties.
(a) $P_\zeta(X)$ is independent of the choice of nth root x of X.
(b) $P_\zeta(1) = 1$, and more generally $P_\zeta(X)P_\zeta(1/X) = 1$ for any $X \in F_n^\times$.
(c) $P_\zeta(X) \in H^\times/H^{\times n}$ is invariant under the action of $\Phi = \mathrm{Gal}(H/F_n)$.
(d) $\sigma(P_\zeta(X)) = P_\zeta(X)^{\chi^{-1}(\sigma)}$ for all $\sigma \in G$.

Proof. First note that, because $P_\zeta(X)$ is defined only up to nth powers, we can replace the definition (16) by

$$P_\zeta(X) \;=\; \prod_{k \bmod n} (1 \,-\, \zeta^k x)^k \quad \bmod H^{\times n} \qquad (x^n = X), \tag{17}$$

where we can now even include the $k = 0$ term that was omitted in (8). Part (a) then follows from the calculation

$$\frac{\prod_{k \bmod n}(1-\zeta^k x)^k}{\prod_{k \bmod n}(1-\zeta^{k+1}x)^k} \;=\; \prod_{k \bmod n}(1-\zeta^k x) \;=\; 1 - X \;\in\; F^\times \;\subset\; H^{\times n}.$$

Similarly, replacing k by $-k$ in the definition of $P_\zeta(1/X)$, gives

$$P_z(X)P_z(1/X) \;=\; \prod_{k \bmod n}(1-\zeta^k x)^k(1-\zeta^{-k}x^{-1})^{-k} \;=\; \prod_{k \bmod n}(-\zeta^k x)^k \;=\; 1 \quad \in\; H^\times/H^{\times n}\,,$$

proving the second statement of (b), and the first statement follows because an element killed by both 2 and the odd number n in any group must be trivial. (It can also be proved more explicitly by evaluating $D_n(1)^n$ itself for $(n,6) = 1$ as the $(-1)^{n(n-1)/2}n^n$, which is an nth power because $(-1)^{(n-1)/2}n$ is a square in $\mathbf{Q}(\zeta_n)$.) For part (c), we note that the effect of an element $\sigma \in \Phi$ on $D_\zeta(x)$ is to replace x by $\zeta^i x$ for some i, so the result follows from part (a). For part (d), we first observe that the statement makes sense because $\Phi = \mathrm{Gal}(H/F_n)$ is a normal subgroup of $\mathrm{Gal}(H/F)$ and hence acts trivially on $P_\zeta(X) \in H^\times/H^{\times n}$ by virtue of (c), so that the quotient $G = \mathrm{Gal}(F_n/F)$ acts on $P_\zeta(X)$. For the proof, we choose a lift of $\sigma \in G$ to $\mathrm{Gal}(H/F)$ that fixes x. Then

$$\sigma P_\zeta(X) \;=\; \prod_k \bigl(1-\sigma(\zeta)^k x\bigr)^k \;=\; \prod_k \bigl(1-\zeta^{k\chi(\sigma)}x\bigr)^k \;=\; \prod_k \bigl(1-\zeta^k x\bigr)^{k\chi(\sigma)^{-1}} \;=\; P_\zeta(X)^{\chi(\sigma)^{-1}}\,,$$

where all products are over k (mod n) and all calculations are modulo $H^{\times n}$. $\square$

Remark 2.3. When n is not prime to 6, then we could also make the calculations above work after replacing the right-hand side of (16) by $P_\zeta(X) = \frac{D_\zeta(x)}{D_\zeta(1)}$. (When $(n,6)=1$ this is not necessary since an elementary calculation shows that then $D_\zeta(1) \in \mathbf{Q}(\zeta)^n$.)

We extend the map P_ζ to the free abelian group $Z(F) = \mathbf{Z}[\mathbf{P}^1(F)]$ by linearity as in (7), with $P_\zeta(0) = P_z(1) = P_\zeta(\infty) = 0$.

2.2. **The map R_ζ.** The next proposition associates an element $R_\zeta(\xi) \in (F_n^\times/F_n^{\times n})^{\chi^{-1}}$ to every element of $B(F)/nB(F)$ as long as $(n, w_F) = 1$. Recall the group $A(F;\mathbf{Z}/n\mathbf{Z})$ from subsection 1.1.

Proposition 2.4. (a) For $\xi \in A(F;\mathbf{Z}/n\mathbf{Z})$, the image of $P_\zeta(\xi)^{w_F}$ lifts to $F_n^\times/F_n^{\times n}$.
(b) The image of $P_\zeta(\xi)^{w_F}$ admits a unique lift to $F_n^\times/F_n^{\times n}$ on which G acts by χ^{-1}. If n is prime to w_F, then $P_\zeta(\xi)$ itself admits a unique lift $R_\zeta(\xi) \in (F_n^\times/F_n^{\times n})^{\chi^{-1}}$.

Proof. For part (a), by Hilbert 90 and inflation-restriction, there is a commutative diagram:

$$\begin{tikzcd}
H^1(\Phi,\mu) \arrow[r] & H^1(F_n,\mu) \arrow[r] \arrow[d, equal] & H^1(H,\mu)^\Phi \arrow[r, "\delta"] \arrow[d, equal] & H^2(\Phi,\mu) \\
 & F_n^\times/F_n^{\times n} \arrow[r] & \left(H^\times/H^{\times n}\right)^\Phi &
\end{tikzcd}$$

That is, there is an obstruction to descending from $\left(H^\times/H^{\times n}\right)^\Phi$ to $F_n^\times/F_n^{\times n}$ which lands in $H^2(\Phi,\mu)$.

We now claim that there is a commutative diagram as follows:

$$\begin{tikzcd}
Z(F) \arrow[r, "P_\zeta"] \arrow[d, "d"'] & (H^\times/H^{\times n})^\Phi \arrow[d, "\delta"] \\
\bigwedge^2(F^\times/F^{\times n}) \arrow[r, hook, "\cup"] & H^2(\Phi,\mu)\ ,
\end{tikzcd}$$

where the left vertical map is the one defined in (2) and the bottom horizontal map is the map induced by the cup product from the isomorphism $F^\times/F^{\times n} \to H^1(\Phi,\mu)$ of Lemma 2.1. Note that the cup product is more naturally a map $\bigwedge^2 H^1(\Phi,\mu) \to H^2(\Phi,\mu^{\otimes 2})$, but can be interpreted as in the theorem by using the trivialization $\mu \simeq \mathbf{Z}/n\mathbf{Z} \simeq \mu^{\otimes 2}$ defined by the choice of the root of unity ζ.

We now show that the above diagram commutes. By linearity, it suffices to prove this for elements ξ of the form $[X]$. Write $X = x^n$ and $1 - X = y^n$. For $Z \in F^\times/F^{\times n}$ and $z^n = Z$, let (following Lemma 2.1),

$$\sigma(z) = \zeta^{\phi(z,\sigma)} z.$$

By definition, we have $P_\zeta([X]) = D_\zeta(x)$ modulo nth powers. The obstruction to lifting $D(x)$ amounts to finding an element $u \in H^\times$ such that $D_\zeta(x)/u^n \in F_n^\times$. Such a u would necessarily satisfy

$$\left(\frac{\sigma u}{u}\right)^n = \frac{\sigma D_\zeta(x)}{D_\zeta(x)} = \frac{D_\zeta(\zeta^{\phi(x,\sigma)}x)}{D_\zeta(x)} = \left(\prod_{k=0}^{\phi(x,\sigma)-1} \frac{1-\zeta^k x}{y}\right)^n.$$

The expression inside the nth power is determined exactly modulo $\langle\zeta\rangle$. Hence we may define a cocycle

$$h = h_X : \Phi \to H^\times/\mu, \qquad h(\sigma) := \prod_{k=0}^{\phi(x,\sigma)-1} \frac{1-\zeta^k x}{y}.$$

This gives an element of $H^1(\Phi, H^\times/\mu)$, which by consideration of the exact sequence

$$H^1(\Phi, H^\times) \longrightarrow H^1(\Phi, H^\times/\langle\zeta\rangle) \longrightarrow H^2(\Phi,\mu)$$

maps to $H^2(\Phi,\mu)$. This is actually an injection, because the first term vanishes by Hilbert 90. This is the image of δ; explicitly, the class $\delta(h) \in H^2(\Phi,\mu)$ is given by

$$\begin{aligned}
\delta(h)(\sigma,\tau) &= \frac{h(\sigma\tau)}{h(\sigma)\sigma h(\tau)} \\
&= \frac{1}{h(\sigma)\sigma h(\tau)} \prod_{k=0}^{\phi(x,\sigma)+\phi(x,\tau)-1} \frac{1-\zeta^k x}{y} \\
&= \frac{1}{h(\sigma)\sigma h(\tau)} \prod_{k=0}^{\phi(x,\sigma)-1} \frac{1-\zeta^k x}{y} \prod_{k=0}^{\phi(x,\tau)-1} \frac{1-\zeta^k\zeta^{\phi(x,\sigma)}x}{y} \\
&= \frac{1}{h(\sigma)\sigma h(\tau)} \prod_{k=0}^{\phi(x,\sigma)-1} \frac{1-\zeta^k x}{y} \prod_{k=0}^{\phi(x,\tau)-1} \frac{1-\zeta^k\zeta^{\phi(x,\sigma)}x}{\zeta^{\phi(y,\sigma)}y} \cdot \zeta^{\phi(y,\sigma)} \\
&= \zeta^{\phi(x,\tau)\phi(y,\sigma)}
\end{aligned}$$

On the other hand, the class in $H^1(\Phi,\mu)$ associated to $X = x^n$ is the map $\tau \mapsto \zeta^{\phi(x,\tau)}$, and the class associated to $1 - X = y^n$ is the map $\sigma \mapsto \zeta^{\phi(y,\sigma)}$, and the exterior product of these two classes in $H^2(\Phi,\zeta)$ is precisely $\delta(h)$. The fact that the cup product gives an injection is an easy fact about the cohomology of abelian groups of exponent n. This concludes the proof of part (a).

For part (b), suppose that $\xi \in A(F; \mathbf{Z}/n\mathbf{Z})$. By the argument above, there certainly exists an element in $F_n^\times/F_n^{\times n}$ which maps to $P_\zeta(\xi)$. Let M denote the image of $F_n^\times/F_n^{\times n}$ in $(H^\times/H^{\times n})^\Phi$, and let $S = F^\times/F^{\times n}$. We have a short exact sequence as follows:

$$0 \longrightarrow S \longrightarrow F_n^\times/F_n^{\times n} \longrightarrow M \longrightarrow 0.$$

Taking χ^{-1}-invariants is the same as tensoring with $\mathbf{Z}/n\mathbf{Z}(1)$ and taking invariants. Hence there is an exact sequence

$$\left(F_n^\times/F_n^{\times n}\right)^{\chi^{-1}} \longrightarrow M^{\chi^{-1}} \longrightarrow H^1(G, S(1)).$$

In particular, the obstruction to lifting to a χ^{-1}-invariant element lies in $H^1(G,S(1))$, and it suffices to prove that this group is annihilated by w_F. By construction, the module S is trivial as a G-module, and hence the action of G on $S(1)$ is via the character χ. Sah's

Lemma ([19, Lem.8.8.1]) implies that the self-map of $H^1(G, S(1))$ induced by $g-1$ for any $g \in Z(G) = G$ is the zero map. On the other hand, since $\chi : G \to (\mathbf{Z}/n\mathbf{Z})^\times$ is the cyclotomic character, the greatest common divisor of $\chi(g) - 1$ for $g \in G$ is $w_F\mathbf{Z}/n\mathbf{Z}$. In particular, the group is annihilated by w_F. The result follows. $\square$

Remark 2.5. Suppose $(w_F, n) = 1$, and let $P \in H^\times$ be a representative of $P_\zeta(\xi) \in H^\times/H^{\times n}$. Then the construction of the element $R_\zeta(\xi)$ whose existence is asserted by Proposition 2.4 reduces to the problem of finding $S \in H^\times$ such that

(a) $P/S^n \in F_n^\times$, and
(b) the image of P/S^n in $F_n^\times/F_n^{\times n}$ lies in the χ^{-1}-eigenspace,

since then $R_\zeta(\xi) = P/S^n \in (F_n^\times/F_n^{\times n})^{\chi^{-1}}$. In practice, S will be constructed via a Hilbert 90 argument as an additive Galois average, and the difficulty is ensuring that $S \neq 0$. See Section 8, where this is done for a particular P constructed as a radial limit of a Nahm sum.

2.3. **Reduction to the case of prime powers.** In this section, we discuss the compatibility of the map R_ζ with the prime factorization of n. This will be important in Section 5, where we consider the relation of our map and the Chern class in K-theory.

Lemma 2.6. Let $(n, w_F) = 1$ and $\zeta = \zeta_n$ as usual. Then the following compatibilities hold:

(1) If $(n, k) = 1$, then $R_{\zeta^k}(X) = R_\zeta(X)^{k^{-1}}$.
(2) Let $n = qr$, and let $\zeta_r = \zeta_n^q$. Then the image of R_{ζ_n} modulo rth powers is equal to the image of $R_{\zeta_r}(X)$ under the map

$$\left(F_r^\times/F_r^{\times r}\right)^{\chi^{-1}} \to \left(F_n^\times/F_n^{\times r}\right)^{\chi^{-1}}$$

induced by the inclusion.

Proof. The first statement reflects the fact that $gR_\zeta = R_{g(\zeta)}$ for $g \in G = \mathrm{Gal}(F_n/F)$. For the second claim, we calculate

$$\begin{aligned} P_{\zeta_n}(X) &= \prod_{k \bmod n} \left(1 - \zeta_n^k x\right)^k = \prod_{\substack{i \bmod q \\ j \bmod r}} \left(1 - \zeta_n^{ri+j} x\right)^{ri+j} \\ &\equiv \prod_{\substack{i \bmod q \\ j \bmod r}} \left(1 - \zeta_q^i \zeta_n^j x\right)^j = \prod_{j \bmod r} \left(1 - \zeta_r^j x^q\right)^j = P_{\zeta_r}(X), \end{aligned}$$

where the congruence is modulo rth powers. $\square$

Next, we discuss a reduction of the map P_{ζ_n} to the case that n is a prime power.

Lemma 2.7. Let $n = ab$ with $(a, b) = 1$ and ζ a primitive nth root of unity. If $X \in A(F; \mathbf{Z}/n\mathbf{Z})$, let $u_n = R_\zeta(X)$, $u_a = R_{\zeta^b}(X)$ and $u_b = R_{\zeta^a}(X)$. Then u_n determines and is uniquely determined by u_a and u_b.

Proof. Part (2) of Lemma 2.6 shows that the image of u_n in $F_n^\times/F_n^{\times a}$ is the image of u_a under the natural map

$$F_a^\times/F_a^{\times a} \to F_n^\times/F_n^{\times a}.$$

Equivalently, u_a determines u_n up to an ath power, and similarly u_b determines u_n up to a bth power. This is enough to determine u_n completely since a and b are coprime. The converse is already shown. $\square$

Remark 2.8. Both lemmas hold also for $(n, w_F) > 1$ if we replace R_ζ by $R_\zeta^{w_F}$.

2.4. **The 5-term relation.** In this section, we use a result of Kashaev, Mangazeev and Stroganov to show that the map R_ζ satisfies the 5-term relation, and consequently descends to a map of the group $B(F; \mathbf{Z}/n\mathbf{Z})$.

Theorem 2.9. *Let F be a field and $F_n = F(\zeta)$, where ζ is a root of unity of order n prime to w_F and to the characteristic prime of F. Then the map R_ζ vanishes on the subgroup $C(F) \subset A(F; \mathbf{Z}/n\mathbf{Z}) \subset Z(F)$ generated by the 5-term relation, and therefore induces a map*

$$B(F) \longrightarrow B(F)/nB(F) \longrightarrow B(F; \mathbf{Z}/n\mathbf{Z}) \xrightarrow{R_\zeta} \left(F_n^\times / F_n^{\times n}\right)^{\chi^{-1}}.$$

Proof. Denote by H the universal Kummer extension as before. Then it suffices to show that the appropriate product of the functions D_ζ is a perfect nth power in H.

Let X, Y, $Z \in F^\times$ be related by $Z = (1-X)/(1-Y)$, and choose nth roots x, y, z of X, Y, Z. Using the standard notation $(x;q)_k = (1-x)(1-qx)\cdots(1-q^{k-1}x)$ (q-Pochhammer symbol) and following the notation of [17] (except that they use $w(x|k)$ for $(x\zeta;\zeta)_k^{-1}$), we set

$$f(x,\, y \,|\, z) = \sum_{k=0}^{n-1} \frac{(\zeta y;\zeta)_k}{(\zeta x;\zeta)_k} = \sum_{k \bmod n} \frac{(\zeta y;\zeta)_k}{(\zeta x;\zeta)_k}\, z^k \ \in H,$$

where the second equality follows from the relation between x, y, and z. By equation (C.7) of [17], we have

$$(\zeta y)^{n(1-n)/2}\, f(x,\, y \,|\, z)^n \ = \ \frac{D_\zeta(1) D_\zeta(y\zeta/x) D_\zeta(x/yz)}{D_\zeta(1/x) D_\zeta(y\zeta) D_\zeta(\zeta/z)}.$$

Considering this modulo nth powers, and using Lemma 2.1, we find

$$1 \ = \ P_\zeta(X)\, P_\zeta(Y)^{-1}\, P_\zeta(Y/X)\, P_\zeta(YZ/X)^{-1}\, P_\zeta(Z)\,.$$

This is precisely the 5-term relation for the map P_ζ, and from the uniqueness clause in Proposition 2.4 implies the same 5-term relation for the map R_ζ. $\square$

2.5. **An eigenspace computation.** As in Section 1.1, we write $G = \mathrm{Gal}(F_n/F)$, identified with a subgroup of $(\mathbf{Z}/n\mathbf{Z})^\times$ via the map χ of Equation (4). Since $F_n^\times/F_n^{\times n}$ is an n-torsion G-module, the χ^{-1} eigenspace makes sense and is given by

$$\left(F_n^\times/F_n^{\times n}\right)^{\chi^{-1}} = \{x \in F_n^\times/F_n^{\times n} \,|\sigma x = x^{\chi(\sigma^{-1})}\,, \text{for all}\, \sigma \in G\}\,,$$

where $x^{\chi(\sigma^{-1})}$ is computed using any lift of $\chi(\sigma^{-1}) \in (\mathbf{Z}/n\mathbf{Z})^\times$ to $\mathbf{Z}$.

In characteristic zero, one can also consider the the action of G on $M \otimes_{\mathbf{Z}} R$, where R is a $\mathbf{Z}[G]$ module that contains the eigenvalues of $\sigma \in G$. For example, one can take $M = \mathcal{O}_n^\times$ and $R = \mathbf{C}$. If $n = p$ is prime, then one can take $R = \mathbf{Z}_p$, which contains the $(p-1)$th roots of unity. In particular, if $n = p$, then one can define $(M \otimes_{\mathbf{Z}} \mathbf{Z}_p)^{\chi^{-1}}$, which will have the property that

$$(M \otimes_{\mathbf{Z}} \mathbf{Z}_p)^{\chi^{-1}} \otimes \mathbf{Z}/p\mathbf{Z} = (M/pM)^{\chi^{-1}}\,.$$

Proposition 2.10. (a) Suppose that F is disjoint from $\mathbf{Q}(\zeta_n)$. Then there exists an isomorphism of G-modules

$$\left(\mathcal{O}_n^\times \otimes \mathbf{C}\right)^{\chi^{-1}} = \mathbf{C}^{r_2(F)} . \tag{18}$$

(b) If, furthermore, $n = p$ is prime, so that $\chi : G \to (\mathbf{Z}/p\mathbf{Z})^\times$ admits a natural Teichmüller lift to $\mathbf{Z}_p^\times$, then

$$\operatorname{rank}_{\mathbf{Z}_p} \left(\mathcal{O}_n^\times \otimes \mathbf{Z}_p\right)^{\chi^{-1}} = r_2(F) .$$

If in addition χ and χ^{-1} are distinct characters of G, then

$$\left(\mathcal{O}_p^\times / \mathcal{O}_p^{\times p}\right)^{\chi^{-1}} = (\mathbf{Z}/p\mathbf{Z})^{r_2(F)} .$$

Proof. Part (b) follows easily from part (a) and the above discussion, together with the fact that if $\chi \neq \chi^{-1}$ then the torsion in the unit group (which just comprises roots of unity) is in the χ-eigenspace and not the χ^{-1}-eigenspace.

For (a), let $\widetilde{F}$ be the Galois closure of F over $\mathbf{Q}$ and let $\Gamma = \mathrm{Gal}(\widetilde{F}/\mathbf{Q})$. By assumption, with $\widetilde{F}_n = \widetilde{F}(\zeta_n)$, we have $\mathrm{Gal}(\widetilde{F}_n/\mathbf{Q}) = \Gamma \times G = \Gamma \times (\mathbf{Z}/n\mathbf{Z})^\times$. From the proof of Dirichlet's unit theorem, the unit group of $\widetilde{F}_n$, tensored with $\mathbf{C}$, decomposes equivariantly as

$$\bigoplus_W W^{\dim(W|c=1)} ,$$

where W runs over all the non-trivial irreducible representations of $\Gamma \times G$ and $c \in \Gamma$ is any complex conjugation, which we may take to be $(c, -1) \in G \times (\mathbf{Z}/n\mathbf{Z})^\times$ for a complex conjugation $c \in \Gamma$. The irreducible representations of W are of the form $U \otimes V$ for irreducible representations U of Γ and V of $G = (\mathbf{Z}/n\mathbf{Z})^\times$. Note that

$$\dim(U \otimes V|(c,-1) = 1) = \dim(U|c = 1) \dim(V|c = 1) + \dim(U|c = -1) \dim(V|c = -1) .$$

If we take the χ^{-1}-eigenspace under the action of the second factor, the only representation V of G which occurs is χ^{-1}, on which -1 acts by -1, and hence we are left with

$$\left(\mathcal{O}_{\widetilde{F}_n}^\times \otimes \mathbf{C}\right)^{\chi^{-1}} = \bigoplus_V V^{\dim(V|c=-1)} ,$$

where the sum runs over all representations V of Γ. In particular, there is an isomorphism in the Grothendieck group of G-modules

$$\left[\, \left(\mathcal{O}_{\widetilde{F}_n}^\times \otimes \mathbf{C}\right)^{\chi^{-1}} \,\right] + \left[\mathcal{O}_{\widetilde{F}}^\times \otimes \mathbf{C}\right] + \left[\mathbf{C}\right] = \left[\mathbf{C}[G]\right] .$$

Now take the $\Delta = \mathrm{Gal}(\widetilde{F}/F) = \mathrm{Gal}(\widetilde{F}_n/F_n)$-invariant part and take dimensions, we obtain the equality

$$\dim_{\mathbf{C}}\Big(\left(\mathcal{O}_{F_n}^\times \otimes \mathbf{C}\right)^{\chi^{-1}} \Big) + (r_1 + r_2 - 1) + 1 = r_1 + 2r_2 ,$$

where (r_1, r_2) is the signature of F. The result follows. ∎

3. Chern Classes for algebraic K-theory

In this section, we will define the Chern class map (11).

3.1. **Definitions.** In the following discussion, it will be important to carefully distinguish canonical isomorphisms from mere isomorphisms. To this end, let $\simeq$ denote an isomorphism and $=$ a canonical isomorphism. Let F be a number field, and let $\mathcal{O} := \mathcal{O}_F$ denote the ring of integers of F. The Tate twist $\mathbf{Z}_p(m)$ is the free $\mathbf{Z}_p$ module on which the Galois group G_F acts via the mth power χ^m of the cyclotomic character. For all $m \geq 1$, there exists a Chern class map:

$$c : K_{2m-1}(F) \to H^1(F, \mathbf{Z}_p(m))\,.$$

These Chern class maps arise as the boundary map of a spectral sequence, specifically, the Atiyah–Hirzebruch spectral sequence for étale K-theory. These maps were originally constructed by Soulé [28], Section II. We may compose this map with reduction mod p^i to obtain a map:

$$c : K_{2m-1}(F) \to H^1(F, \mathbf{Z}/p^i\mathbf{Z}(m)).$$

By the Chinese remainder theorem, we may also piece these maps together to obtain a map:

$$c : K_{2m-1}(F) \to H^1(F, \mathbf{Z}/n\mathbf{Z}(m))$$

for any integer n. Let ζ be a primitive nth root of unity, $F_n = F(\zeta)$ and write G for the (possibly trivial) Galois group $\mathrm{Gal}(F_n/F)$. Let μ denote the module of n roots of unity. There is a canonical injection

$$\chi : G \to \mathrm{Aut}(\mu_n) = (\mathbf{Z}/n\mathbf{Z})^\times\,.$$

By inflation–restriction, there is a canonical map:

$$H^1(F, \mathbf{Z}/n\mathbf{Z}(m)) \to H^1(F_n, \mathbf{Z}/n\mathbf{Z}(m))^G = H^1(F_n, \mathbf{Z}/n\mathbf{Z}(1))^{\chi^{1-m}}. \tag{19}$$

For $i \geq 1$, there is an invariant $w_i(F) \in \mathbf{N}$ that we will need. It is defined in terms of Galois cohomology by

$$w_i(F) \;=\; \prod_p \bigl|H^0(F, \mathbf{Q}_p/\mathbf{Z}_p(m))\bigr|\,,$$

Note that $w_1(F)$ is equal to w_F, the number of roots of unity in F, and $w_2(F)$ agrees with (10). We also define

$$\widetilde{w}_F \;=\; \prod_p \bigl|H^0(\widetilde{F}(\zeta_p + \zeta_p^{-1}), \mathbf{Q}_p/\mathbf{Z}_p(1))\bigr|\,. \tag{20}$$

where $\widetilde{F}$ is the Galois closure of F over $\mathbf{Q}$. Thus $\widetilde{w}_F$ is divisible only by the finitely many primes p such ζ_p belongs to $\widetilde{F}(\zeta_p + \zeta_p^{-1})$. If $p|\widetilde{w}_F$ and $p > 2$, then p necessarily ramifies in F. Note that $\widetilde{w}_F$ is always divisible by w_F.

Lemma 3.1. The map (19) is injective for integers n prime to $w_i(F)$.

Proof. The kernel of this map is $H^1(F_n/F, \mathbf{Z}/n\mathbf{Z}(m))$. Assume that this is non-zero. By Sah's lemma, this group is annihilated by $\chi^i(g) - 1$ for any $g \in G$. Equivalently, the kernel has order divisible by $p|n$ if and only if the elements $a^i - 1$ are divisible by p for all $(a, p) = 1$. Yet this is equivalent to saying that $H^0(F, \mathbf{Z}/p\mathbf{Z}(m)) \subset H^0(F, \mathbf{Q}_p/\mathbf{Z}_p(m))$ is non-zero, and hence $p|w_i(F)$. $\square$

There is an isomorphism $\mathbf{Z}/n\mathbf{Z}(1) = \mu$ coming from the choice of a given nth root of unity ζ. By Hilbert 90, for a number field L, there is a canonical isomorphism $H^1(L, \mu) = L^\times/L^{\times n}$, and hence c and ζ give rise to a map:

$$c_\zeta : K_{2m-1}(F) \to \left(F_n^\times/F_n^{\times n}\right)^{\chi^{1-i}} . \tag{21}$$

3.2. **The relation between étale cohomology and Galois cohomology.** There are isomorphisms that can be found in Sections 5.2 and 5.4 of [36]

$$K_{2m-1}(F) \otimes \mathbf{Z}_p \simeq K_{2m-1}(\mathcal{O}_F[1/p]) \simeq K_{2m-1}(\mathcal{O}_F) \otimes \mathbf{Z}_p$$

for $m > 1$. These isomorphisms are also reflected in the following isomorphism between étale cohomology groups and Galois cohomology groups:

$$H^1_{\text{ét}}(\mathcal{O}_F[1/p], \mathbf{Z}_p(m)) \simeq H^1(F, \mathbf{Z}_p(m))$$

for $i \geq 2$. In particular, we may also view the Chern class maps considered above as morphisms

$$c : K_{2m-1}(F) \otimes \mathbf{Z}_p \simeq K_{2m-1}(\mathcal{O}_F) \otimes \mathbf{Z}_p \to H^1_{\text{ét}}(\mathcal{O}_F[1/p], \mathbf{Z}_p(m)).$$

Theorem 3.2. *For $p > 2$, there is an isomorphism*

$$c : K_3(F) \otimes \mathbf{Z}_p \simeq K_{2m-1}(\mathcal{O}_F) \otimes \mathbf{Z}_p \to H^1(\mathcal{O}_F[1/p], \mathbf{Z}_p(2)).$$

The rank of $K_3(F)$ is r_2.

Sketch. This follows from the Quillen–Lichtenbaum conjecture, as proven by Voevodsky and Rost (see [36], [35]). In this case, it can also be deduced from the description of torsion in $K_3(F)$ by Merkur'ev and Suslin [21] (described in terms of $w_2(F)$ above) combined with Borel's theorem for the rank (see also Theorem 6.5 of [32]), and the result of Soulé that the Chern class map is surjective. □

Lemma 3.3. Suppose that $p \nmid w_2(F)$. Then the map

$$c_\zeta : K_3(F) \to K_3(F)/nK_3(F) \to \left(F_n^\times/F_n^{\times n}\right)^{\chi^{-1}} \tag{22}$$

is injective.

Proof. By the Chinese Remainder Theorem, it suffices to consider the case $n = p^m$. In light of Theorem 3.2, it suffices to show that the map

$$H^1(F, \mathbf{Z}_p(2))/n \to H^1(F, \mathbf{Z}/n\mathbf{Z}(2)) \to F_n^\times/F_n^{\times n}$$

is injective. The kernel of the first map is $H^0(F, \mathbf{Z}_p(2))/n = 0$. The kernel of the second map is, via inflation–restriction, the group $H^1(\mathrm{Gal}(F_n/F), H^0(F, \mathbf{Z}/n\mathbf{Z}(2)))$. This group is certainly zero unless

$$H^0(F, \mathbf{Z}/n\mathbf{Z}(2)) \subset H^0(F, \mathbf{Q}_p/\mathbf{Z}_p(2))$$

is non-zero, or in other words, unless p divides $w_2(F)$. □

3.3. **Upgrading from $F_n^\times$ to $\mathcal{O}_{F_n}[1/S]^\times$.** The following is a consequence of the finite generation of $K_3(F)$:

Lemma 3.4. For any field F, there exists a finite set S of primes which avoids any given finite set of primes not dividing n such that the image of c_ζ on $K_3(F)/nK_3(F)$ may be realized by an element of $\mathcal{O}_{F(\zeta)}^\times[1/S]$.

Proof. Note that without the requirement that S avoids any given finite set of primes not dividing n, the result is a trivial consequence of the fact that $K_3(F)$ is finitely generated. The construction of c as a map to units in $F_n^\times$ proceeded via Hilbert 90. In light of Theorem 3.2 above, it suffices to do the same with $H^1(F_n, \mu)$ replaced by $H^1_{\text{ét}}(\mathcal{O}_{F_n}[1/S], \mu)$ for some set S containing $p|n$. However, in this case, the class group intervenes, as there is an exact sequence ([22], p.125):

$$\mathcal{O}_{F_n}[1/S]^\times/\mathcal{O}_{F_n}[1/S]^{\times n} \to H^1_{\text{ét}}(\mathcal{O}_{F_n}[1/S], \mu) \to \text{Pic}(\mathcal{O}_{F_n}[1/S])[n]$$

where $M[n]$ denotes the n-torsion of M and Pic is the Picard group, which may be identified with the class group of $\mathcal{O}_{F_n}[1/S]$. On the other hand, it is well known that one can represent generators in the class group by a set of primes avoiding any given finite set of primes, and hence for a set S including primes for each generator of the class group, the last term vanishes. □

3.4. **Upgrading from S-units to units.** We give the following slight improvement on Lemma 3.4.

Lemma 3.5. Suppose that any prime divisor p of n is odd and divides neither the discriminant of F nor the order of $K_2(\mathcal{O}_F)$. Then the image of c_ζ on $K_3(F)/nK_3(F)$ may be realized by an element of $\mathcal{O}_n^\times$.

Proof. By Lemma 2.7, it suffices to consider the case when $n = n$ is a power of p. Let $\zeta = \zeta_n$. The fact that p is prime to the discriminant of F implies that $F(\zeta)/F$ is totally ramified at p. The image of c_ζ factors through $H^1_{\text{ét}}(\mathcal{O}[1/p], \mathbf{Z}/n\mathbf{Z}(2))$, and, via inflation–restriction, through $H^1_{\text{ét}}(\mathcal{O}_{F(\zeta)}[1/p], \mathbf{Z}/n\mathbf{Z}(1))$. The Kummer sequence for étale cohomology gives a short exact sequence:

$$\mathcal{O}_{F(\zeta)}[1/p]^\times/\mathcal{O}_{F(\zeta)}[1/p]^{\times n} \to H^1_{\text{ét}}(\mathcal{O}_{F(\zeta)}[1/p], \mathbf{Z}/n(1)) \to \text{Pic}(\mathcal{O}_F(\zeta)[1/p])[n]\,.$$

The image of c_ζ lands in the χ^{-1}-invariant part of the second group. The χ^{-1}-invariant part $M^{\chi^{-1}}$ of a G-module M is non-zero if and only if the largest χ^{-1}-invariant quotient $M_{\chi^{-1}}$ is non-zero. However, by results of Keune [18], there is an injection

$$(\text{Pic}(\mathcal{O}_F(\zeta)[1/p])/p^m)_{\chi^{-1}} \to K_2(\mathcal{O}_F)/p^m\,.$$

In particular, the pushforward of the image of c_ζ to the Picard group is trivial whenever $K_2(\mathcal{O}_F)\otimes\mathbf{Z}_p$ is trivial. Since we are assuming that p does not divide the order of $K_2(\mathcal{O}_F)$, we deduce that the image of c_ζ is realized by p-units. We now upgrade this to actual units. There is an exact sequence:

$$(\mathcal{O}_{F(\zeta)})^\times/(\mathcal{O}_{F(\zeta)})^{\times n} \to (\mathcal{O}_{F(\zeta)}[1/p])^\times/(\mathcal{O}_{F(\zeta)}[1/p])^{\times n} \to \bigoplus_{v|p} \mathbf{Z}/n\mathbf{Z}\,,$$

where the last map is the valuation map. Since p is totally ramified in $F(\zeta)/F$, the action of G on the final term is trivial. By assumption, the quotient $\mathrm{Gal}(F(\zeta_p)/F)$ is non-trivial, and hence the χ^{-1}-invariants of the final term are zero. Hence, after taking χ^{-1}-invariants, we see that the image of c_ζ comes from a unit. □

3.5. **Proof of Theorem 1.5.** We have all the ingredients to give a proof of Theorem 1.5. Fix a natural number n and a primitive nth root of unity ζ. Consider the Chern class map

$$c_\zeta : K_3(F)/nK_3(F) \to \left(F_n^\times/F_n^{\times n}\right)^{\chi^{-1}} .$$

from (22). When n is coprime to $w_2(F)$, the above map is injective by Lemma 3.3. When n is coprime to the discriminant Δ_F of F and the order of $K_2(\mathcal{O}_F)$, Lemma 3.5 implies the above map factors through a map

$$c_\zeta : K_3(F)/nK_3(F) \to \left(\mathcal{O}_n^\times/\mathcal{O}_n^{\times n}\right)^{\chi^{-1}} ,$$

where $\mathcal{O}_n$ is the ring of integers of F_n. When n is square-free and coprime to $w_2(F)$, then (9) and Proposition 2.10 imply that both sides of the above equation are abelian groups isomorphic to $(\mathbf{Z}/n\mathbf{Z})^{r_2(F)}$. It follows that when n is square-free and coprime to $w_2(F)\,\Delta_F|K_2(\mathcal{O}_F)|$, then the above map is an injection of finite abelian groups of the same order, and hence an isomorphism. This concludes the proof of Theorem 1.5. □

4. Reduction to finite fields

As we will see in Section 5, the comparison of the maps c_ζ and R_ζ and the proof of Theorem 1.6 require a reduction of both maps to the case of finite fields. In this section, we review the local Chern classes and the Bloch groups of finite fields, and introduce local (finite field) versions of the maps c_ζ and R_ζ. We will be considering the case that n is a prime power p^m, and will denote by ζ a primitive nth root of unity.

4.1. **Local Chern class maps.** Let $\mathfrak{q}$ be a prime of norm $q \equiv -1 \bmod n$ in $\mathcal{O}_F$. The residue field of $\mathcal{O}_F$ at $\mathfrak{q}$ is $\mathbf{F}_q$, and the residue field of $\mathcal{O}_{F(\zeta)}$ at a prime $\mathfrak{Q}$ above $\mathfrak{q}$ is $\mathbf{F}_{q^2} = \mathbf{F}_q(\zeta)$. Following Lemma 3.4, suppose that S does not contain any primes dividing q.

Lemma 4.1. There exists a commutative diagram of Chern class maps as follows:

$$\begin{tikzcd}
K_3(F)/nK_3(F) \arrow[r, "c_\zeta"] \arrow[d] & (\mathcal{O}[1/S]^\times_{F(\zeta)}/\mathcal{O}[1/S]^{\times n}_{F(\zeta)})^{\chi^{-1}} \arrow[d] \\
K_3(\mathbf{F}_q)/nK_3(\mathbf{F}_q) \arrow[r, equal, "c_{\zeta,\mathfrak{q}}"] & \mathbf{F}_{q^2}^\times/\mathbf{F}_{q^2}^{\times n}.
\end{tikzcd}$$

Proof. By the Chinese Reminder Theorem, we may reduce to the case when $n = p^m$. There is an isomorphism $K_3(F)\otimes\mathbf{Z}_p \simeq K_3(\mathcal{O}_F)\otimes\mathbf{Z}_p$ (see Theorem 3.2). Let $\mathcal{O}_{F,\mathfrak{q}}$ be the completion of $\mathcal{O}_F$ at $\mathfrak{q}$. We have a more general diagram as follows:

$$\begin{tikzcd}
K_3(\mathcal{O}_F)/nK_3(\mathcal{O}_F) \arrow[r, "c_\zeta"] \arrow[d] & H^1_{\text{ét}}(\mathcal{O}_F[1/p], \mathbf{Z}/n\mathbf{Z}(2)) \arrow[r] \arrow[d] & (\mathcal{O}[1/S]^\times_{F(\zeta)}/\mathcal{O}[1/S]^{\times n}_{F(\zeta)})^{\chi^{-1}} \arrow[d] \\
K_3(\mathcal{O}_{F,\mathfrak{q}}; \mathbf{Z}_p)/nK_3(\mathcal{O}_{F,\mathfrak{q}}; \mathbf{Z}_p) \arrow[r, equal, "c_{\zeta,\mathfrak{q}}"] \arrow[d, equal] & H^1_{\text{ur}}(\mathcal{O}_{F,\mathfrak{q}}, \mathbf{Z}/n\mathbf{Z}(2)) \arrow[r, equal] \arrow[d, equal] & \mathbf{F}^\times_{q^2}/\mathbf{F}^{\times n}_{q^2} \arrow[d, equal] \\
K_3(\mathbf{F}_q)/nK_3(\mathbf{F}_q) \arrow[r, equal, "c_{\zeta,\mathfrak{q}}"] & H^1(\mathbf{F}_q, \mathbf{Z}/n\mathbf{Z}(2)) \arrow[r, equal] & \mathbf{F}^\times_{q^2}/\mathbf{F}^{\times n}_{q^2}.
\end{tikzcd}$$

The image of $H^1_{\text{ét}}(\mathcal{O}_F[1/p])$ in the cohomology of $\mathcal{O}_{F,\mathfrak{q}}$ for $\mathfrak{q}$ prime to p lands in the subgroup H^1_{ur} of unramified classes. This subgroup is precisely the image of $H^1(\mathcal{O}/\mathfrak{q}, \mathbf{Z}/n\mathbf{Z}(2))$ under inflation. The maps on the right hand side of the diagram are just what one gets when unwinding the application of Hilbert's Theorem 90. The identification of the two lower horizontal lines is a reflection of Gabber rigidity, which implies that $K_3(\mathcal{O}_{F,\mathfrak{q}}; \mathbf{Z}_p) \simeq K_3(\mathbf{F}_\mathfrak{q}) \otimes \mathbf{Z}_p$. $\square$

Proposition 4.2. Let $\widetilde{F}$ denote the Galois closure of F, and suppose that $\zeta \notin \widetilde{F}(\zeta + \zeta^{-1})$. (Equivalently, suppose that n is prime to $\widetilde{w}_F$ of equation 20.)
(a) There is a map:

$$K_3(F)/nK_3(F) \xrightarrow{\bigoplus c_{\zeta,\mathfrak{q}}} \bigoplus \mathbf{F}^\times_{q^2}/\mathbf{F}^{\times n}_{q^2}$$

where the sum ranges over all primes $\mathfrak{q}$ of prime norm $q \equiv -1 \bmod n$ which split completely in F, or alternatively runs over all but finitely many primes $q \equiv -1 \bmod n$ which split completely in F.
(b) The image of this map is isomorphic to the image of the global map c_ζ, which is injective if $(n, w_2(F)) = 1$.
(c) For $\xi \in K_3(F)$, the set

$$\{\mathfrak{q} \subset \mathcal{O}_F[1/S] \mid c_{\zeta,\mathfrak{q}}(\xi) = 0\}$$

(for any finite S) determines the image of ξ up to a scalar.

Proof. It suffices to consider the case when $n = p^m$. Let $\xi \in K_3(F)$, and let the class of $c_\zeta(\xi)$ be represented by an S-unit ϵ. Because of the Galois action, this gives rise via Kummer theory to a $\mathbf{Z}/n\mathbf{Z}$-extension H of $F(\zeta+\zeta^{-1})$, and such that the reduction mod $\mathfrak{q}$ of ϵ determines the element $\mathrm{Frob}_\mathfrak{q} \in \mathrm{Gal}(H/F(\zeta+\zeta^{-1}))$. (Explicitly, we have $H(\zeta) = F(\zeta, \epsilon^{1/n})$.) Hence our assumptions imply that any prime q which splits completely in $F(\zeta+\zeta^{-1})$ (which forces $q \equiv \pm 1 \bmod n$) and is additionally congruent to $-1 \bmod p$ must split in H. Let $\widetilde{H}$ denote the Galois closure of H over $\mathbf{Q}$, and $\widetilde{F}$ the Galois closure of F over $\mathbf{Q}$. Note that the Galois closure of $F(\zeta+\zeta^{-1})$ is $\widetilde{F}(\zeta+\zeta^{-1})$. A prime q splits completely in H if and only if it splits completely in $\widetilde{H}$, and splits completely in $F(\zeta+\zeta^{-1})$ if and only if it splits completely in $\widetilde{F}(\zeta+\zeta^{-1})$. We have a diagram of fields as follows:

$$\begin{tikzcd}
\widetilde{H} \arrow[r, dash] \arrow[d, dash] & \widetilde{H}(\zeta) \arrow[d, dash] \\
\widetilde{F}(\zeta+\zeta^{-1}) \arrow[r, dash] & \widetilde{F}(\zeta)
\end{tikzcd}$$

By assumption, we have $\zeta \notin \widetilde{F}(\zeta+\zeta^{-1})$. Since $H/F(\zeta+\zeta^{-1})$ is cyclic of degree n, it follows that $\mathrm{Gal}(\widetilde{H}/\widetilde{F}(\zeta+\zeta^{-1}))$ is an abelian p-group. On the other hand, $\mathrm{Gal}(\widetilde{F}(\zeta)/\widetilde{F}(\zeta+\zeta^{-1})) = \mathbf{Z}/2\mathbf{Z}$. so $\mathrm{Gal}(\widetilde{H}(\zeta)/\widetilde{F}(\zeta+\zeta^{-1}))$ is the direct sum of $\mathbf{Z}/2\mathbf{Z}$ with a p-group, Let $\sigma \in \mathrm{Gal}(\widetilde{H}(\zeta)/\widetilde{F}(\zeta+\zeta^{-1})) \subset \mathrm{Gal}(\widetilde{H}/\mathbf{Q})$ denote an element of order $2p$. By the Cebotarev density theorem, there exist infinitely many primes $q \in \mathbf{Q}$ with Frobenius element in $\mathrm{Gal}(\widetilde{H}/\mathbf{Q})$ corresponding to σ. By construction, the prime q splits completely in $\widetilde{F}(\zeta+\zeta^{-1})$ because the corresponding Frobenius element is trivial in $\mathrm{Gal}(\widetilde{F}(\zeta+\zeta^{-1})/\mathbf{Q})$. On the other hand, since σ has order divisible by 2 and by p, it is non-trivial in both $\mathrm{Gal}(\widetilde{F}(\zeta)/\widetilde{F}(\zeta+\zeta^{-1})) = \mathrm{Gal}(\mathbf{Q}(\zeta)/\mathbf{Q}(\zeta+\zeta^{-1}))$ and $\mathrm{Gal}(\widetilde{H}/\widetilde{F}(\zeta+\zeta^{-1}))$. The first condition implies that $q \equiv -1 \bmod n$, and the second condition implies that q does not split completely in H, a contradiction. The injectivity (under the stated hypothesis) follows from Lemma 3.1. $\square$

Remark 4.3. The condition that $\zeta \notin \widetilde{F}(\zeta+\zeta^{-1})$ is automatic if p is unramified in F, because then the ramification degree of $\mathbf{Q}(\zeta)$ is $p-1$ whereas the ramification degree of $\widetilde{F}(\zeta+\zeta^{-1})$ is $(p-1)/2$ for p odd. If $\zeta \in \widetilde{F}(\zeta+\zeta^{-1})$, then there are no primes q which split completely in F and have norm $-1 \bmod n$. In particular, when $\zeta \in \widetilde{F}(\zeta+\zeta^{-1})$, we have $B(\mathbf{F}_q)\otimes\mathbf{F}_p = 0$ for every prime q which splits completely in F.

4.2. **The Bloch group of $\mathbf{F}_q$.** In order to make our maps explicit, we must relate the Chern class map to the Bloch group. Let $p > 2$ and $q > 2$ be odd primes such that $q \equiv -1 \bmod n$, where $n = p^m$. For a finite field $\mathbf{F}_q$, the group $\mathbf{F}_q^\times$ is cyclic, so $\bigwedge^2 \mathbf{F}_q^\times$ is a 2-torsion group. Hence the Bloch group $B(\mathbf{F}_q)$ coincides with the pre-Bloch group after tensoring with $\mathbf{F}_p$, where the pre-Bloch group is defined as the quotient of the free abelian group on $\mathbf{F}_q \smallsetminus \{0,1\}$ by the 5-term relation. By [16], the Bloch group $B(\mathbf{F}_q)$ is a cyclic group of order $q+1$ up to 2-torsion. Moreover, following [16], one may relate $B(\mathbf{F}_q)$ to the cohomology of $\mathrm{SL}_2(\mathbf{F}_q)$ in degree three, as we now discuss.

There is an isomorphism

$$H_3(\mathrm{SL}_2(\mathbf{F}_q),\mathbf{Z})\otimes\mathbf{Z}/n\mathbf{Z} \;\simeq\; \mathbf{Z}/n\mathbf{Z}.$$

Let us describe this isomorphism more carefully. By a computation of Quillen, we know that $H^3(\mathrm{SL}_2(\mathbf{F}_q),\mathbf{Z})$ is cyclic of order q^2-1. It follows that the p-part of this group comes from the p-Sylow subgroup. If one chooses an isomorphism

$$\mathbf{F}_{q^2} \simeq (\mathbf{F}_q)^2$$

of abelian groups, one gets a well defined map:

$$\mathbf{F}_{q^2}^\times \;=\; C \;=\; \mathrm{Aut}_{\mathbf{F}_q}(\mathbf{F}_{q^2}) \to \mathrm{GL}_2(\mathbf{F}_q)$$

which is well defined up to conjugation. There is, correspondingly, a map $C^1 \to \mathrm{SL}_2(\mathbf{F}_q)$, where

$$C^1 \;=\; \mathrm{Ker}\big(N : \mathbf{F}_{q^2}^{\times} \to \mathbf{F}_q^{\times}\big)\,.$$

We refer to both C and C^1 as the non-split Cartan subgroup. By Quillen's computation, we deduce that there is a canonical map:

$$C^1 \;=\; H^3(C^1,\mathbf{Z}) \to H^3(\mathrm{SL}_2(\mathbf{F}_q),\mathbf{Z})$$

which is an isomorphism after tensoring with $\mathbf{Z}/n\mathbf{Z}$. There is a canonical isomorphism $C^1[n] = \mu$, where μ denotes the nth roots of unity. Hence to give an element of order p in $H^3(\mathrm{SL}_2(\mathbf{F}_q),\mathbf{Z})$ up to conjugation is equivalent to giving a primitive nth root of unity $\zeta \in C^1 \subset C = \mathbf{F}_{q^2}^{\times}$. From [16], there is a canonical map:

$$H^3(\mathrm{SL}_2(\mathbf{F}_q),\mathbf{Z}) \to B(\mathbf{F}_q)\,,$$

at least away from 2-power torsion, which is an isomorphism after tensoring with $\mathbf{Z}/n\mathbf{Z}$. Given a root of unity ζ, let t denote the corresponding element of $\mathrm{SL}_2(\mathbf{F}_q)$. The corresponding element of $B(\mathbf{F}_q)$, up to six-torsion, is given (see [16], p.36) by:

$$\sum_{k=1}^{n-1}\left[\frac{t(\infty)-t^{k+1}(\infty)}{t(\infty)-t^{k+2}(\infty)}\right]\,.$$

This construction yields the same element for ζ and ζ^{-1}. We may represent t by its conjugacy class in $\mathrm{GL}_2(\mathbf{F}_q)$, which has determinant one and trace $\zeta+\zeta^{-1} \in \mathbf{F}_q$. The choice of ζ up to (multiplicative) sign is given by this trace. Note that the congruence condition on q ensures that the Chebyshev polynomial with roots $\zeta+\zeta^{-1}$ has distinct roots which split completely over $\mathbf{F}_q$. Explicitly, we may choose

$$t \;=\; \begin{pmatrix} 0 & 1 \\ -1 & \zeta+\zeta^{-1}\end{pmatrix} = A\begin{pmatrix}\zeta^{-1} & 0 \\ 0 & \zeta\end{pmatrix}A^{-1}, \qquad A \;=\; \begin{pmatrix}\zeta & \zeta^{-1} \\ 1 & 1\end{pmatrix}.$$

Let F_k be the Chebyshev polynomials, so $F_k(2\cos\phi) = \dfrac{\sin k\phi}{\sin\phi}$. Then

$$t^k(\infty) \;=\; \frac{F_{k-1}(\zeta+\zeta^{-1})}{F_k(\zeta+\zeta^{-1})}\,,$$

and an elementary computation then shows that the corresponding element in $B(\mathbf{F}_q)\otimes\mathbf{Z}_p$ is given by

$$\sum_{k=1}^{n-1}\left[1-\frac{1}{F_k(\zeta+\zeta^{-1})^2}\right] \;\sim\; \sum_{k=1}^{n-1}\left[\left(\frac{\zeta^k-\zeta^{-k}}{\zeta-\zeta^{-1}}\right)^2\right]\,,$$

where $\sim$ denotes equality in $B(\mathbf{F}_q)\otimes\mathbf{Z}_p$, since $[x]=[1-1/x]$ up to 3-torsion. (When $p=3$, one may verify directly that the latter term is also a 3-torsion element.)

4.3. **The local Chern class map c_ζ.** In this section, q will denote a prime with $q \equiv -1 \bmod p^m$ which splits completely in F. Let $\mathfrak{q}$ be a prime above q. There is a natural map $B(F) \to B(\mathcal{O}_F/\mathfrak{q}) = B(\mathbf{F}_q)$. The elements $[0]$, $[1]$, and $[\infty]$ are trivial elements of $B(F)$ and $B(\mathbf{F}_q)$; the reduction map then sends $[x]$ to $[\overline{x}]$ under the natural reduction map $\mathbf{P}^1(F) \to \mathbf{P}^1(\mathbf{F}_q)$.

Lemma 4.4. Let $p > 2$. There is a commutative diagram as follows:

$$\begin{tikzcd} B(F)/nB(F) \arrow[r] \arrow[d, equal] & \bigoplus B(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z} \arrow[d, dash, "\simeq"] & \\ K_3(F)/nK_3(F) \arrow[r] & \bigoplus K_3(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z} \arrow[r, equal, "\bigoplus c_{\zeta,\mathfrak{q}}"] & \bigoplus \mathbf{F}_{q^2}^\times/\mathbf{F}_{q^2}^{\times n}, \end{tikzcd}$$

where the product runs over all primes $\mathfrak{q}$ of norm $q \equiv -1 \bmod n$ which split completely in F, or alternatively all but finitely many such primes.

Proof. The isomorphism of the left vertical map is a theorem of Suslin [31], and the isomorphism of the right vertical map follows from [16]. The fact that the diagram commutes is a consequence of the fact that both constructions are compatible (and can be seen in group cohomology). ☐

Recall that an element x of an abelian group G is p-saturated if $x \notin [p]G$, where $[p] : G \to G$ is the multiplication by p map.

Corollary 4.5. There is an algorithm to prove that a set of generators of $B(F)$ is p-saturated for $p > 2$.

Proof. Computing $B(\mathbf{F}_q)$ is clearly algorithmically possible. Moreover, we can *a priori* compute $B(F) \otimes \mathbf{Z}_p$ as an abstract $\mathbf{Z}_p$-module. Hence it suffices to find sufficiently many distinct primes $\mathfrak{q}$ such that the image of a given set of generators has the same order as $B(F)/nB(F)$. ☐

In light of the commutative diagram of Lemma 4.4, we also use c_ζ to denote the Chern class map on $B(F)/nB(F)$.

4.4. **The local R_ζ map.** Suppose that $q \equiv -1 \bmod p$. It follows that the field $\mathbf{F}_q$ does not contain ζ_p, and so Proposition 2.4 applies to give maps P_ζ and R_ζ which are well defined over this field. In particular, since $(p, q-1) = 1$, all elements of $\mathbf{F}_q$ are p-th powers, and hence the Kummer extension H is given by $H = F_n$ and R_ζ and P_ζ coincide.

5. Comparison between the maps c_ζ and R_ζ

The main goal of this section, carried out in the first subsection, is to prove Theorem 1.6. The main result here is Theorem 5.2, which says that that our mod n local regulator map $R_{\zeta,q}$ gives an isomorphism from $B(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z}$ to $\mathbf{F}_{q^2}^\times \otimes \mathbf{Z}/n\mathbf{Z}$ for any prime power n and prime $q \equiv -1 \pmod{n}$. This implies in particular the existence of a curious "mod-p-q dilogarithm map" from $\mathbf{F}_q$ to $\mathbf{Z}/n\mathbf{Z}$, and in Section 5.2, we digress briefly to give an explicit formula for this map. In the final subsection, we describe the expected properties of the Chern class map that would imply the conjectural equality (13) and hence, in conjunction with (12), the evaluation $\gamma = 2$ of the comparison constant γ occurring in Theorem 1.6.

5.1. **Proof of Theorem 1.6.** Throughout this section, we set $n = p^m$, and let ζ denotes a primitive nth root of unity. For a prime $q \equiv -1 \bmod n$ that splits completely in F, and for a corresponding prime $\mathfrak{q}$ above q, let $R_{\zeta,\mathfrak{q}}$ denote the map $B(\mathcal{O}_F/\mathfrak{q}) = B(\mathbf{F}_q) \to \mathbf{F}_{q^2}^\times/\mathbf{F}_{q^2}^{\times n}$.

We have two maps we wish to compare. One of them is

$$c_\zeta : \ B(F)/nB(F) \ \to \ \left(F_n^\times/F_n^{\times n}\right)^{\chi^{-1}}.$$

Because $B(F)$ is a finitely generated abelian group, we may represent the generators of the image by S-units for some fixed S (at this point possibly depending on n) and consider the map

$$c_\zeta : \ B(F)/nB(F) \to (\mathcal{O}_{F(\zeta)}[1/S]^\times/\mathcal{O}_{F(\zeta)}[1/S]^{\times n})^{\chi^{-1}} \hookrightarrow \bigoplus \mathbf{F}_{q^2}^\times/\mathbf{F}_{q^2}^{\times n} \simeq \bigoplus B(\mathbf{F}_q),$$

where the final sum is over all but finitely may primes $\mathfrak{q}$ of norm $q \equiv -1 \bmod n$ which split completely in F. We have the diagram

$$\begin{tikzcd}
B(F)/nB(F) \arrow[r, "R_\zeta"] \arrow[d] & (\mathcal{O}_{F(\zeta)}[1/S]^\times/\mathcal{O}_{F(\zeta)}[1/S]^{\times n})^{\chi^{-1}} \arrow[d] \\
\bigoplus B(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z} \arrow[r, "R_{\zeta,\mathfrak{q}}"] & \bigoplus \mathbf{F}_{q^2}^\times/\mathbf{F}_{q^2}^{\times n}.
\end{tikzcd}$$

We have already shown, by Cebotarev (Proposition 4.2(b)), that $c_\zeta(\xi)$ for $\xi \in K_3(F)$ is determined up to scalar by the set of primes for which $c_{\zeta,\mathfrak{q}}(\xi) = 0$. Hence the result is a formal consequence of knowing that the maps $R_{\zeta,\mathfrak{q}}$ are isomorphisms for all $\mathfrak{q}$ of norm $q \equiv -1 \bmod n$. This is exactly Theorem 5.2 below. □

By (9), the p-torsion subgroup of $K_3(\mathbf{Q}(\zeta + \zeta^{-1}))$ is isomorphic to $\mathbf{Z}/n\mathbf{Z}$. On the other hand, since $\mathbf{Q}(\zeta + \zeta^{-1})$ is totally real, we have an isomorphism:

$$K_3(\mathbf{Q}(\zeta + \zeta^{-1})) \otimes \mathbf{Z}_p \simeq \mathbf{Z}/n\mathbf{Z}.$$

Lemma 5.1. Let $p > 2$ and $n = p^m$. Suppose that $q \equiv -1 \bmod n$ and $q \not\equiv -1 \bmod pn$. The prime q splits completely in $\mathbf{Q}(\zeta + \zeta^{-1})$. Let $\mathbf{F}_q$ denote the residue field at one of the primes above q. Then the map

$$K_3(\mathbf{Q}(\zeta + \zeta^{-1})) \otimes \mathbf{Z}_p \to B(\mathbf{F}_q) \otimes \mathbf{Z}_p$$

is an isomorphism.

Proof. A generator of $B(\mathbf{Q}(\zeta + \zeta^{-1}))[n] \simeq K_3(\mathbf{Q}(\zeta + \zeta^{-1})) \otimes \mathbf{Z}_p$ is given explicitly by the element

$$\eta_\zeta \ := \ \sum_{\ell=1}^{n-1} \left[\left(\frac{\zeta^\ell - \zeta^{-\ell}}{\zeta - \zeta^{-1}} \right)^2 \right] \tag{23}$$

This follows from Theorem 1.3 of [40]. On the other hand, the reduction modulo any prime above q generates the latter group, as follows from the discussion in Section 4.2. $\square$

We now prove Theorem 5.2 as mentioned above:

Theorem 5.2. *Let n be a prime power and $q \equiv -1 \bmod n$. Then the map*

$$R_{\zeta,q} : B(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z} \to \mathbf{F}_{q^2}^{\times} \otimes \mathbf{Z}/n\mathbf{Z}$$

is an isomorphism, where ζ is an nth root of unity.

Proof. Note that $B(\mathbf{F}_q)$ is cyclic of order $q+1$ up to 2-torsion, and $\mathbf{F}_{q^2}^{\times}$ is cyclic of order q^2-1. In particular, for odd primes p with $q \equiv -1 \bmod p$, the groups $B(\mathbf{F}_q) \otimes \mathbf{Z}_p$ and $\mathbf{F}_{q^2}^{\times} \otimes \mathbf{Z}_p$ are isomorphic to each other and to $\mathbf{Z}_p/(q+1)\mathbf{Z}_p$. We begin with the following:

Lemma 5.3. $R_\zeta(\eta_\zeta) = \zeta^\gamma \in (\mathbf{Q}(\zeta)^\times/\mathbf{Q}(\zeta)^{\times n})^{\chi^{-1}}$ for some $\gamma \in \mathbf{Z}_p$.

Proof. Write $\zeta_n = \zeta$ and let ζ' be an n^2th root of unity. Consider the image of $R_{\zeta'}(\eta_{\zeta'})$. Because η_ζ is divisible by n in $B(\mathbf{Q}(\zeta')^+)$, the image is a nth power. Hence, by the compatibility of the maps R for varying n (Lemma 2.6 (2)), it follows that $R_\zeta(\eta_\zeta)$ lies in the kernel of the map

$$\left(\mathbf{Q}(\zeta)^\times/\mathbf{Q}(\zeta)^{\times n}\right)^{\chi^{-1}} \to \left(\mathbf{Q}(\zeta')^\times/\mathbf{Q}(\zeta')^{\times n}\right)^{\chi^{-1}}.$$

But this kernel consists precisely of nth roots of unity. $\square$

Let $\eta_{\zeta,q} \in B(\mathbf{F}_q)$ denote the reduction of η_ζ in $B(\mathbf{F}_q)$. By Lemma 5.1, the image also generates $B(\mathbf{F}_q) \otimes \mathbf{Z}/n\mathbf{Z}$. Since all primes $q \equiv -1 \bmod n$ split completely in $\mathbf{Q}(\zeta)^+$, if $\gamma \not\equiv 0 \bmod p$, the result above follows by specialization. We proceed by contradiction and assume that $\gamma \equiv 0 \bmod p$, which means that the image of the map $P_{\zeta,\mathfrak{q}}$ is divisible by p for all $\mathfrak{q}$ of norm q satisfying $q \equiv -1 \bmod n$. In particular, to prove the result, it suffices to find a *single* such $\mathfrak{q}$ for which $R_{\zeta,\mathfrak{q}}$ is an isomorphism.

Choose a completely split prime $\mathfrak{r}$ in $\mathbf{Q}(\zeta)$. Assume that

$$\zeta \equiv a^{-1} \bmod \mathfrak{r}, \qquad \zeta \not\equiv a^{-1} \bmod \mathfrak{r}^2$$

for some integer $a \neq 1$. The splitting assumption means that an a satisfying the first condition exists, replacing a^{-1} by $(a+N(\mathfrak{r}))^{-1}$ if necessary implies the second, because

$$\frac{1}{a} - \frac{1}{a+N(\mathfrak{r})} = \frac{N(\mathfrak{r})}{a(a+N(\mathfrak{r}))} \not\equiv 1 \bmod \mathfrak{r}^2.$$

Let

$$\tau = \prod_{k=0}^{n-1}(1-\zeta^k a)^k \in \mathbf{Q}(\zeta)^\times.$$

Lemma 5.4. $\tau \cdot \zeta^i$ is not a perfect pth power for any i.

Proof. The assumption on $\mathfrak{r}$ implies that all the pth roots of unity are distinct modulo $\mathfrak{r}$, and hence the only factor of τ divisible by $\mathfrak{r}$ is $(1-a\zeta)$, which has valuation one. $\square$

The element τ gives rise, via Kummer theory, to a $\mathbf{Z}/n\mathbf{Z}$-extension $F/\mathbf{Q}(\zeta)^+$. By the Lemma above, it is non-trivial. Let $q \equiv -1 \bmod n$ be prime. Then, for a prime $\mathfrak{q}$ above q, the element $\mathrm{Frob}_{\mathfrak{q}} \in \mathrm{Gal}(F/\mathbf{Q}(\zeta)^+)$ fails to generate $\mathbf{Z}/n\mathbf{Z}$ if and only if τ is a perfect pth power modulo $\mathfrak{q}$. This is equivalent to saying that $\mathrm{Frob}_{\mathfrak{q}}$ generates $\mathrm{Gal}(F/\mathbf{Q}(\zeta)^+)$ if and only if

$$R_{\zeta,\mathfrak{q}}([a^n]) = P_{\zeta,\mathfrak{q}}([a^n]) = \prod_{k=0}^{n-1}(1-a\zeta^k)^k \in \mathbf{F}_{q^2}^\times \otimes \mathbf{Z}/n\mathbf{Z}$$

is a generator. Hence it suffices to find a single $q \equiv -1 \bmod n$ and $q \not\equiv -1 \bmod np$ with the desired Frobenius. Such a q exists by Cebotarev density unless $\langle\tau\rangle = \langle\zeta\rangle \bmod \mathbf{Q}(\zeta)^{\times p}$. However, this cannot happen by Lemma 5.4. ∎

Proof of Theorem 1.2. Assume that n is prime to $w_2(F)$. It follows that the Chern class map gives an injection

$$K_3(F)/nK_3(F) \to \mathcal{O}_{F_n}[1/S]^\times/\mathcal{O}_{F_n}[1/S]^{\times n}$$

for some finite set of primes S. If, in addition, we assume that p does not divide $\widetilde{w}_F$, then we deduce from Proposition 4.2 that this map can be extended to an injection into the group $\bigoplus_{\mathfrak{q}} B(\mathbf{F}_q)/nB(\mathbf{F}_q)$. By Theorem 1.5, this agrees with the map R_ζ defined on $B(F)$, which is thus injective. If one additionally assumes that n is prime to $|\Delta_F||K_2(\mathcal{O}_F)|$, then by Lemma 3.4 one may additionally assume that the image is precisely the χ^{-1}-invariants of $\mathcal{O}_{F_n}^\times/\mathcal{O}_{F_n}^{\times n}$. ∎

5.2. **Digression: the mod-p-q dilogarithm.** Let q be prime, and $q+1 \equiv 0 \bmod n$ with n a power of p as before. Fix an nth root of unity ζ in $\mathbf{F}_{q^2}$. Then there is a trivialization $\log_\zeta : \mathbf{F}_{q^2}^\times \otimes \mathbf{Z}/n\mathbf{Z} \simeq \mathbf{Z}/n\mathbf{Z}$ sending ζ to 1. The isomorphism $B(\mathbf{F}_q)\otimes\mathbf{Z}_p \simeq \mathbf{Z}/n\mathbf{Z}$ of Theorem 5.2 now gives a curious function, the p-q *dilogarithm*, which is a function

$$L:\ \mathbf{F}_q \ \to\ \mathbf{F}_{q^2}^\times \otimes \mathbf{Z}/n\mathbf{Z} \ \overset{\log_\zeta}{\to}\ \mathbf{Z}/n\mathbf{Z}$$

satisfying the 5-term relation. What is perhaps surprising is that the quantum *logarithm* suffices to give an explicit formula, as follows.

Proposition 5.5. The function L is given by the formula

$$L(a) \ = \ \sum_{b^n=a} \log_\zeta(b)\log_\zeta(1-b) \qquad (a \in \mathbf{F}_q^\times),$$

where the sum is over the nth roots b of a in $\mathbf{F}_{q^2}^\times$.

Proof. Since $\mathbf{F}_q^\times$ has order prime to n, the element a has a unique nth power $c \in \mathbf{F}_q^\times$. Then (17) can be rewritten as $L(a) = \sum_{k \bmod n} k\log_\zeta(1-\zeta^k c)$. (Note that $R_\zeta = P_\zeta$ for finite fields.) The elements $b = \zeta^k c$ are the nth roots of a in $\mathbf{F}_{q^2}^\times$, and $\log_\zeta(b) = k$ because c has order prime to n and thus $\log_\zeta(c) = 0$. ∎

5.3. **The Chern class map on n-torsion in $\mathbf{Q}(\zeta)^+$.** (The following section contains a speculative digression and is not used elsewhere in the paper.) We have proved that the maps c_ζ and R_ζ agree up to an invertible element of $\mathbf{Z}_p^\times$. To determine the value of this ratio, whose conjectural value is 2, we need to compute the images of specific elements of the Bloch group. More specifically, as explained in the introduction, we need the two statements (12) and (13). The first of these will be proved below (Theorem 8.5). Here we want to show that the second is not pure fancy. We shall give a heuristic justification of why the image of the Chern class map on η_ζ should be ζ — at least up to a sign and a small power of 2 in the exponent. We hope that the arguments of this section could, with care, be made into a precise argument. However, since the main conjecture of this section is somewhat orthogonal to the main purpose of this paper, and correctly proving everything would (at the very least) involve establishing that several diagrams relating the cohomology of SL_2 and PSL_2 and GL_2 and PGL_2 commute up to precise signs and factors of 2. Thus we content ourselves with a sketch, and enter the happy land where all diagrams commute.

The first subtle point is that the relation between $K_3(F)$ and $B(F)$ as established by Suslin is not an isomorphism. There is always an issue with 2-torsion coming from the image of Milnor K_3. However, even for primes p away from 2, there is an exact sequence of Suslin ([31], Theorem 5.2; here F is a number field so certainly infinite):

$$0 \to \mathrm{Tor}_1(\mu_F, \mu_F) \otimes \mathbf{Z}[1/2] \to K_3(F) \otimes \mathbf{Z}[1/2] \to B(F) \otimes \mathbf{Z}[1/2] \to 0,$$

and hence when $p | w_F = |\mu_F|$, the comparison map is not an isomorphism. (This is one of the headaches which implicitly us to assume that $\zeta \notin F$ when computing the Chern class map on $B(F)$.) This issue arises in the following way. Over the field $\mathbf{Q}(\zeta)$, the Bott element provides a direct relationship between $K_1(F, \mathbf{Z}/n\mathbf{Z})$ and $K_3(\mathbf{Z}, \mathbf{Z}/n\mathbf{Z})$. This suggests we should push forward η_ζ to $\mathbf{Q}(\zeta)$ and compute the Chern class there. However, since in $B(\mathbf{Q}(\zeta))$, the class η_ζ may (and indeed does) become trivial, we instead consider η_ζ as an element of $K_3(\mathbf{Q}(\zeta))$, and then compute the Chern class map directly in K-theory.

By Theorem 4.10 of Dupont–Sah [7], the diagonal map

$$x \to \begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$$

induces an injection

$$\mu_{\mathbf{C}} \simeq H_3(\mu_{\mathbf{C}}, \mathbf{Z}) \to H_3(\mathrm{SL}_2(\mathbf{C}), \mathbf{Z})$$

whose image is precisely the torsion subgroup. (We shall be more precise about this first isomorphism below.) Let n be odd, and let ζ be a primitive nth root of unity, let $E = \mathbf{Q}(\zeta)$, and let $E^+ = \mathbf{Q}(\zeta)^+$. If μ_E is the group of nth roots of unity, the map $\mu_E \to \mathrm{SL}_2(E)$ is conjugate to a map

$$\mu_E \to \mathrm{SL}_2(E^+)$$

as follows; send ζ to

$$t = A \begin{pmatrix} \zeta^{-1} & 0 \\ 0 & \zeta \end{pmatrix} A^{-1}, \quad \text{where } A = \begin{pmatrix} \zeta & \zeta^{-1} \\ 1 & 1 \end{pmatrix}.$$

The cohomology of μ_E with coefficients in $\mathbf{Z}/n\mathbf{Z}$ is (non-canonically) isomorphic to $\mathbf{Z}/n\mathbf{Z}$ in all degrees. More precisely, there is a canonical isomorphism

$$H_1(\mu_E, \mathbf{Z}) = H_1(\mu_E, \mathbf{Z}/n\mathbf{Z}) = \mu_E\,,$$

we have $H_2(\mu_E, \mathbf{Z}) = 0$, and thus via the Bockstein map $H_2(\mu_E, \mathbf{Z}/n\mathbf{Z}) = H_1(\mu_E, \mathbf{Z})[n] = \mu_E$. A choice of ζ leads to a choice of element $\beta \in H_2(\mu_E, \mathbf{Z}/n\mathbf{Z}) = \mu_E$, and hence to an isomorphism

$$\mu_E = H_1(\mu_E, \mathbf{Z}/n\mathbf{Z}) \xrightarrow{*\beta} H_3(\mu_E, \mathbf{Z}/n\mathbf{Z}) = H_3(\mu_E, \mathbf{Z})$$

where the isomorphism is given by the Pontryagin product of μ_E with $\beta \in H_2(\mu_E, \mathbf{Z}/n\mathbf{Z})$. These choices induce a map

$$\mu_E \to H_3(\mu_E, \mathbf{Z}) \to H_3(\mathrm{SL}_2(E^+), \mathbf{Z}) \to K_3(E^+) \to B(E^+)$$

which sends ζ to η_ζ. That the image of ζ is η_ζ follows (for example) by §8.1 of [40]). Implicit in this statement also is that the Pontryagin product of $1 \in \mathbf{Z}/n\mathbf{Z} = H_1(\mathbf{Z}/n\mathbf{Z}, \mathbf{Z}/n\mathbf{Z})$ with $1 \in H_2(\mathbf{Z}/n\mathbf{Z}, \mathbf{Z}/n\mathbf{Z})$ is exactly the class constructed in Proposition 3.25 of Parry and Sah [27]. (The maps above are only properly defined modulo 2-torsion, since μ has odd order this issue can safely be ignored). Denote by η_{E^+} the corresponding element in $K_3(E^+)$. The Chern class maps are compatible with base change, so to compute $c(\eta_{E^+})$ it suffices to compute $c(\eta_E)$ where $\eta_E \in K_3(E)$ is the image of η_{E^+} under the map $K_3(E^+) \to K_3(E)$. The Chern class map on $K_1(E) = E^\times$ canonically sends $\zeta \in E^\times$ to ζ; we would like to directly connect the Chern class map on K_1 with the one on K_3 using the Bott element. The Bott element $\beta \in K_2(E; \mathbf{Z}/n\mathbf{Z})$ is defined as follows. There is an isomorphism:

$$\mu_E = \ker\left(E^\times \xrightarrow{n} E^\times\right) = \pi_2(E^\times; \mathbf{Z}/n\mathbf{Z})\,.$$

The element β is defined as the image of ζ under the composition

$$\pi_2(\mathrm{BGL}_1(E); \mathbf{Z}/n\mathbf{Z}) \to \pi_2(\mathrm{BGL}(E); \mathbf{Z}/n\mathbf{Z}) \to \pi_2(\mathrm{BGL}(E)^+; \mathbf{Z}/n\mathbf{Z}) = K_2(E; \mathbf{Z}/p\mathbf{Z})\,.$$

The Bott element induces an isomorphism:

$$\beta : K_1(E; \mathbf{Z}/n\mathbf{Z}) \to K_3(E; \mathbf{Z}/n\mathbf{Z})\,.$$

Hence there is, given our choice of $\zeta \in E$, a canonically defined map:

$$\begin{tikzcd}
K_3(E) \arrow[d] & E^\times/E^{\times n} \arrow[dd, equal] \\
K_3(E; \mathbf{Z}/n\mathbf{Z}) \arrow[ur, "c_\zeta"] \arrow[d, equal, "\beta^{-1}"'] & \\
K_1(E, \mathbf{Z}/n\mathbf{Z}) \arrow[r, "c"] & E^\times/E^{\times n}
\end{tikzcd}$$

Here c_ζ is the composition of the Chern class map to $H^1(E, \mathbf{Z}/n\mathbf{Z}(2))$ which can be identified with $E^\times/E^{\times n}$ after a choice of $\zeta \in E$. Note that the definition of β also requires a similar choice. Thus it makes sense to make the following:

Assumption 5.6. The diagram above commutes.

We believe that it should be possible to prove this assumption, at least up to a choice of sign and a power of 2.

Using Assumption 5.6, we would like to show that $c_\zeta(\eta_E) = \zeta$, and hence that $c_\zeta(\eta_{E^+})$ and thus $c_\zeta(\eta_\zeta)$ are also both equal to ζ. This will follow if, under the Bott element, the class η_E corresponds to $\zeta \in K_1(E; \mathbf{Z}/n\mathbf{Z})$. To prove this, one roughly has to show that the following square commutes:

$$\begin{tikzcd}
\mu_E = H_1(\mu_E, \mathbf{Z}/n\mathbf{Z}) \arrow[r, "*\beta"] \arrow[d] & H_3(\mu_E, \mathbf{Z}/n\mathbf{Z}) \arrow[d] \\
E^\times/E^{\times n} = K_1(E, \mathbf{Z}/n\mathbf{Z}) \arrow[r, "\beta"] & K_3(E; \mathbf{Z}/n\mathbf{Z}).
\end{tikzcd}$$

The top line comes from the Pontryagin product structure of $H_1(\mu_E, \mathbf{Z}/n\mathbf{Z}) = \mu_E$ with

$$H_2(\mu_E, \mathbf{Z}/n\mathbf{Z}) = \ker(\mu_E \xrightarrow{[n]} \mu_E)\,,$$

and the bottom line comes from Pontryagin product with the Bott element β coming via the Bockstein map from

$$\ker(E^\times \xrightarrow{[n]} E^\times)\,.$$

We conveniently denote both maps by essentially the same letter in order to be more suggestive. One caveat is that the maps from $E^\times \to \mathrm{GL}_2(E)$ and $\mu_E \to \mathrm{SL}_2(E)$ considered above differ slightly in that x is sent to $\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$ respectively; since n is odd such maps can be compared by comparing the cohomologies of GL, PGL, SL, and PSL respectively; it is quite possible that such comparisons might require that the maps above include a factor of 2 or -1 at some point.

The above discussion above makes the conjectured equation (13) plausible.

6. The connecting homomorphism to K-theory

In this section, we give a proof of Theorem 1.7. Assume that F is a field of characteristic prime to p which does not contain a pth root of unity. Recall that $Z(F)$ is the free abelian group on $F \smallsetminus \{0, 1\}$ and $C(F)$ the subgroup generated by the 5-term relation.

Definition 6.1. Let $A(F; \mathbf{Z}/n\mathbf{Z})$ be the kernel of the map

$$d : Z(F) \longrightarrow \textstyle\bigwedge^2 F^\times \otimes \mathbf{Z}/n\mathbf{Z}\,, \qquad [X] \mapsto X \wedge (1 - X)\,.$$

The étale Bloch group $B(F; \mathbf{Z}/n\mathbf{Z})$ is the quotient

$$B(F; \mathbf{Z}/n\mathbf{Z}) = A(F; \mathbf{Z}/n\mathbf{Z})/(nZ(F) + C(F))\,.$$

It is annihilated by n.

There is a tautological exact sequence

$$0 \to B(F) \to Z(F)/C(F) \to \bigwedge^2 F^\times \to K_2(F) \to 0\,.$$

For appropriately defined R, we may break this into the two short exact sequences as follows:

$$\begin{tikzcd}
0 \arrow[r] & A(F) \arrow[r] \arrow[d] & Z(F) \arrow[r] \arrow[d] & R \arrow[r] \arrow[d, equal] & 0, \\
0 \arrow[r] & B(F) \arrow[r] & Z(F)/C(F) \arrow[r] & R \arrow[r] & 0,
\end{tikzcd}$$

$$0 \to R \to \bigwedge^2 F^\times \to K_2(F) \to 0. \tag{24}$$

Similarly, for some Q, we have corresponding short exact sequences:

$$\begin{tikzcd}
0 \arrow[r] & A(F) \arrow[r] \arrow[d] & A(F;\mathbf{Z}/n\mathbf{Z}) \arrow[r] \arrow[d] & Q \arrow[r] \arrow[d] & 0, \\
0 \arrow[r] & A(F)/(nZ(F)+C(F)) \arrow[r] \arrow[d, equal] & A(F;\mathbf{Z}/n\mathbf{Z})/(nZ(F)+C(F)) \arrow[r] \arrow[d, equal] & Q/nR \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & B(F)/nB(F) \arrow[r] & B(F;\mathbf{Z}/n\mathbf{Z}) \arrow[r] & Q/nR \arrow[r] & 0,
\end{tikzcd}$$

We have inclusions $Q \subseteq R$ and $nR \subseteq Q \subseteq n\bigwedge^2 F^\times$. From now on, we make the assumption that the number field F does not contain a pth root of unity for any p dividing n. This implies from the previous inclusions that Q and R are all p-torsion free for $p|n$. Tensor the exact sequence (24) with $\mathbf{Z}/n\mathbf{Z}$. The group $\mathrm{Tor}^1(\mathbf{Z}/n\mathbf{Z}, \wedge^2 F^\times)$ vanishes by our assumption. Hence we have an exact sequence:

$$0 \to K_2(F)[n] \to R/nR \to \bigwedge^2 F^\times \otimes \mathbf{Z}/n\mathbf{Z} \to K_2(F)/nK_2(F) \to 0. \tag{25}$$

Recall that R is the image of $Z(F)$ in $\bigwedge^2 F^\times$ and Q is the image of $A(F;\mathbf{Z}/n\mathbf{Z})$, which is precisely the kernel of the map from R to $\bigwedge^2 F^\times \otimes \mathbf{Z}/n\mathbf{Z}$. It follows that the image of Q in R/nR is the kernel of the map from R/nR to $\bigwedge^2 F^\times \otimes \mathbf{Z}/n\mathbf{Z}$. From the short exact sequence (25), this may be identified with $K_2(F)[n]$. Since the image of Q in R/nR is precisely Q/nR, however, this shows that $Q/nR \simeq K_2(F)$, we obtain the exact sequence:

$$0 \longrightarrow B(F)/nB(F) \longrightarrow B(F;\mathbf{Z}/n\mathbf{Z}) \longrightarrow K_2(F)[n] \longrightarrow 0\,,$$

completing the proof of Theorem 1.7.

The previous result was a diagram chase. The map $\delta : B(F;\mathbf{Z}/n\mathbf{Z}) \to K_2(F)$ can be given explicitly as follows: Lift $[x] \in B(F;\mathbf{Z}/n\mathbf{Z})$ to an element x of $A(F;\mathbf{Z}/n\mathbf{Z})/C(F)$, which is unique up to an element of $nZ(F)$. The image of x in $\bigwedge^2 F^\times \otimes \mathbf{Z}/n\mathbf{Z}$ is zero by definition. Hence, because $\bigwedge^2 F^\times$ is p-torsion free for $p|n$, there exists an element $y \in \bigwedge^2 F^\times$ such that the image of z in $\bigwedge^2 F^\times$ is ny, and now y is unique up to an element in the image of $C(F)$. Yet the projection z of $y \in \bigwedge^2 F^\times$ to $K_2(F)$ sends this ambiguity $C(F)$ to zero, and so $\delta([x]) := z \in K_2(F)$ is well defined.

If we assume that n is not divisible by any prime p which divides $w_2(F)$, we have constructed a map

$$R_\zeta : B(F; \mathbf{Z}/n\mathbf{Z}) \to (F_n^\times / F_n^{\times n})^{\chi^{-1}} \simeq H^1(F, \mathbf{Z}/n\mathbf{Z}(2))\,. \tag{26}$$

Taking $n = p^m$ for various m, and using the fact that $B(F)$ is finitely generated and so $\operatorname{proj\,lim} B(F)/p^m B(F) = B(F) \otimes \mathbf{Z}_p$, we obtain a commutative diagram as follows:

$$\begin{tikzcd} 0 \arrow[r] & B(F)/p^m B(F) \arrow[r] \arrow[d] & B(F; \mathbf{Z}/n\mathbf{Z}) \arrow[r] \arrow[d] & K_2(F)[n] \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & H^1(F, \mathbf{Z}_p(2))/p^m \arrow[r] & H^1(F, \mathbf{Z}/n\mathbf{Z}(2)) \arrow[r] & H^2(F, \mathbf{Z}_p(2))[n] \arrow[r] & 0, \end{tikzcd} \tag{27}$$

The first vertical map is an isomorphism by Theorem 3.2, and the last vertical map is also an isomorphism by a theorem of Tate [32]. It follows that the map R_ζ in equation 26 is an isomorphism for n prime to $w_2(F)$. This gives a link between our explicit construction of Chern class maps for $K_3(F)$ and the explicit construction of $K_2(F)$ in Galois cohomology by Tate [32].

We end this section with a remark on circular units. Let $F = \mathbf{Q}(\zeta_D)$. Associated to a primitive Dth root of unity ζ_D, Beilinson (see §9 of [15]) constructed special generating elements of $K_3(F)$, which correspond, on the Bloch group side, to the classes $D\,[\zeta_D] \in B(F)$. Soulé [29] proved that the images of these classes under the Chern class map consist exactly of the circular units. On the other hand, for p not dividing D, we see that the images of $D[\zeta_D]$ under the maps R_ζ are unit multiples of the elements

$$\prod_{k=0}^{p^{m-1}} (1 - \zeta^k \, \zeta_D)^k \,;$$

these are exactly the compatible sequences of circular units which yield a finite index subgroup of $H^1(F, \mathbf{Z}_p(2))$ — the index being directly related to $K_2(\mathcal{O}_F)$ via the Quillen–Lichtenbaum conjectures.

7. Relation to quantum knot theory

As was mentioned in the introduction, the initial motivation for expecting a map as in (5) was the Quantum Modularity Conjecture, which concerns the asymptotics of a twisted version of the Kashaev invariant of a knot at roots of unity and a subtle transformation property (verified numerically for many knots and proved in the case of the figure 8 knot) of certain associated formal power series under the group $\mathrm{SL}(2, \mathbf{Z})$. In this section, we give a summary of this conjecture (a much more detailed discussion is given in [13]) and compare the near units appearing there with the ones studied in this paper.

Let K be a hyperbolic knot, i.e., an embedded circle in S^3 for which the 3-manifold $M_K = S^3 \smallsetminus K$ has a hyperbolic structure. This structure is then unique and gives several invariants: the *volume* $\mathrm{V}(K) \in \mathbf{R}_{>0}$ and *Chern-Simons-invariant* $\mathrm{CS}(K) \in \mathbf{R}/4\pi^2\mathbf{Z}$ of the 3-manifold M_K, the *trace field* $F_K = \mathbf{Q}\big[\{\mathrm{tr}(\gamma)\}_{\gamma\in\Gamma}\big]$ where $M_K = \mathbb{H}^3/\Gamma$ with $\Gamma \subset \mathrm{SL}(2,\mathbf{C})$ (the finitely generated group Γ is only unique up to conjugacy, but the set of traces of its elements is well-defined), and a *fundamental class* ξ_K in the Bloch group $B(F_K)$, defined as the class of $\sum[z_j]$ in $B(F_K)$, where $\sqcup\Delta_j = M_K$ is any ideal triangulation of M_K and z_j the cross-ratio of the four vertices of Δ_j. We also have two quantum invariants, the (normalized) *colored Jones polynomial* $J_N^K(q) \in \mathbf{Z}[q,q^{-1}]$ and the *Kashaev invariant* $\langle K\rangle_N \in \overline{\mathbf{Q}}$, which are computable expressions defined for any $N \in \mathbf{N}$ whose precise definitions, not needed here, we omit. The **Volume Conjecture**, due to Kashaev, says that the limit of $\frac{1}{N}\log|\langle K\rangle_N|$ as $N \to \infty$ equals $\frac{1}{2\pi}\mathrm{V}(K)$, the **Complexified Volume Conjecture** is the more precise statement $\langle K\rangle_N = e^{v(K)N+\mathrm{o}(N)}$ as $N\to\infty$, where $v(K) = \frac{1}{2\pi}(\mathrm{V}(K) - i\mathrm{CS}(K))$ (this makes sense because $v(K)$ is well-defined modulo $2\pi i$), and the yet stronger **Arithmeticity Conjecture**, stated in [6] and [9], says that there is a full asymptotic expansion

$$\langle K\rangle_N \;\sim\; \mu_8\delta(K)^{-1/2}\, N^{3/2}\, e^{v(K)N}\Big(1 \,+\, \kappa_1(K)\,\frac{2\pi i}{N} \,+\, \kappa_2(K)\Big(\frac{2\pi i}{N}\Big)^2 \,+\, \cdots\Big) \tag{28}$$

as $N \to \infty$, where $\delta(K)$ is a non-zero number related to the Ray-Singer torsion of K and where $\delta(K)$ and $\kappa_j(K)$ $(j \geq 1)$ belongs to the trace field F_K. An example, one of the few that are known rigorously (some other cases have now been proved by Ohtsuki et al; see [26]), is the expansion

$$\langle 4_1\rangle_N \;\sim\; \frac{N^{3/2}}{\sqrt[4]{3}}\, e^{v(4_1)N}\Big(1 \,+\, \frac{11\pi}{36\sqrt{3}\,N} \,+\, \frac{697\pi^2}{7776N^2} \,+\, \frac{724351\pi^3}{4199040\sqrt{3}\,N^3} \,+\, \cdots\Big)$$

for the knot $K = 4_1$ (figure 8), for which $F_K = \mathbf{Q}(\sqrt{-3})$.

Equation (28) is already a strong refinement of the Volume Conjecture. An even stronger is the **Modularity Conjecture** given in [39] and discussed further in [8] and [13]. The starting point is the famous theorem of H. Murakami and J. Murakami [23] saying that the Kashaev invariant $\langle K\rangle_N$, originally defined by Kashaev as a certain state sum, coincides with the value of the colored Jones polynomial $J_N^K(q)$ at $q = \zeta_N = e^{2\pi i/N}$. We define a function $\mathbf{J}^K : \mathbf{Q} \to \overline{\mathbf{Q}}$ by setting $\mathbf{J}^K(x) = J_N^K(e^{2\pi i x})$ for any $N \in \mathbf{N}$ with $Nx \in \mathbf{Z}$. This is independent of the choice of N since the values of the colored Jones polynomial $J_N^K(q)$ at an nth root of unity q is periodic in N of period n. From the Murakami-Murakami theorem and since the function $\mathbf{J}^K$ is Galois-invariant by its very definition, we have

$$\mathbf{J}^K\big(\frac{a}{N}\big) \;=\; \sigma_a\big(\langle K\rangle_N\big) \qquad \text{for any } a \in (\mathbf{Z}/N\mathbf{Z})^\times,$$

where σ_a is the automorphism of $\mathbf{Q}(\zeta_N)$ sending ζ_N to ζ_N^a, so the function $\mathbf{J}^K$ can be thought of as a Galois-twisted version of the Kashaev invariant. The Modularity Conjecture describes its asymptotic behavior of $\mathbf{J}^K(x)$ as the argument $x \in \mathbf{Q}$ approaches any fixed rational number α, not just 0. Specifically, it asserts that there exist formal power series $\Phi_\alpha^K(h) \in \overline{\mathbf{Q}}[[h]]$ ($\alpha \in \mathbf{Q}$, periodic in α with period 1), such that for any $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbf{Z})$, we have

$$\mathbf{J}^{K^*}\Big(\frac{aX+b}{cX+d}\Big) \;\sim\; (cX+d)^{3/2}\,\mathbf{J}^{K^*}(X)\, e^{v(K)(X+d/c)}\,\Phi^K_{a/c}\Big(\frac{2\pi i}{cX+d}\Big) \tag{29}$$

to all orders in $1/X$ as $X \to \infty$ in $\mathbf{Q}$ with bounded denominator. Here, K^* is the mirror of K. If we take $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $X = N \to \infty$, then (29) reduces to $\langle K \rangle_N = \mathbf{J}^K(1/N) = \mathbf{J}^{K^*}(-1/N) \sim N^{3/2} e^{v(K)N} \Phi_0^K\left(\frac{2\pi i}{N}\right)$, so (29) is a generalization of (28) with $\Phi_0^K(h) = \mu_8\, \delta(K)^{-1/2}\, (1 + \kappa_1(K)h + \cdots)$.

The Quantum Modularity Conjecture, as already mentioned, was first given in [39], with detailed numerical evidence for the case of the figure 8 knot, and is further discussed in [8] and then in much more detail in [13], where this case is proved completely and numerical examples for several more knots are given. It is perhaps worth mentioning that regarding numerical evidence, one cannot use the standard `SnapPy` or `Mathematica` programs to compute the colored Jones polynomial (hence the Kashaev invariant) of a knot, since these work for small values of N (say, $N \sim 20$). Instead, we used a finite recursion for the colored Jones polynomial, whose existence was proven in [9], and concretely computed in several examples summarized in [8] and [13]. We could then compute the Kashaev invariant numerically to high precision up to N of the order of 5000, which with suitable numerical extrapolation techniques made it possible to compute and recognize several terms of the series $\Phi_{a/c}^K(h)$ with high confidence.

But already in the constant term of the series $\Phi_{a/c}^K(h)$, mysterious roots of algebraic units appear and those led to the main theorems of this paper. For instance, for the 4_1 knot, when a is an integer prime to 5, we have

$$\Phi_{a/5}^{4_1}(h) \;=\; 3^{\frac{1}{4}}\left(\varepsilon^{(a)}\right)^{\frac{1}{10}} \left(\left(2-\varepsilon_1^{(a)}+\varepsilon_2^{(a)}+2\varepsilon_3^{(a)}\right) + \frac{2678 - 943\varepsilon_1^{(a)} + 1831\varepsilon_2^{(a)} + 2990\varepsilon_3^{(a)}}{2^3 3^2 5^2 \sqrt{-3}}\, h + \cdots\right),$$

where $\varepsilon^{(a)} = \varepsilon_2^{(a)}/(\varepsilon_1^{(a)})^3 \varepsilon_3^{(a)}$ and $\varepsilon_k^{(a)} = 2\cos\frac{2\pi(6a-5)k}{15}$. (See [39], p. 670 except that the formula is given there in terms of $\log\Phi$, which makes its coefficients much more complicated.) The number $\varepsilon_k^{(a)}$ is an algebraic unit in $F_5 = \mathbf{Q}(\zeta_{15})$, and it is the appearance of the 10th root of $\varepsilon^{(a)}$ that was the origin of the present investigation. In fact, although the unit $\varepsilon^{(a)}$ is not a square in the field $\mathbf{Q}(\zeta_{15})^+$ which it generates, its negative is a square in the larger field F_5: $\sqrt{-\varepsilon^{(a)}} = 2i\sin\frac{2\pi(6a-5)}{15}\varepsilon_2^{(a)}/\varepsilon_1^{(a)}$.

More generally, when we do numerical calculations for arbitrary knots K, we find:

- the power series $\Phi_\alpha^K(h)$ ($\alpha \in \mathbf{Q}$) belongs to $\overline{\mathbf{Q}}[[h]]$ and has a factorization of the form
$$\Phi_\alpha^K(h) \;=\; C_\alpha(K)\, \phi_\alpha^K(h) \tag{30}$$
with $C_\alpha(K) \in \overline{\mathbf{Q}}$ and $\phi_\alpha^K(h) \in F_c[[h]]$, where the number field F is independent of α (and is in fact is conjecturally the trace field F_K of K) and $F_c = F(\zeta_c)$;
- the constant $C_\alpha(K)$ factors as
$$C_\alpha(K) \;=\; \mu_{8c}(K)\delta(K)^{-1/2}\, \varepsilon_\alpha(K)^{1/c}$$
where μ_{8c} is a $8c$ root of unity and $\varepsilon_\alpha(K)$ is a unit of F_c;
- the units $\varepsilon_\alpha(K)$ for different rational numbers α and $\beta = k\alpha$ with the same denominator (assumed prime to some fixed integer depending on K) are related by both $\varepsilon_\beta(K) = \sigma_k(\varepsilon_\alpha(K))$ and $\varepsilon_\beta(K)^k = \varepsilon_\alpha(K)$ (the latter equality holding modulo cth powers), where $\sigma_k \in \mathrm{Gal}(F_c/F)$ is the map sending ζ_α to ζ_β. Compare this double Galois invariance with (6).

Notice that the factorization (30) is not canonical, since we can change both the constant C_α and the power series by a unit of F_c and its inverse, so that the formula involves a slight abuse of notation. Notice also that v depends only on the element of the Bloch group associated to the knot [25, 14] and so does ε_α, but not Φ_α. Specifically, as seen in [13], sister (or partner) knots do *not* have the same power series, but do (experimentally) have the same unit. It is this observation that led us to search for the map (5).

8. Nahm's conjecture and the asymptotics of Nahm sums at roots of unity

In the previous section, we saw that the near units constructed in this paper from elements of the Bloch group appear naturally (although in general only conjecturally) in connection with the asymptotic properties of the Kashaev invariant of knots and its Galois twists. A second place where these units appear is in the radial asymptotics of so-called Nahm sums, as was shown in [11] and is quoted (in a simplified form) in Theorem 8.1 below. In this section, we explain this and give two applications, the proof of Theorem 8.5 and the proof of one direction of Nahm's conjecture relating the modularity of Nahm sums to the vanishing of certain elements in Bloch groups.

Nahm sums are special q-hypergeometric series whose summand involves a quadratic form, a linear form and a constant. They were introduced by Nahm [24] in connection with characters of rational conformal field theories, and led to his above-mentioned conjecture concerning their modularity. They have also appeared recently in quantum topology in relation to the stabilization of the coefficients of the colored Jones polynomial (see Garoufalidis-Le [10]), and they are building blocks of the 3D-index of an ideally triangulated manifold due to Dimofte-Gaiotto-Gukov [5, 4]. Further connections between quantum topological invariants and Nahm sums are given in [12], where one sees once again the appearance of the units $R_\zeta(\xi)^{1/n}$.

In the first subsection of this section, we review Nahm sums and the Nahm conjecture and state Theorem 8.1 relating the asymptotics of Nahm sums at roots of unity to the near units of Theorem 1.2. This is then applied in §8.2 to a particular Nahm sum (namely, the famous Andrews-Gordon generalization of the Rogers-Ramanujan identities) to prove equation (12) of the introduction (Theorem 8.5). In the final subsection, we use Theorem 8.1 together with Theorem 1.2 to give a proof of one direction of Nahm's conjecture.

8.1. **Nahm's conjecture and Nahm sums.**

Nahm's conjecture gives a very surprising connection between modularity and algebraic K-theory. More precisely, it predicts that the modularity of certain q-hypergeometric series ("Nahm sums") is controlled by the vanishing of certain associated elements in the Bloch group $B(\overline{\mathbf{Q}}) = K_3(\overline{\mathbf{Q}})$.

The definition of Nahm sums and the question of determining when they are modular were motivated by the famous Rogers-Ramanujan identities, which say that

$$G(q) := \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q)_n} = \prod_{\substack{n>0 \\ (\frac{n}{5})=1}} \frac{1}{1-q^n}, \qquad H(q) := \sum_{n=0}^{\infty} \frac{q^{n^2+n}}{(q)_n} = \prod_{\substack{n>0 \\ (\frac{n}{5})=-1}} \frac{1}{1-q^n},$$

where $(q)_n = (1-q)\cdots(1-q^n)$ is the q-Pochhammer symbol or quantum n-factorial. These identities imply via the Jacobi triple product formula that the two functions $q^{-1/60}G(q)$ and

$q^{11/60}H(q)$ are quotients of unary theta-series by the Dedekind eta-function and hence are modular functions. (Here and from now on we will allow ourselves the abuse of terminology of saying that a function $f(q)$ is modular if the function $\widetilde{f}(\tau) = f(e^{2\pi i\tau})$ is invariant under the action of some subgroup of finite index of $\mathrm{SL}(2,\mathbf{Z})$.) To see how general this phenomenon might be, Nahm [24] considered the three-parameter family

$$f_{A,B,C}(q) = \sum_{m\ge 0} \frac{q^{\frac{A}{2}m^2+Bm+C}}{(q)_m} \qquad (A\in\mathbf{Q}_{>0},\ B,\ C\in\mathbf{Q}) \tag{31}$$

These are formal power series with integer coefficients in some rational power of q, and are analytic in the unit disk $|q|<1$, but they are very seldom modular: apart from the two Rogers-Ramanujan cases $(A,B,C) = (2,0,-\frac{1}{60})$ or $(2,1,\frac{11}{60})$, only five further cases $(1,0,-\frac{1}{48})$, $(1,\pm\frac{1}{2},\frac{1}{24})$, $(\frac{1}{2},0,-\frac{1}{40})$ and $(\frac{1}{2},\frac{1}{2},\frac{1}{40})$ were known for which $f_{A,B,C}$ is modular, and it was later proved ([33], [38]) that these are in fact the only ones. Since this list of seven examples is not very enlightening, Nahm introduced also a higher-order version

$$f_{A,B,C}(q) = \sum_{m\in\mathbf{Z}^r_{\ge 0}} \frac{q^{\frac{1}{2}m^tAm+Bm+C}}{(q)_{m_1}\cdots(q)_{m_r}}\,, \tag{32}$$

where $A=(a_{ij})$ is a symmetric positive definite $r\times r$ matrix with rational entries, $B\in\mathbf{Q}^r$ a column vector, and $C\in\mathbf{Q}$ a scalar, and asked for which triples (A,B,C) the function $\widetilde{f}_{A,B,C}(\tau) = f_{A,B,C}(e^{2\pi i\tau})$ is modular. His conjecture gives a partial answer to this question.

To formulate this conjecture, Nahm made two preliminary observations.

(i) Let $X=(X_1,\dots,X_r)\in\mathbf{C}^r$ be a solution of *Nahm's equations*

$$1\ -\ X_i\ =\ \prod_{j=1}^{r} X_j^{a_{ij}} \qquad (1\le j\le r) \tag{33}$$

(or symbolically $1-X=X^A$), and let F be the field they generate over $\mathbf{Q}$, which will typically be a number field since (33) is a system of r equations in r unknowns and generically defines a 0-dimensional variety. Then the element $[X]=[X_1]+\cdots[X_r]$ of $\mathbf{Z}[F]$ belongs to the kernel of the map (2), because

$$d\big([X]\big)\ =\ \sum_i (X_i)\wedge(1-X_i)\ =\ \sum_{i,j} a_{ij}\,(X_i)\wedge(X_j)\ =\ 0$$

by virtue of the symmetry of A. (This calculation makes sense as it stands if A has integer entries; if the entries are only rational, we have to tensor everything with $\mathbf{Q}$.) Therefore $[X]$ determines an element of the Bloch group $B(F)\otimes\mathbf{Q}$ and it makes sense to ask whether this element vanishes. This is equivalent to the vanishing of the numbers $D(\sigma X)=\sum D(\sigma X_i)$ for all embeddings $\sigma: F\hookrightarrow\mathbf{C}$, where $D(x)$ is the Bloch-Wigner dilogarithm function, and this condition can be either tested numerically to any precision or else verified rigorously by writing a multiple of $[X]$ as a linear combination of 5-term relations.

(ii) The first remark applies to any symmetric matrix A. If A is positive definite, then there is a distinguished solution of the Nahm equations, namely the unique solution $X^A = (X_1^A,\dots,X_r^A)$ with $0<X_i^A<1$ for all i. We denote by ξ_A the corresponding element $[X^A]$

of the Bloch group. Then since X^A is real, we obtain a further characteristic property when this element is torsion, namely that the real number $\mathrm{L}(\xi_A) = \sum \mathrm{L}(X_i)$, where $\mathrm{L}(x)$ is the Rogers dilogarithm function as defined below, is a rational multiple of π^2. But it can be shown fairly easily that $f_{A,B,C}(e^{-h})$ has an asymptotic expansion as $e^{\mathrm{L}(\xi_A)/h+\mathrm{O}(1)}$ as $h \to 0^+$ for any B and C (in fact, a full asymptotic expansion of the form $e^{\mathrm{L}(\xi_A)/h+c_0+c_1h+\cdots}$ is given in [38]). Since a modular function must have an expansion $e^{c/h+\mathrm{O}(1)}$ with $c \in \mathbf{Q}\pi^2$, this already gives a strong indication of a relation between the modularity of Nahm sums and the vanishing (up to torsion) of the associated elements of Bloch groups.

Based on these observations, one can consider the following three properties of a matrix A as above:

(a) The class $[X] \in B(\mathbf{C})$ vanishes for all solutions X of the Nahm equations (33).

(b) The special class $\xi_A \in B(\mathbf{C})$ associated to the solution X^A of (33) vanishes.

(c) The function $f_{A,B,C}(q)$ is modular for some $B \in \mathbf{Q}^r$ and $C \in \mathbf{Q}$.

Trivially (a) $\Rightarrow$ (b). Nahm's conjecture (see [24] and [38]) says that (a) $\Rightarrow$ (c) and (c) $\Rightarrow$ (b). (The possible stronger hypothesis that (b) alone might already imply (c) was eliminated in [38] using the 2×2 matrix $A = \left(\begin{smallmatrix} 8 & 5 \\ 5 & 4 \end{smallmatrix}\right)$, and the other possible stronger assertion that (c) might require (a) was shown to be false by Vlasenko and Zwegers [34] with the counterexample $A = \left(\begin{smallmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{smallmatrix}\right)$.) This conjecture had a dual motivation: on the one hand, the above-mentioned fact that both (b) and (c) force the rationality of $L(\xi_A)/\pi^2$, which is most unlikely to happen "at random," and on the other hand, a large number of supporting examples coming from the characters of rational conformal field theories, which are always modular functions and where the condition in the Bloch group can also be verified in many cases. Here we are concerned with an extension of the first of these two aspects, namely the asymptotics of the Nahm sum $f_{A,B,C}(q)$ as q tends radially to *any* root of unity, not just to 1.

In order to state the asymptotic formula, we need to define the Rogers dilogarithm. In our normalization (which is $\pi^2/6$ minus the standard one as given, for instance, in [38], §II.1A), this is the function defined on $\mathbf{R} \smallsetminus \{0,1\}$ by

$$\mathrm{L}(x) \;=\; \begin{cases} \frac{\pi^2}{6} \,-\, \mathrm{Li}_2(x) \,-\, \frac{1}{2}\log(x)\log(1-x) & \text{if } 0<x<1, \\ -\,\mathrm{L}(1/x) & \text{if } x>1, \\ \frac{\pi^2}{6} \,-\, \mathrm{L}(1-x) & \text{if } x<0 \end{cases}$$

(here $\mathrm{Li}_2(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}$ is the standard dilogarithm) and extended by continuuity to a function $\mathbf{P}^1(\mathbf{R}) \to \mathbf{R}/\frac{\pi^2}{2}\mathbf{Z}$ by sending the three points 0, 1 and ∞ to $\frac{\pi^2}{6}$, 0, and $-\frac{\pi^2}{6}$. Its linear extension to $Z(\mathbf{R})$ vanishes on the group $C(\mathbf{R})$ as defined at the beginning of §1.1. (We comment here that there are several definitions of the Bloch group in the literature, all the same up to 6-torsion, and that the specific choice made in Definition 1.1, which forces $3[0]=0$, $[X]+[1/X]=0$ and $[X]+[1-X]=[0]$ for any field F and any element X of $\mathbf{P}^1(F)$, was chosen precisely so that L is well-defined on $B(\mathbf{R})$ and takes values in the full circle group $\mathbf{R}/\frac{\pi^2}{2}\mathbf{Z}$ rather than just its quotient $\mathbf{R}/\frac{\pi^2}{6}\mathbf{Z}$.)

Specifically, let A, B and C be as above let $X = X^A$ be the distinguished solution of (33) as in (ii) and F the corresponding number field, and for each integer n choose a primitive

nth root of unity ζ, set $F_n = F(\zeta)$ and denote by $H = H_n$ the Kummer extension of F_n obtained by adjoining the positive nth roots x_i of the X_i. We are interested in the asymptotic expansion of $f_{A,B,C}(\zeta e^{-h/n})$ as $h \to 0^+$. Strictly speaking, this only makes sense if A has integral coefficients, B is congruent to $\frac{1}{2}\mathrm{diag}(A)$ modulo $\mathbf{Z}^r$, and $C \in \mathbf{Z}$, since otherwise the quadratic function $q^{\frac{1}{2}nAn^t+nB+C}$ occurring in the definition of $f_{A,B,C}$ is not uniquely defined. We get around this by picking a representation of ζ as $\mathbf{e}(a/n)$ for some $a \in \mathbf{Z}$ and interpreting $f_{A,B,C}(\zeta e^{-h/n})$ as $\widetilde{f}_{A,B,C}\big(\frac{a+i\hbar}{n}\big)$, where $\hbar = \frac{h}{2\pi}$. The full asymptotic expansion of $f_{A,B,C}(\zeta e^{-h/n})$ as $h \to 0^+$ was calculated in [12] using the Euler-Maclaurin formula, generalizing an earlier result in [38] for the case $n = 1$. We do not give the complete formula here, but only the simplified form as needed for the applications we will give. In the statement of the theorem we have abbreviated by Δ_X the diagonal matrix whose diagonal is a vector X.

Theorem 8.1. [12] *Let (A, B, C) be as above. Then for every positive integer n (coprime to a finite set of primes that depend on A and B) and for every primitive nth root of unity ζ, we have*

$$f_{A,B,C}\big(\zeta\, e^{-h/n}\big) \;=\; \mu\,\omega\, e^{\mathrm{L}(\xi_A)/nh}\,\big(\Phi_\zeta(h) \,+\, O(h^K)\big) \tag{34}$$

for all $K > 0$ as $h \to 0^+$, where $\omega^2 \in F^\times$, $\mu^{24n} = 1$ and $\Phi_\zeta(h) = \Phi_{A,B,C,\zeta}(h)$ is an explicit power series satisfying the two properties $\Phi_\zeta(h)^n \in F_n[[h]]$ and $P_\zeta(\xi_A)^{1/n}\,\Phi_\zeta(h) \in H_n[[h]]$. Moreover, if $\Phi_\zeta(0)^n \neq 0$, then its image in $F_n^\times/F_n^{\times n}$ belongs to the χ^{-1} eigenspace.

Remark 8.2. If n is prime to 6, then we can choose μ to be a 24th root of unity, since the nth roots of unity are contained in F_n and can be absorbed into the power series Φ.

Corollary 8.3. If $\Phi_\zeta(0) \neq 0$, then the product of the power series $\Phi_\zeta(h)$ with $\varepsilon^{1/n}$ for any unit ε representing $R_\zeta(\xi_A)$ belongs to $F_n[[h]]$.

Proof. Let $\varepsilon \in F_n^\times$ denote a representative of $R_\zeta(\xi_A)$. On the one hand, Theorem 8.1 and Remark 2.5 imply that $\Phi_\zeta(0)\varepsilon^{1/n} \in F_n^\times$. On the other hand, Theorem 8.1 and our assumption implies that $(\Phi_\zeta(h)/\Phi_\zeta(0))^n \in F_n[[h]]$. Since $\Phi_\zeta(h)/\Phi_\zeta(0)$ is a power series with constant term 1, it follows that $\Phi_\zeta(h)/\Phi_\zeta(0) \in F_n[[h]]$. Combining both conclusions, it follows that $\varepsilon^{1/n}\Phi_\zeta(h) \in F_n[[h]]$. ☐

Remark 8.4. In the theorem, we do *not* assert that the power series Φ cannot vanish identically (which is why we wrote an equality sign and $\Phi(h) + \mathrm{O}(h^K)$ in (34) rather than writing an asymptotic equality sign and putting simply $\Phi(h)$ on the right), and indeed this often happens, for instance, when $f_{A,B,C}$ is modular and we are expanding at a cusp not equivalent to 0. Of course, the corollary is vacuous if Φ vanishes.

8.2. **Application to the calculation of $R_\zeta(\eta_\zeta)$.** In this subsection, we apply Theorem 8.1 and its corollary to a specific Nahm sum to prove equation (12) in the introduction.

Theorem 8.5. *Let n be positive and prime to 6 and η_ζ be the n-torsion element in $B(\mathbf{Q}(\zeta)^+)$ defined by (23), where ζ is a primitive nth root of unity. Then $R_\zeta(\eta_\zeta)^4 \;=\; \zeta$.*

Proof. Set $A_n = \big(2\min(i,j)\big)_{1\le i,j\le r}$, where $r = \frac{n-3}{2}$, and let f_n be the Nahm sum $f_{A_n,0,0}$ of order r. By a famous identity of Andrews and Gordon [1], which reduces to the first

Ramanujan-Rogers identity when $n = 5$, we have the product expansion

$$f_n(q) \;=\; \prod_{\substack{k>0 \\ 2k\not\equiv 0,\pm 1 \,(\bmod n)}} \frac{1}{1-q^k}. \tag{35}$$

and this is modular up to a power of q for the same reason as for $G(q) = f_5(q)$ (quotient of a theta series by the Dedekind eta-function). This modularity allows us to compute its asymptotics as $q \to \zeta_n$, and by comparing the result with the general asymptotics of Nahm sums as given in 8.1, we will obtain the desired evaluation of η_n. We now give details.

It is easy to check that all solutions X of the Nahm equation $1 - X = X^{A_n}$ have the form

$$X = (X_1, \ldots, X_r), \qquad X_k \;=\; \frac{(1-\zeta^{2k})(1-\zeta^{2k+4})}{(1-\zeta^{2k+2})^2}$$

with ζ a primitive n root of unity, and hence form a single Galois orbit. The distinguished solution $X^{A_n} \in (0,1)^r$ corresponds to $\zeta = \mathbf{e}(1/n) = \zeta_n$. From $1 - X_k = (\frac{\zeta-\zeta^{-1}}{\zeta^{k+1}-\zeta^{-k-1}})^2$ and the functional equation $L(1-X) = \frac{\pi^2}{6} - L(X)$ we find

$$\mathrm{L}(X^{A_n}) \;=\; \frac{1}{2}\sum_{0<\ell<n}\left(\frac{\pi^2}{6} - \mathrm{L}\left(\frac{\sin^2(\pi/n)}{\sin^2(\ell\pi/n)}\right)\right) \;=\; \frac{(n-3)\pi^2}{6n}\,,$$

the final equality being a well-known identity for the Rogers dilogarithm of which a proof can be found at the end of [38], §II.2C. Denote the right-hand side of this by $-4\pi^2 C_n$ and set $\widetilde{f}_n(\tau) = q^{C_n} f_n(q)$. Using the Jacobi theta function and Jacobi triple product formula

$$\theta(\tau,z) \;=\; \sum_{n\in\mathbf{Z}+1/2} (-1)^{[n]}\, q^{n^2/2}\, y^n \;=\; q^{1/8}\, y^{1/2} \prod_{n=1}^{\infty} \bigl(1-q^n\bigr)\bigl(1-q^n y\bigr)\bigl(1-q^{n-1}y^{-1}\bigr)$$

(where $\Im(\tau) > 0$, $z \in \mathbf{C}$, $q = \mathbf{e}(\tau)$, and $y = \mathbf{e}(z)$), together with the Dedekind eta-function $\eta(\tau) = q^{1/24}\prod_{n>0}(1-q^n)$, we can rewrite (35) as

$$f_n(\tau) \;=\; -q^{(r+1)^2/2n}\,\frac{\theta(n\tau,\,(r+1)\tau)}{\eta(\tau)}\,,$$

which in conjunction with the standard transformation properties of θ and η implies that $f_n(\tau)$ is a modular function (with multiplier system) on the congruence subgroup $\Gamma_0(n)$ of $\mathrm{SL}(2,\mathbf{Z})$. We need only the special case $\tau \mapsto \frac{\tau}{n\tau+1}$, where the transformation law is given by

$$f_n\Bigl(\frac{\tau}{n\tau+1}\Bigr) \;=\; \mathbf{e}\Bigl(\frac{n-3}{24}\Bigr)\, f_n(\tau)\,, \tag{36}$$

whose proof we sketch for completeness. The well-known modular transformation properties of θ and η under the generators $T = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ of $\mathrm{SL}(2,\mathbf{Z})$ are given by

$$\theta(\tau+1,z) \;=\; \mathbf{e}(1/8)\,\theta(\tau,z)\,, \quad \theta(-1/\tau,\,z/\tau) \;=\; \sqrt{\tau/i}\;\mathbf{e}(z^2/2\tau)\,\theta(\tau,z)$$
$$\eta(\tau+1) \;=\; \mathbf{e}(1/24)\,\eta(\tau)\,, \qquad \eta(-1/\tau) \;=\; \sqrt{\tau/i}\;\eta(\tau)\,.$$

Hence, using $\overset{T}{\sim}$ and $\overset{S}{\sim}$ to denote an equality up to an elementary factor (the product of a power of τ with the exponential of a linear combination of 1, τ and z^2/τ) that can be deduced from the T- or S-transformation behavior of the function in question, we have

$$\theta\Big(\frac{n\tau}{n\tau+1},\frac{(r+1)\tau}{n\tau+1}\Big) \overset{T}{\sim} \theta\Big(\frac{-1}{n\tau+1},\frac{(r+1)\tau}{n\tau+1}\Big) \overset{S}{\sim} \theta\left(n\tau+1,(r+1)\tau\right) \overset{T}{\sim} \theta\left(n\tau,(r+1)\tau\right) ,$$

$$\eta\Big(\frac{\tau}{n\tau+1}\Big) \overset{S}{\sim} \eta\Big(-n-\frac{1}{\tau}\Big) \overset{T}{\sim} \eta\Big(-\frac{1}{\tau}\Big) \overset{S}{\sim} \eta(\tau)\,.$$

Inserting all omitted factors and dividing the first equations by the second, we obtain (36).

Now applying (36) to $\tau=\frac{1+i\hbar}{n}$, with $\hbar=\frac{h}{2\pi}$, where h positive and small, we find

$$\begin{aligned} f_{A_n,0,C_n}\big(\zeta_n e^{-h/n}\big) &= f_n\Big(\frac{1+i\hbar}{n}\Big) = \mathbf{e}\Big(\frac{n-3}{24}\Big)\, f_n\Big(\frac{-1+i/\hbar}{n}\Big) \\ &= \mathbf{e}\Big(\frac{n}{24}-\frac{1}{8}+\frac{1}{24n}-\frac{1}{8n^2}\Big)\, e^{\mathrm{L}(X^{A_n})/nh}\left(1+O\big(e^{-4\pi^2/nh}\big)\right). \end{aligned} \tag{37}$$

Taking the $8n$-th power of this and combining with Theorem 8.1 and its Corollary 8.3, we find that $R_\zeta(\xi_{A_n})^8=\mathbf{e}(1/n)\in(F_n^\times/F_n^{\times n})^{\chi^{-1}}$. On the other hand, using the same identity $1-X_k=(\frac{\zeta-\zeta^{-1}}{\zeta^{k+1}-\zeta^{-k-1}})^2$ as before, we find that the Bloch element ξ_{A_n} associated to the distinguished real solutions X^{A_n} of the Nahm equation is *equal* to twice the Bloch element η_ζ defined in (23). This completes the proof of Theorem 8.5. ◻

8.3. **Application to Nahm's conjecture.** In this final subsection, we give an application of the asymptotic Theorem 8.1 and Theorem 1.2 to proving one direction of Nahm's conjecture about the modularity of Nahm sums. The notations and assumptions are as before, but for convenience we repeat them here.

Let $A\in M_r(\mathbf{Q})$ be a positive definite symmetric matrix, $B\in\mathbf{Q}^r$, and $C\in\mathbf{Q}$. We denote $X^A=(X_1,\dots,X_r)$ denote the unique solution in $(0,1)^r$ to the Nahm equation, by $F=F_A$ the real number field generated by the X_i and by $\xi_A=\sum_i[X_i]\in B(F_A)$ the corresponding element of the Bloch group. Finally, when we say that $F_{A,B,C}$ is modular, we mean that the function $\widetilde{f}(\tau)=f_{A,B,C}(\mathbf{e}(\tau))$ is invariant with respect to a subgroup of finite index of $\mathrm{SL}(2,\mathbf{Z})$.

Theorem 8.6. *If $f_{A,B,C}(\tau)$ is a modular function, then $\xi_A\in B(F_A)$ is a torsion element.*

Proof. On p. 56 of [38] it is shown that any Nahm sum has an expansion near $q=1$ of the form

$$f_{A,B,C}(e^{-\epsilon}) = e^{\mathrm{L}(\xi_A)/\epsilon}\left(K+\mathrm{O}(\epsilon)\right) \qquad (\epsilon\to 0), \tag{38}$$

where K (given explicitly in eq. (29) of [38]) is a non-zero algebraic number some power of which belongs to $F=F_A$ and where the error term $\mathrm{O}(\epsilon)$ can be replaced by $\mathrm{O}(e^{-c/\epsilon})$ with some $c>0$ if $f_{A,B,C}$ is assumed to be modular ([38], eq. (28)). Notice that in this case the number $\lambda=\mathrm{L}(\xi_A)/4\pi^2$ must be rational, since the modularity of $\widetilde{f}(\tau)=f_{A,B,C}(\mathbf{e}(\tau))$ implies that the function $\widetilde{f}(-1/\tau)$ is invariant under some power of $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$.

Now assume that $\widetilde{f}$ is modular with respect to a finite index subgroup Γ of $\mathrm{SL}(2,\mathbf{Z})$. Then for $h \to 0^+$, $\hbar = \frac{h}{2\pi}$, and any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, taking $\epsilon = \frac{dh}{1-ic\hbar}$, we find

$$f_{A,B,C}(e^{-\epsilon}) \;=\; \widetilde{f}\Big(\frac{i\epsilon}{2\pi}\Big) \;=\; \widetilde{f}\Big(\frac{ai\epsilon/2\pi+b}{ci\epsilon/2\pi+d}\Big) \;=\; \widetilde{f}\Big(\frac{b+i\hbar}{d}\Big) \;=\; f_{A,B,C}(\zeta e^{-h/d}),$$

where $\zeta = \mathbf{e}(b/d)$, and now comparing the asymptotic formulas (38) and (34) (with $n = d$), we find

$$\mu\, e^{\mathrm{L}(\xi_A)/hd}\, \Phi(h) \;=\; e^{L(\xi_A)/dh}\, \big(K\mathbf{e}(\lambda c/d) \,+\, \mathrm{O}(h)\big)$$

or $\Phi_\zeta(0) = \mu^{-1}K\mathbf{e}(\lambda c/d)$, with $\lambda \in \mathbf{Q}$ as above. This implies in particular that $\Phi_\zeta(0) \neq 0$, and now, using that some bounded power of both μ and K belong to F_n, we deduce that $\Phi(0)^r$ belongs to F_n for some fixed integer $r > 0$ independent of $n = d$. We can also assume that d is prime to M for any fixed integer M, since by intersecting Γ with the full congruence subgroup $\Gamma(M)$, we may assume that Γ is contained in $\Gamma(M)$. This shows that there are infinitely many integers n and primitive nth roots of unity ζ for which $\Phi_\zeta(0)^r$ in Theorem 8.1 is a non-zero element of F_n. Now Corollary 8.3 implies that the rth power of $R_\zeta(\xi_A)$ has trivial image in $F_n^\times/F_n^{\times n}$ for infinitely many n, and in view of the injectivity statement in Theorem 1.2 this proves that ξ_A is a torsion element in the finitely generated group $B(F)$. ☐

Remark 8.7. The proof of the theorem would have been marginally shorter if we had assumed that $f_{A,B,C}$ was a modular function on a congruence subgroup, rather than just a subgroup of finite index of $\mathrm{SL}(2,\mathbf{Z})$. We did not make this assumption since it was not needed, but should mention that $f_{A,B,C}$, if modular at all, is expected automatically to be modular for a congruence subgroup, because it has a Fourier expansion with integral coefficients in some rational power of q and a standard conjecture says that the Fourier expansion of a modular function on a non-congruence subgroup of $\mathrm{SL}(2,\mathbf{Z})$ always has unbounded denominators.

Remark 8.8. Conversely, we could have stated Theorem 8.6 in an apparently more general form by writing "modular form" instead of "modular function." We did not do this since it is easy to see that if a Nahm sum is modular at all, it is actually a modular function, because if it were a modular form of non-zero rational weight k, there would be an extra factor h^{-k} in the right-hand side of (38).

Acknowledgements. F.C. and S.G. were supported in part by grants DMS-1404620 and DMS-1406419 of the US National Science Foundation.

References

[1] George Andrews. *On the general Rogers-Ramanujan theorem.* American Mathematical Society, Providence, R.I., 1974. Memiors of the American Mathematical Society, No. 152.

[2] Spencer Bloch. *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, volume 11 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2000.

[3] Frank Calegari, Stavros Garoufalidis, and Don Zagier. In preparation.

[4] Tudor Dimofte, Davide Gaiotto, and Sergei Gukov. 3-Manifolds and 3d indices. *Adv. Theor. Math. Phys.*, 17(5):975–1076, 2013.

[5] Tudor Dimofte, Davide Gaiotto, and Sergei Gukov. Gauge theories labelled by three-manifolds. *Comm. Math. Phys.*, 325(2):367–419, 2014.

[6] Tudor Dimofte, Sergei Gukov, Jonatan Lenells, and Don Zagier. Exact results for perturbative Chern-Simons theory with complex gauge group. *Commun. Number Theory Phys.*, 3(2):363–443, 2009.
[7] Johan Dupont and Chih Han Sah. Scissors congruences. II. *J. Pure Appl. Algebra*, 25(2):159–195, 1982.
[8] Stavros Garoufalidis. Quantum Knot Invariants. Mathematische Arbeitstagung 2012.
[9] Stavros Garoufalidis and Thang T. Q. Lê. The colored Jones function is q-holonomic. *Geom. Topol.*, 9:1253–1293 (electronic), 2005.
[10] Stavros Garoufalidis and Thang T. Q. Lê. Nahm sums, stability and the colored Jones polynomial. *Res. Math. Sci.*, 2:Art. 1, 55, 2015.
[11] Stavros Garoufalidis and Don Zagier. Asymptotics of Nahm sums. In preparation.
[12] Stavros Garoufalidis and Don Zagier. Knots and their related q-series. In preparation.
[13] Stavros Garoufalidis and Don Zagier. Quantum modularity of the Kashaev invariant. In preparation.
[14] Sebastian Goette and Christian Zickert. The extended Bloch group and the Cheeger-Chern-Simons class. *Geom. Topol.*, 11:1623–1635, 2007.
[15] Annette Huber and Jörg Wildeshaus. Classical motivic polylogarithm according to Beilinson and Deligne. *Doc. Math.*, 3:27–133 (electronic), 1998.
[16] Kevin Hutchinson. A Bloch-Wigner complex for SL_2. *J. K-Theory*, 12(1):15–68, 2013.
[17] Rinat Kashaev, Vladimir Mangazeev, and Yuri Stroganov. Star-square and tetrahedron equations in the Baxter-Bazhanov model. *Internat. J. Modern Phys. A*, 8(8):1399–1409, 1993.
[18] Frans Keune. On the structure of the K_2 of the ring of integers in a number field. In *Proceedings of Research Symposium on K-Theory and its Applications (Ibadan, 1987)*, volume 2, pages 625–645, 1989.
[19] Serge Lang. *Fundamentals of Diophantine geometry*. Springer-Verlag, New York, 1983.
[20] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.
[21] Alexander Merkur′ev and Andrei Suslin. The group K_3 for a field. *Izv. Akad. Nauk SSSR Ser. Mat.*, 54(3):522–545, 1990.
[22] James Milne. *Étale cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1980.
[23] Hitoshi Murakami and Jun Murakami. The colored Jones polynomials and the simplicial volume of a knot. *Acta Math.*, 186(1):85–104, 2001.
[24] Werner Nahm. Conformal field theory and torsion elements of the Bloch group. In *Frontiers in number theory, physics, and geometry. II*, pages 67–132. Springer, Berlin, 2007.
[25] Walter Neumann. Extended Bloch group and the Cheeger-Chern-Simons class. *Geom. Topol.*, 8:413–474 (electronic), 2004.
[26] Tomotada Ohtsuki. On the asymptotic expansion of the Kashaev invariant of the 5_2 knot. *Quantum Topol.*, 7(4):669–735, 2016.
[27] Walter Parry and Chih-Han Sah. Third homology of $\mathrm{SL}(2, \mathbf{R})$ made discrete. *J. Pure Appl. Algebra*, 30(2):181–209, 1983.
[28] Christophe Soulé. K-théorie des anneaux d'entiers de corps de nombres et cohomologie étale. *Invent. Math.*, 55(3):251–295, 1979.
[29] Christophe Soulé. Éléments cyclotomiques en K-théorie. *Astérisque*, (147-148):225–257, 344, 1987. Journées arithmétiques de Besançon (Besançon, 1985).
[30] Andrei Suslin. Torsion in K_2 of fields. *K-Theory*, 1(1):5–29, 1987.
[31] Andrei Suslin. K_3 of a field, and the Bloch group. *Trudy Mat. Inst. Steklov.*, 183:180–199, 229, 1990. Translated in Proc. Steklov Inst. Math. **1**991, no. 4, 217–239, Galois theory, rings, algebraic groups and their applications (Russian).
[32] John Tate. Relations between K_2 and Galois cohomology. *Invent. Math.*, 36:257–274, 1976.
[33] Michael Terhoeven. Dilogarithm identities, fusion rules and structure constants of CFTs. *Modern Phys. Lett. A*, 9(2):133–141, 1994.
[34] Masha Vlasenko and Sander Zwegers. Nahm's conjecture: asymptotic computations and counterexamples. *Commun. Number Theory Phys.*, 5(3):617–642, 2011.

[35] Vladimir Voevodsky. On motivic cohomology with $\mathbf{Z}/l$-coefficients. *Ann. of Math. (2)*, 174(1):401–438, 2011.

[36] Charles Weibel. Algebraic K-theory of rings of integers in local and global fields. In *Handbook of K-theory. Vol. 1, 2*, pages 139–190. Springer, Berlin, 2005.

[37] Don Zagier. Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields. In *Arithmetic algebraic geometry (Texel, 1989)*, volume 89 of *Progr. Math.*, pages 391–430. Birkhäuser Boston, Boston, MA, 1991.

[38] Don Zagier. The dilogarithm function. In *Frontiers in number theory, physics, and geometry. II*, pages 3–65. Springer, Berlin, 2007.

[39] Don Zagier. Quantum modular forms. In *Quanta of maths*, volume 11 of *Clay Math. Proc.*, pages 659–675. Amer. Math. Soc., Providence, RI, 2010.

[40] Christian Zickert. The extended Bloch group and algebraic K-theory. *J. Reine Angew. Math.*, 704:21–54, 2015.

Department of Mathematics, University of Chicago, Chicago, IL 60637, USA
http://math.uchicago.edu/~fcale
E-mail address: fcale@math.uchicago.edu

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA
http://www.math.gatech.edu/~stavros
E-mail address: stavros@math.gatech.edu

Max Planck Institute for Mathematics, 53111 Bonn, Germany
http://people.mpim-bonn.mpg.de/zagier
E-mail address: dbz@mpim-bonn.mpg.de